

## **Effectiveness and Durability of Digital Preservation and Curation Services**

Ithaka S+R proposes an 18-month research project to conduct case studies that will examine and assess how digital preservation and curation systems (DPCS) are developed, deployed, and sustained. We will analyze the business approaches of community-based and commercial initiatives, offer lessons learned through case studies, and propose alternative sustainability models for long-term maintenance and development. The study aims not only to further increase our understanding of sustainability principles but also to design actionable recommendations to help the sector implement research findings. This **National Digital Infrastructures and Initiatives** project will culminate in a series of in-person and/or virtual forums to share the findings and engage the community in designing actionable recommendations. The ultimate goal is to apply the empirical data gained through research to guide the design choices of future systems and inform the users' selections and grantmakers' policies. A guiding principle for the proposed research is exploring ways to engage the community to move from the articulation of problems to the design of strategies addressing challenges to long-term success. This process will take into consideration the varying needs and resources of cultural institutions that serve user communities with diverse geographic, cultural, and socioeconomic backgrounds.

Ithaka S+R's team will be led by Oya Y. Rieger, Senior Strategist for Ithaka S+R. Roger Schonfeld, Director of Libraries, Scholarly Communication, and Museums at Ithaka S+R, will serve as an advisor to the project, working closely with Rieger. Given Ithaka S+R's broad experience with issues related to collaboration, sustainability, funding models, funders' practices, and OS and commercial software for heritage institutions, we have the necessary skills and expertise to undertake this work.<sup>1</sup> We will form and work closely with an Advisory Board, which will include representation from the leadership of various cultural heritage organizations, DPCS development projects, and sustainability/business experts. At the conclusion of the project, we will convene a series of in-person and/or virtual forums to share findings with key audience groups and design actionable recommendations for implementing best practices. This project will produce interim and final outputs including blog posts, webinars, and a final report.

# **1. Statement of National Need**

## ***1.1 National Need***

Digital preservation and curation process involves a series of technical, intellectual, and managerial activities in support of creating, acquiring, appraising, repurposing, accessing, and preserving digitized and born-digital information assets.<sup>2</sup> For the purposes of the proposed study, we define DPCS as the tools and services used by cultural heritage organizations to undertake digital preservation and curation work. Examples include Archivematica, Cell Signalling Alliance (defunct), Chronopolis, ContentDM, Digital Preservation Network (defunct), DuraSpace, Dryad, EPrints, figshare, Greenstone, HathiTrust,

---

<sup>1</sup> Ithaka S+R conducts research and provides convening and advisory services. Please see our work on [funding and sustainability](#) and [cross-institutional collaboration and open source software](#). See also Schonfeld's [relevant pieces](#) at the Scholarly Kitchen.

<sup>2</sup> Digital curation is a broad concept that involves creating, acquiring, appraising, repurposing, accessing, and preserving digitized and born-digital information assets. Preservation is considered one of the important life-cycle events of curation to ensure enduring access to digital content. In this study, we are particularly interested in systems that facilitate the management and maintenance of digital objects to ensure the authenticity, accuracy, and functionality of content over time in the face of technological and administrative changes. For the purposes of our study, we will focus on preservation and curation systems with multiple/integrated functionality, rather than individual tools designed for specific purposes, such as ingesting content or supporting fixity checks.

Invenio, LOCKSS/CLOCKSS, Mukurtu, National Public Library of America (DPLA), Portico, Rosetta, Samvera/Hyku, and Zenodo.

During the last two decades, cultural heritage institutions have participated in the development of community-based<sup>3</sup> DPCS, such as DuraSpace, Samvera/Hyku, and Mukurtu, to foster cooperation among cultural institutions that share a common vision of safeguarding and increasing access to our cultural heritage. These investments have led to a scholarly communications sector that is brimming with small, independent start-ups and membership organizations. Libraries, archives, and museums have a growing dependency on these digital platforms to support the curation, discovery, and long-term management of digital content. Yet, some of the systems and tools libraries use for curation have been shown to have substantial sustainability challenges. As we witness the recent organizational challenges faced by services such as Digital Preservation Network and Digital Public Library of America, we are reminded of the importance of creating sustainable services with clear visions and value-propositions that are aligned with the marketplace and available resources.

