The California Digital Library (CDL) and the Association of Research Libraries (ARL) request $668,048 for a 2-year National Leadership Grant in the Implementation category in alignment with Goal 3 of the NLG program for Libraries. Our proposal addresses the urgent needs of academic and research libraries to accelerate their abilities to respond to increasing requirements for sharing federally funded research data. Our project extends the current research data management digital infrastructure by expanding existing technology and infrastructure services, focusing on the critical infrastructure of the DMPTool and persistent identifier (PID) registries to meet funding agency requirements for creating machine-actionable data management plans (maDMP). The project also will support strategic planning for libraries as they implement policies, workflows, and technical solutions that build local capacity for effective data management, publication, and reuse. Stakeholder forums, consultations, and educational resources will inform this strategic planning. Paired with the project's technical work enabling maDMPs, the data management plan (DMP) will gain traction as a strategic communication and knowledge management tool which will aid data findability, interoperability, and reuse. This project will enable institutions of varying sizes and research budgets to meet new federal and foundation requirements for data management and sharing.

Project Justification: There is a growing list of federal requirements for data sharing and the creation of maDMPs. Recent significant federal policy updates regarding data sharing and reproducibility include the *CHIPS & Science Act* and the *Office of Science and Technology Policy (OSTP) Ensuring Free, Immediate, and Equitable Access to Federally Funded Research* policy guidance. The *CHIPS & Science Act* introduced new requirements for National Science Foundation (NSF) researchers to create maDMPs. The updated OSTP public access policy guidance increases the number of federal agencies with data sharing requirements, including IMLS and the National Endowment for the Humanities. The next step for the academic library community, which fundamentally supports the academic community, in supporting these new requirements, is to perform technical work to enable implementation and compliance by re-tooling existing research data management infrastructure. Additionally, libraries need to develop and enhance institutional research data services strategies that incorporate technology-agnostic tools and services that utilize the existing open infrastructure. Our proposal seizes the momentum of the rapidly changing federal guidelines by expanding library-created research data management digital infrastructure and pairing this implementation with intra-institutional research support coordination.

In January 2023, twelve of the largest U.S. federal agencies, including the National Science Foundation, the National Institutes of Health, and the Department of Energy, among others, joined with the White House to affirm their commitment to advance Open and Equitable Research. In doing so, these agencies made commitments to spark change and inspire open science engagement to advance the adoption of open, equitable, and secure science. Concomitantly, the Office of Science and Technology Policy (OSTP) released guidance in August 2022 to federal agencies to update or create new public access plans—including increasing requirements for research data management and sharing that result from federally funded research (White House Office of Science and Technology Policy 2022). While the specific implications for these policies and initiatives are forthcoming, the downstream impact on

institutions of higher education and, specifically, professionals who support and collaborate with researchers on their research data needs will be significant. While some research data management and sharing policies have been in place for a decade, the broader remit, as described by OSTP, will likely result in increased investments in infrastructure and services by institutions. Most recently, with the newly released National Institutes of Health Data Management and Sharing Policy, the Council on Governmental Relations (COGR) (2023) found that institutions will see an increased financial burden of approximately $1 million in meeting the policy requirements. This is an early analysis of one policy, but at the minimum, we will see eleven more updated or new federal agency policies by 2025. Thus, efficient and effective coordination of research data services at the institutional level is critical. In their 2021 publication Accelerating Public Access to Research Data, the Association of Public and Land Grant Universities (APLU) and the Association of American Universities (AAU) further highlighted the importance of campus coordination—indicating it to be one of the first and most significant steps for institutions to make towards meeting research data management, sharing, and curation requirements.