Our cultural, historic, and scientific heritage is increasingly being produced and shared in digital forms. The COVID-19 pandemic and its potential operational and financial implications for cultural heritage organizations further underscores the importance of deploying operationally and financially durable and effective systems. Although there are a variety of sustainability models and some helpful guidebooks to promote best practices (see Literature Review), understanding the complexities of implementing these practices requires further research.

## ***1.2 Literature Review***

In preparation for this proposal, we conducted an initial review of key initiatives that explore both technical and sociotechnical aspects of the sustainability challenge. This preliminary research has revealed a number of themes that provide a foundation for what we propose to do.<sup>4</sup>

A desire for “innovation” in new digital library services during the early 2000s led to the development of many novel tools without sufficient emphasis on long-term business planning. Sustainability, in terms of long-term business planning, includes anticipating issues such as governance, covering operating costs, and continuing maintenance and innovation. However, DPCS rarely engage in this type of long-term strategic planning, instead depending on soft-money, struggling with limited resources, and underestimating the importance of assessing the marketplace for future financial stability and growth.

Furthermore, there is a lack of programs to provide guidance, mentorship, and training to instill managerial, financial, and marketing skills for DPCS leaders and staff. Most community-based DPCS are led by individuals who are technology-savvy but lack experience or training in developing and maintaining business operations. This sometimes leads to privileging innovation as the driving factor

---

<sup>3</sup> Community-based implies the engagement of mission-driven institutions (such as libraries, archives, museums, and other non-for-profit services and advocacy organizations) in conceptualization, design, maintenance and development of services and systems. While we differentiate between community-based and commercial, we recognize that in practice there are hybrid approaches.

<sup>4</sup> There are several studies that aim to strengthen planning, promotion, and assessment of system and service sustainability. Examples include: [National Digital Platform at Three, It Takes a Village: Open Source Sustainability](#), [Community Cultivation: A Field Guide](#), [Community Health Analytics Open Source Software](#), [Understanding What Constitutes a Vibrant Open Source Community](#), [Why Are So Many Scholarly Communication Infrastructure Providers Running a Red Queen’s Race?](#), [Sustaining the Open Sector: A Brief Look Back](#), [Invest in Open Infrastructure](#), [Mapping the Scholarly Communication Landscape](#), [Restructuring Library Collaboration](#), [Architecting Sustainable Futures: Exploring Funding Models in Community-Based Archives](#), [Preserving Digital Objects With Restricted Resources](#).

behind the service or product, at the cost of user' needs, resources, and actual workflows. Innovation, moreover, tends to be divorced from the needs of cultural institutions, which operate on limited resources and serve diverse user communities.

Collaboration, even among mission-driven OS communities, is difficult to establish and maintain due to competing local priorities, limited resources, and differing branding needs. Additionally, DPCS leaders may lack an adequate understanding of the role of a governance system in providing overall strategic guidance in various stages, from an initial R&D phase to production, to mergers, and to sunset. Likewise, they may underutilize assessment methods that would increase internal risk management and external accountability to various stakeholders to monitor and control development and deployment processes. Open source solutions have an especially precarious balance to maintain between community governance and strategic agility. Yet open source solutions compete in the same marketplace, where the pace of innovation is relentless.

To have a competitive edge in the library and scholarly communications sector, stand-alone applications with comparatively static product definitions are increasingly giving way to integrated solutions, with fast-moving boundaries and a growing emphasis on data and analytics. This raises two complexities. First, community-based initiatives must strike a balance between agile governance and inclusive governance. Second, not-for-profit initiatives compete with commercial enterprises in the marketplace. While a strong commitment to mission is a vital underpinning for any not-for-profit, recognizing the marketplace dynamics, implementing sound and transparent business processes, and being willing to operate within the reality of their constraints are equally essential for long-term success. Business planning should include realistic risk assessments and should happen transparently, rather than being buried in budgeting processes. For instance, it is necessary to expand the community's understanding of the value of mergers and other organizational strategies that are necessary to maintain an effective and efficient administrative, fiscal, and social infrastructure.

## **2. Project Design**

### ***2.1 Project Objectives***

Given the substantial related work discussed above in the literature review, we have identified two key issues that we plan to address in the current project. First, there has been limited examination of how community-based initiatives develop sufficient capital and agility to thrive in sectors that include for-profit competitors. Second, there has been very little sustained engagement with the funder community to help it consider how altered programmatic guidelines or investment strategies might improve outcomes.

Ithaka S+R proposes a project that will investigate how DPCS are designed, implemented, and assessed in order to identify key success factors and impediments related to the long-term development, deployment, and sustainability of these systems. The ultimate goal is to support the design choices of future systems, offer alternative sustainability models, and inform the users' selections and grantmakers' policies by taking into consideration the diverse needs and requirements of cultural institutions. We define the user group as individuals (customers) within cultural heritage institutions who evaluate, select, acquire, deploy, and manage systems (technology services and applications) for curating and preserving various digital content.

This project is aligned with the National Digital Infrastructures and Initiative's directive to enhance the sustainability and accessibility of digital content and collections to provide long-term impact to and value for diverse and evolving user communities. One of our goals is to explore novel approaches for

engaging the related communities in the interpretation and uptake of our research findings so that the outcomes of this project will be actionable, not limited to creating new research insights.

## ***2.2 Target Audience and Engagement Plan***

The outputs of this research will have the most value to the following audiences:

- Leaders of existing DPCS, defined to include project directors, other appropriate staff leaders, and governing board members, as well as those that might be interested in creating DPCS; and
- Organizations responsible for providing capital funding in support of DPCS, defined to include major relevant grant-making organizations as well as emergent groups of libraries and others interested in capitalizing this work.

Our Advisory Board<sup>5</sup> is intentionally designed to represent them, helping to ensure that our work speaks to their experiences and needs. The members will be instrumental in designing our outreach, engagement, communication, and implementation plans, in addition to reviewing our draft findings to ensure they are relevant and accessible to the target audience.

At the conclusion of the project, we will convene a series of forums to share the findings and engage each of the two primary audiences in designing actionable recommendations to leverage research findings (see National Impact). For DPCS leaders, we will focus on how to apply the findings of the project on key success factors that can help to inform their work going forward. For capital funders, we will focus on how to apply the outcomes of this research to support sustainability planning at the point of capital funding.

We have also identified an important secondary stakeholder group: DPCS customers, defined to include individuals in a library, archives, or museum, with direct responsibility for selecting DPCS for their organizations' use. We plan to engage this audience through tailored outputs, which may include blogposts and webinars, in addition to the final report.

## ***2.3 Research Questions***

Our examination of the development, implementation, and sustainability of digital preservation and curation systems will be guided by the following research questions:

1. Business planning includes the utilization of long-term strategic thinking, governance models, usability studies, market research, risk assessment and mitigation, and agility and flexibility to navigate and adapt to change in time of operational challenges, changing priorities, evolving leadership, and shifting funding streams. **What business approaches are used to plan and implement community-based DPCS and how do they compare and relate to strategies followed by commercial entities offering similar products?**
2. DPCS need to strike a balance between agility, inclusivity, and the diverse needs of users. Our study will examine what it means for DPCS to be inclusive and accessible. **How are related requirements assessed and incorporated based on a variety of cultural heritage institutions that deploy these tools or rely on them for enduring access?**

## ***2.4 Methodological Approach***

---

<sup>5</sup> Confirmed advisors include Mike Furlough, Executive Director, HathiTrust; Carol A. Mandel Distinguished Presidential Fellow, Dean Emerita, New York University (NYU) Libraries; Robert Miller, Chief Executive Officer, Lyris; Veronica Reyes-Escuerdo, Katherlyne B. Willock Head of Special Collections, University of Arizona Libraries; and Katherine Skinner, Executive Director, Educopia Institute.

## 2.4.1 Environmental Scan

We will begin by reviewing reports, project wikis, and social media about technical, managerial, and socioeconomic aspects of DPCS with the particular goal of enabling us to develop a practical taxonomy of functions offered by DPCS. We will use the taxonomy to identify potential case study candidates. For the sake of efficiency, we will heavily rely on the existing inventories, outcomes of the existing environmental scans, and expertise of our Advisory Board.<sup>6</sup>

## 2.4.2 Case Studies

The choices that impact the sustainability of DPCS are made in complex settings, with organizational, technical, and collaborative and other social dynamics informing one another with contingency. As a result, we have selected a case study method to allow us to document the rich complexity of each of the selected DPCS.<sup>7</sup> The case study approach is particularly useful in conducting an in-depth and holistic investigation of complex issues without a predetermined hypothesis.<sup>8</sup> The research approach lends itself to capturing contextual information on more explanatory 'how', 'what,' and 'why' questions in everyday settings.