This proposal supports IMLS National Leadership Grant program Goal 3—Improve the ability of libraries and archives to provide broad access to and use of information and collections, emphasizing collaboration to avoid duplication and maximize reach; Objective 3.1—Advance digital inclusion … [by] "enhancing digital infrastructures, platforms, technologies, online services, connectivity, digital literacy, privacy, and security, as well as creating new processes and procedures needed to sustain a robust online environment". Research data and associated outputs are distributed in many and varied repositories and hence difficult to track and manage. DMPs are currently static unstructured narrative text documents that are not updated over time. The promise of the machine-actionable DMP is to be a vehicle for reporting on the intentions and outcomes of a research project that enable information exchange across relevant stakeholders and systems. Machine-actionable DMPs contain an inventory of key information about a project and its outputs with a change history that stakeholders can query for updated information about the project over its lifetime. By incorporating open persistent identifiers (PIDs) into DMPs and leveraging all DMP metadata for interoperability across infrastructures, libraries will be better equipped to track and manage their institutional research data products. By linking technical development with stakeholder engagement and coordination, CDL and ARL will make a substantial contribution to institutional infrastructure and workflows for providing broad access to information about research. This project will build a simplified approach for the entire community to track information about research project outputs from inception through publication and sharing results.

CDL and ARL have a history of collaboration on applying PIDs and DMPs to support data sharing and reuse. The two organizations were awarded an NSF grant (NSF 1945938 Implementing Effective Data Practices) in 2019, which resulted in published recommendations to support the broad adoption of PIDs and maDMPs (Chodacki et al., 2020). This proposed project will implement these recommendations by combining infrastructure and strategic planning to address the research community's needs as it continues evolving services to support institutional compliance for research data management and sharing policies.

IMLS was one of the original funders of the DMPTool with a 2012 grant to support training in research data management for librarians (LG-07-12-0505-12). CDL has continued to support and develop the application and remains committed to assuring the long-term sustainability of this resource. Since 2017, CDL has been awarded several innovation grants to explore ways to connect the research outputs described in a DMP to the broader ecosystem. This initial work led CDL to explore how to best capture information about associated outputs (preprints, datasets, protocols, instruments, samples, etc.) and map these resources to other related research outputs. Throughout this exploratory phase, CDL consulted with the broader scholarly infrastructure community and within the Research Data Alliance. In response, CDL built a prototype version to generate PIDs for DMPs minted via the DMPTool, and to record updates to the DMP that occur over time. Our proposal builds on this exploratory research to create a fully scalable workflow, equipping institutions to comply effectively with new data sharing mandates.

The vision and power of maDMPs have been well articulated in the scholarly community. For example, in the 2020 report, *Implementing Effective Data Practices: Stakeholder Recommendations for Collaborative Research Support*, ARL, CDL, and other research data management stakeholders, highlighted how DMPs contain essential information related to many institutional and research functions. Creating machine-actionable DMPs, would allow information to be shared programmatically and support integrations and information exchange in areas such as within the lab, among campus resource units, and with repositories and funding agencies (Chodacki et al., 2020). By tying technical development to institutionally-based workflows this project can accomplish that goal.

Background of the DMPTool

The DMPTool was developed from a grassroots effort of eight institutions beginning in January 2011. These institutions came together in response to new requirements by agencies such as the National Science Foundation (NSF) mandating that researchers submit plans to manage their research data. The DMPTool was developed to consolidate expertise and reduce costs in addressing data management needs at their respective institutions. This original mission still serves as the guiding principle of the DMPTool, as the application continues to be free and community-supported. Since its inception, CDL has maintained the tool, providing dedicated in-kind support of 2 FTEs for the project throughout the life of the service.

The tool's usage has increased significantly in the past several years due to increasing data management and sharing plan requirements, such as the new NIH Data Management and Sharing Policy (see Appendix Item 1). Many newer members of the DMPTool community come from smaller institutions with limited library staff expertise in research data management. With the rollout of the NIH policy (and other similar policies resulting from the 2022 OSTP Memo), these smaller research institutions are now faced with supporting their researchers in developing data management and sharing plans for the first time. This project will extend work on machine-actionable DMPs to allow data librarians and researchers they serve to quickly and effectively track and manage their institutional research data products. The project will expand capacity by building new workflows and tracking systems that generate optimal, structured

plans. Building these new workflows within the DMPTool will also provide a mechanism to communicate institutional-level guidance around local requirements and resources for research data management.