We anticipate conducting 8 case studies. Each case study will involve a close examination of an individual DPCS selected from the list of potential case study candidates identified during the Environmental Scan. Although we will primarily focus on OS and community-driven systems, we will include at least two commercial products (e.g. Rosetta, Arkivum, Archivemata, etc.) in our case studies to enable a more nuanced understanding of the range of tools available, as well as any significant comparisons that emerge across these types of products. While we differentiate between community-based and commercial applications, our approach recognizes that in practice some institutions adopt hybrid deployment approaches, such as combining tools from different vendors and communities (thus requiring tools to have a certain level of interoperability and modularity).

For each case study, we will triangulate several sources of data, including:<sup>9</sup>

- Information from the recent sustainability studies<sup>10</sup> and various web and social media resources about technical, managerial, and socioeconomic aspects of the systems selected for case studies;
- Interviews with an array of leaders and other contributors and observers;

---

<sup>6</sup> As the [Mapping the Scholarly Communication Landscape - 2019 Census](#) revealed, there is not a standardized taxonomy for the various functions performed by systems that comprise the scholarly communication ecology. There are a number of registries to help cultural heritage organizations identify curation and preservation systems and tools. Examples include [Community Owned digital Preservation Tool Registry](#), [AV Preserve Tools List](#), [Digital Curation Centre Preservation and Curation Tools and Services](#), [The Digital Cultural Heritage Roadmap for Preservation Tools Registry](#), [Inventory of Free/Libre Open-Source Software for Cultural Heritage](#), [Library of Congress NDIIPP Tools Showcase](#).

<sup>7</sup> Ithaka S+R has deep experience in conducting case-study based research. Since 2006, we have explored the sustainability requirements of digital resources and systems, investigating both the factors that lead to success and the reasons behind setbacks and failures through numerous case studies. In 2018, we conducted a series of museum case studies to assess how the museums think about diversity, equity, accessibility, and inclusion in relation to collections, programs, community engagement, partnerships, and board development. Other examples of Ithaka S+R case studies can be found at: [https://sr.ithaka.org/publications/?fwp\\_publication\\_types=case-study](https://sr.ithaka.org/publications/?fwp_publication_types=case-study)

<sup>8</sup> Robert K. Yin. *Case Study Research and Applications*, Sage, 2018; Bill Gillham, *Case Study Research Methods*, Bloomsbury Publishing, 2000; John W. Cresswell. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. Thousand Oaks, CA: Sage, 2014.

<sup>9</sup> Triangulation in research involves combining different data gathering methods to provide a more detailed and balanced assessment, leading to richer explanations.

<sup>10</sup> We intend to reuse available data from related sustainability studies mentioned in Literature Review to avoid replicating.

- Direct gathering of financial statements, funding proposals and awards, and similar materials that elucidate sustainability and reinvestment; and
- Interview and email exchanges with users and non-users for each case study, focusing on their experiences and perceptions with the DPCS examined.<sup>11</sup>

For interviews, we will develop a semi-structured guide with a blend of closed- and open-ended questions. This approach is strongly preferable to a static straightforward question and answer format because it allows for greater discussion and a more nuanced understanding of the interviewee's perspective. Our interviews will address our Research Questions, examining from a variety of perspectives how and why decisions were made, taking into consideration historical context. Given current circumstances due to the COVID-19 pandemic, we have adjusted our approach and budget to conduct virtual interviews. Ithaka S+R has extensive experience in conducting research studies based on virtual interactions. Each interview will be recorded and transcribed manually (not verbatim but capturing all salient information).

### 2.4.3. Data Analysis

To analyze the information gathered, we will use a qualitative content/data approach.<sup>12</sup> The thematic analysis will start with a close examination of the verbatim interview transcripts to identify themes and assign codes (labels with few words or short phrases) to evoke the relevant features of the data. We will condense and organize codes into categories informed by the research questions and environmental scan. We will use NVIVO, a qualitative analysis software, to organize codes and categories manually due to the interpretive nature of the study.<sup>13</sup> The analysis will result in case studies that profile each DPCS individually. We will also produce a final report that synthesizes data across each of the individual case studies, offers lessons learned, and proposes alternative sustainability models for long-term maintenance and development.

### 2.4.3. Moving Findings into Practice

The study aims not only to further increase our understanding of sustainability principles but also design actionable recommendations to help the sector implement research findings. To this end, we will convene a series of in-person and/or virtual forums to share the findings with the members of the relevant DPCS and higher education community, funders, and policy makers to facilitate community-based discussions of the research findings, implications, and alternative models. Such deliberations will promote the findings and recommendations as well as investigate the potential impediments and enablers behind putting recommendations into practice.

For DPCS leaders, we will host up to two in-person convenings of 10-12 individuals at ITHAKA's offices in New York and/or in conjunction with a conference (such as CNI), focusing on how to apply the findings of the project on key success factors that can help to inform their work going forward<sup>14</sup>. We are especially interested in engaging the forum participants in considering financial models for community-based initiatives to develop sufficient capital and agility to thrive in sectors that include

---

<sup>11</sup> The term non-users denote groups and individuals who choose not to use the service, stopped using the service, preferred a comparable service, or are unaware of the service's existence. We will identify users and non-users for interviews through social media, Advisory Board members, and Ithaka S+R's wide network of cultural heritage partners.

<sup>12</sup> Greg Guest, Kathleen M. MacQueen, and Emily E. Namey. *Applied Thematic Analysis*. Thousand Oaks, CA: Sage, 2012.

<sup>13</sup> NVivo is a qualitative research tool used to organize and analyze findings from interviews, focus groups, etc.

<sup>14</sup> We anticipate that circumstances relating to COVID-19 will have diminished by the point in this project where we will seek to implement these convenings. If, however, in-person convenings are not possible, we will adjust our approach and budget to accommodate virtual convenings.

for-profit competitors. The meeting will provide an opportunity to test new models for governance, usability studies, market research, and risk assessment and mitigation.

For capital funders, we will host an in-person workshop for 10-12 individuals, including IMLS staff and other invited funders, in New York or Washington DC. Based on our initial research, it appears that there has been limited sustained engagement with the funder community to alter programmatic guidelines or investment strategies for improved outcomes. The workshop will result in actionable recommendations on how to apply the outcomes of this research to support the long-term success of initiatives that start with grants and gifts.

A guiding principle of this project will be to educate participants in these convenings about the importance of taking into consideration the varied capabilities and stewardship responsibilities of cultural heritage organizations when developing DPCS. All types of cultural heritage organizations face increasing competition for dwindling resources; the financial consequences of COVID-19 will further aggravate the resource shortages. The convening will help us build consensus around the need for designing inclusive, accessible, and sustainable DPCS, by allowing us to engage closely with key stakeholders. This process will ensure that the guidance this project offers will be foremost actionable, rather than merely aspirational.

## **2.5 Timeline**

### *Phase 1: Environmental Scan and Research Initiation (September 2020 - November 2020)*

- Project kickoff.
- Review reports, project wikis, and social media about technical, managerial, and socioeconomic aspects of DPCS.
- Hold a virtual meeting with selected leads of the previous sustainability studies (see Literature Review) to get their input on how to build on existing insights and data.
- Develop brief taxonomy of functions offered by DPCS based on the landscape analysis to enable us to identify initial case study candidates.<sup>15</sup>
- Convene Advisory Board to review our research methodology and desired outcomes.
- Select a subset of DPCS (in collaboration with the Advisory Board) to be explored in greater depth through case studies.
- Plan and launch a strategic communications plan, which will outline our plan to engage different stakeholders and share information through the duration of the project through blog posts, webinars, interviews, etc.

### *Phase 2: Conduct Case Studies and Initial Analysis (December 2020 - May 2021)*

- Schedule case studies.
- Gather and review information from various web and social media resources about technical, managerial, and socioeconomic aspects of the systems selected for case studies.
- Conduct interviews with leadership, governing boards, and staff involved in developing and maintaining each DPCS selected as case study.
- Seek input from users and non-users of each DPCS selected as case study through interviews and email.

---

<sup>15</sup> Initial case study candidates include: Archivematica, Cell Signaling Alliance (defunct), Chronopolis, ContentDM, DPN (defunct), DuraSpace, Dryad, EPrints, figshare, Greenstone, HathiTrust, Invenio, LOCKSS, Mukurtu, Portico, Rosetta, Samvera/Hyku, Zenodo. Although our focus will be on open source and community-build systems, we'll include two commercial products in our case studies to allow comparison.

- Analyze data gathered.
- Draft individual case studies for each DPCS.<sup>16</sup>
- Share first set of case study reports with Advisory Board members for review, feedback, and discussion.

*Phase 3: Conclude Case Study Analysis (June 2021 - September 2021)*

- Share the second set of case study reports with Advisory Board members for review, feedback, and discussion.
- Share individual case studies with participants for feedback and validation of findings.
- Synthesize case studies and draft final report.
- Review the final report with the Advisory Board.
- Plan logistics for in-person convenings (for DPCS leaders and capital funders). Develop agenda and invitee list. Event coordination and logistics. Send invitations.

*Phase 4: October 2021 - February 2022 (Dissemination, Convening, and Final Report)*

- Hold in-person convenings for DPCS leaders and capital funders to share project findings and design strategies to implement recommendations (See National Impact).
- Plan and implement webinar for DPCS users (cultural heritage organizations).
- Publish and disseminate a final report and case studies (see Communications).
  - Present findings broadly in related meetings, conferences, etc.
  - Publish final report, including the final report and case studies.

## ***2.6 Indicators of Success and Risk Mitigation***

The key indicator of success for our exploratory research in service to practice project will be our ability to build on existing research to move the community forward by designing strategies (actionable recommendations) for implementing best practices pertaining to access and sustainability. We will assess the short-term success by our ability to engage different communities into this process by taking into consideration different needs and resources of cultural institutions that serve user communities with diverse geographic, cultural, and socioeconomic backgrounds. In the long-term, the project's success will be evaluated by assessing how contributions of the outcomes in changing the ways that DPCS are capitalized, the business models they use and the governance models and practices they utilize, to improve sustainability.

The only risk we are aware of at this point is running into challenges in recruiting project participants (interviewees) associated with each case study. Given Ithaca S+R's broad network and that of the Advisory Board, we are confident that we will succeed in engaging sufficient stakeholders to produce high-quality data and evidence. Our Advisory Board will periodically evaluate our accomplishments and provide feedback in making the necessary adjustments to accomplish the project goals we have set for this study. Furthermore, we will carefully differentiate our study from other studies so that invited participants understand the added value of our project and do not perceive it as reduplicating existing studies.

# **3. National Impact**

## ***3.1 Audiences, Outcomes, and Outputs***

---

<sup>16</sup> We anticipate each case study to be structured similar to this example: [https://sr.ithaca.org/wp-content/uploads/2015/08/SCA\\_IthacaSR\\_CaseStudies\\_SEP\\_2011.pdf](https://sr.ithaca.org/wp-content/uploads/2015/08/SCA_IthacaSR_CaseStudies_SEP_2011.pdf)



The primary audiences for this project are leaders of existing DPCS, those that might be interested in creating DPCS, and organizations responsible for providing capital funding in support of DPCS. Additionally, the secondary audience for this project is DPCS customers, including individuals in a library, archives, or museum with direct responsibility for selecting DPCS for their organizations' use.

The findings from this project will contribute to realistic and sustainable approaches for the stewardship of new and complex content types and emerging digital formats. Our ultimate goal is to produce empirical data to guide the design of future DPCS systems and the organizations that support them, ensuring that these services and tools remain sustainable and accessible to users who depend on them. While we recognize that influencing behavior is a long-term goal (longer than the period of this grant), this research represents an essential first-step towards this larger goal. Because the proposed research relates to a pressing area that requires the development of alternative strategies in a timely manner, we will engage the key stakeholders at the project onset and develop a communication strategy to engage them throughout the project.

To support these outcomes, we will produce the following deliverables:

- A strategic communications plan, including a series of interim webinars, blog posts, presentations, and interviews which will take place throughout the project targeting audiences who design, develop, select, use, assess, and fund DPCS to keep them informed of how the initiative is evolving, share interim outcomes, and seek feedback.
- A final report written for a general audience at the conclusion of the project accompanied with a series of virtual forums to broadly engage the primary audiences in strategies to effectively leverage research insights.
- Series of community-based in-person and/or virtual forums to present findings as well as recommendations customized for each set of participants that will generate discussion and potentially adoption:
  - Up to two in-person convenings of 10-12 DPCS leaders, focused on how to apply the findings of the project on key success factors that can help to inform their work going forward.
  - In-person convening of 10-12 funders focused on how to apply the outcomes of this research to support the long-term success of initiatives that start with grants and gifts and support sustainability planning at the point of capital funding.
  - Webinar for DPCS users (cultural heritage organizations), on how to apply the outcomes of this research to support as they select, deploy and manage DPCS for their institutions' use.

### ***3.2 Communications***

All publications that result from this project will be published on the Ithaka S+R website under an open access CC BY-NC 4.0 license. Each will be assigned an individual DOI and promoted widely through our email list and blog, through a variety of appropriate listservs, and via a number of social media channels. Through these techniques, Ithaka S+R's research outputs regularly secure extensive reach in several relevant sectors, including academic libraries and associated software/tool providers, and we expect to see a similar impact for this project. Members of the Advisory Board will play a role in ensuring the national dissemination of research outputs from this work.

Throughout the project, and especially as findings are developed, researchers will seek out opportunities to write blog posts and opinion pieces, to engage and promote research findings on social media, to speak at conferences, and to organize conference panels and webinars on the topics of the

research project. We will especially be targeting the following conferences: Coalition of Networked Information, Confederation of Open Access Repositories, Digital Library Federation, International Conference on Digital Preservation. As discussed above, the project will conclude with in-person and/or virtual forums to share project findings and design strategies to implement recommendations.

Effectiveness and Durability of Digital Preservation and Curation Services																		
Schedule of Completion	2020				2021				2021				2021				2022	
	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb
<b>1 Environmental Scan and Research Initiation</b>																		
1.1 Review of reports, project wikis, and social media	█	█																
1.2 Seek input from lead of previous related studies		█																
1.3 Develop a brief taxonomy of functions offered by DPCS			█															
1.4 Convene Advisory Board to review research methodology				█														
1.5 Select a subset of DPCS for case studies																		
1.6 Plan and launch a series of webinars and blog posts						█		█		█		█		█				
<b>2 Conduct Case Studies and Initial Analysis</b>	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb
2.1 Schedule case studies (in-person and virtual interviews)				█	█	█												
2.2 Gather information about DPCS selected for case study					█	█	█											
2.3 Conduct interviews with DPCS staff and governing boards					█	█	█	█	█									
2.4 Seek input from users and non-users of DPCS examined							█	█										
2.5 Analyze cases individually and draft case study reports							█	█	█									
2.6 Share 1st set of case studies with Advisory Board for feedback									█									
<b>3 Conclude Case Study Analysis</b>	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb
3.1 Share 2nd set of case studies with Advisory Board for feedback										█								
3.2 Share findings with case study participants for feedback										█								
3.3 Synthesize case studies and draft capstone report										█	█	█						
3.4 Review the capstone report with Advisory Board													█					
3.5 Plan in-person community discussion on alternative models											█	█	█	█				
3.6 Plan logistics for in-person convenings									█	█	█	█	█					
<b>4 Dissemination, Convening, and Final Report</b>	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb
4.1 Hold in-person convenings for DPCS leaders and funders														█	█			
4.2 Hold webinar for DPCS users (cultural heritage organizations)														█	█			
4.3 Present findings broadly in related meetings, conferences, etc.														█	█	█	█	
4.4 Publish the formal final report (8 case studies and synopsis)																█	█	



## DIGITAL PRODUCT FORM

### INTRODUCTION

The Institute of Museum and Library Services (IMLS) is committed to expanding public access to digital products that are created using federal funds. This includes (1) digitized and born-digital content, resources, or assets; (2) software; and (3) research data (see below for more specific examples). Excluded are preliminary analyses, drafts of papers, plans for future research, peer-review assessments, and communications with colleagues.

The digital products you create with IMLS funding require effective stewardship to protect and enhance their value, and they should be freely and readily available for use and reuse by libraries, archives, museums, and the public. 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

If you propose to create digital products in the course of your IMLS-funded project, you must first provide answers to the questions in **SECTION I: INTELLECTUAL PROPERTY RIGHTS AND PERMISSIONS**. Then consider which of the following types of digital products you will create in your project, and complete each section of the form that is applicable.

#### **SECTION II: DIGITAL CONTENT, RESOURCES, OR ASSETS**

Complete this section if your project will create digital content, resources, or assets. These include both digitized and born-digital products created by individuals, project teams, or through community gatherings during your project. Examples include, but are not limited to, still images, audio files, moving images, microfilm, object inventories, object catalogs, artworks, books, posters, curricula, field books, maps, notebooks, scientific labels, metadata schema, charts, tables, drawings, workflows, and teacher toolkits. Your project may involve making these materials available through public or access-controlled websites, kiosks, or live or recorded programs.

#### **SECTION III: SOFTWARE**

Complete this section if your project will create software, including any source code, algorithms, applications, and digital tools plus the accompanying documentation created by you during your project.

#### **SECTION IV: RESEARCH DATA**

Complete this section if your project will create research data, including recorded factual information and supporting documentation, commonly accepted as relevant to validating research findings and to supporting scholarly publications.

## **SECTION I: INTELLECTUAL PROPERTY RIGHTS AND PERMISSIONS**

**A.1** We expect applicants seeking federal funds for developing or creating digital products to release these files under open-source licenses to maximize access and promote reuse. What will be the intellectual property status of the digital products (i.e., digital content, resources, or assets; software; research data) you intend to create? What ownership rights will your organization assert over the files you intend to create, and what conditions will you impose on their access and use? Who will hold the copyright(s)? Explain and justify your licensing selections. Identify and explain the license under which you will release the files (e.g., a non-restrictive license such as BSD, GNU, MIT, Creative Commons licenses; RightsStatements.org statements). 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.

**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.

## **SECTION II: 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 the format(s) you will use.

**A.2** List the equipment, software, and supplies that you will use to create the digital 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, OBJ, DOC, PDF) you plan to use. If digitizing content, describe the quality standards (e.g., resolution, sampling rate, pixel dimensions) you will use for the files you will create.

### **Workflow and Asset Maintenance/Preservation**

**B.1** Describe your quality control plan. How will you monitor and evaluate your workflow and products?

**B.2** Describe your plan for preserving and maintaining digital assets during and after the award period. Your plan should 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).

## **Metadata**

**C.1** Describe how you will produce any and all technical, descriptive, administrative, or preservation metadata or linked data. Specify which standards or data models you will use for the metadata structure (e.g., RDF, BIBFRAME, 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).

### **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, delivery enabled by IIIF specifications).

**D.2.** Provide the name(s) and URL(s) (Universal Resource Locator), DOI (Digital Object Identifier), or other persistent identifier for any examples of previous digital content, resources, or assets your organization has created.



## **SECTION III: SOFTWARE**

### **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 or similar functions, and explain how the software you intend to create is different, and justify why those differences are significant and necessary.

### **Technical Information**

**B.1** List the programming languages, platforms, frameworks, 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), URL(s), and/or code repository locations for examples of any previous software your organization has created.

## Access and Use

**C.1** Describe how you will make the software and source code available to the public and/or its intended users.

**C.2** Identify where you will deposit the source code for the software you intend to develop:

Name of publicly accessible source code repository:

URL:

## SECTION IV: RESEARCH DATA

As part of the federal government's commitment to increase access to federally funded research data, Section IV represents the Data Management Plan (DMP) for research proposals and should reflect data management, dissemination, and preservation best practices in the applicant's area of research appropriate to the data that the project will generate.

**A.1** Identify the type(s) 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.

**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 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.

**A.4** What technical (hardware and/or software) requirements or dependencies would be necessary for understanding retrieving, displaying, processing, or otherwise reusing the data?

**A.5** 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?

**A.6** What is your plan for managing, disseminating, and preserving data after the completion of the award-funded project?

**A.7** Identify where you will deposit the data:

Name of repository:

URL:

**A.8** When and how frequently will you review this data management plan? How will the implementation be monitored?