Current development of the DMPTool

For the past several years, DMPTool technical development has centered on transforming the DMP into a critical component of networked research data management. In 2017, CDL was awarded a National Science Foundation EAGER innovation grant to explore ways of mapping research project outputs as described in a data management plan to the broader ecosystem (NSF 1745675). This initial work led CDL to explore how to best capture information about associated outputs (preprints, datasets, protocols, instruments, samples, etc.) and map these resources to other related research outputs. After several years of consultation with the broader scholarly infrastructure community and within the Research Data Alliance (RDA), a standardized, structured metadata application profile was developed for communicating information about these research project outputs in a machine-actionable form. Following the development of the RDA Common Standard for DMPs, CDL partnered with DataCite to support the creation of a unique identifier for DMPs, the DMP-ID. Building on the release of DMP-IDs, version 1 of the maDMP feature set was released in April of 2021, featuring the generation of DMP-IDs for DMPs.

Workplan: This work plan includes a description of the new features of the DMPTool being built as part of this project and a project plan for implementing the features utilizing machine-actionable DMP workflows in institutional settings.

DMPTool Technical Development

The proposal aims to extend work on maDMPs by focusing on developing three key feature areas. These features are currently in the prototyping stage; the project requires further funding to support their full development and enable user feedback and pilot testing. To ensure that the features align with community needs, they will undergo multiple rounds of testing and iterative design, with feedback gathered from the pilot institutions.

1. New workflows for externally created DMPs: Most DMPs are not created with the DMPTool; however, there is still a need to structure these DMPs to generate DMP-IDs for DMPs created outside of the DMPTool. With this new feature, users will upload PDFs of existing DMPs and enter basic metadata about the project. Connections to funder APIs will automatically transfer relevant metadata to avoid duplication of data entry. Uploaded DMPs will get a DMP-ID. With these DMP-IDs, researchers and administrators can connect outputs related to their project, ultimately allowing for notifications and compliance reporting. The Project Manager (ARL based), funded by this proposal, will coordinate user testing, gather feedback acquired during the upload process, and streamline this feedback for iterative improvements of workflows for uploading these documents.

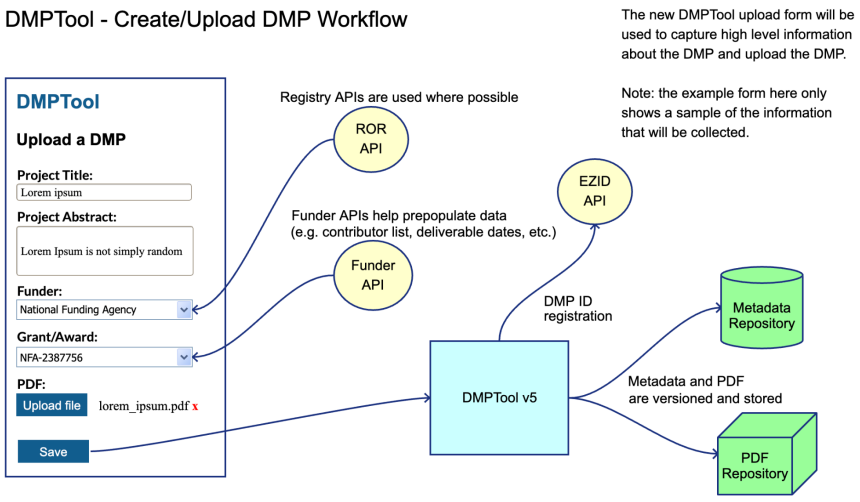

Figure 1: New workflow for structuring, generating DMP-IDs, and adding PIDs to DMPs created outside the DMPTool.

2. Automated updates to the DMP: The new system utilizing maDMPs will facilitate updating information associated with DMP-IDs via API-based integrations with external funders and research systems. For example, CDL is working with the NSF and NIH awards APIs to harvest funding and project information for inclusion in the DMP. CDL is also exploring integrations with the OpenAIRE Research Graph and DataCite Commons, which have robust APIs providing metadata and links between research outputs, organizations, funders, and datasets. Systems integration is a key feature of the machine-actionable DMP. It supports our goal of reducing the burden on researchers by generating updates to a plan and facilitates integration with systems and groups that support research. The Project Manager, funded by this proposal, will survey existing API services and coordinate user testing to iterate on what information the DMPTool consumes about research outputs from these external sources.

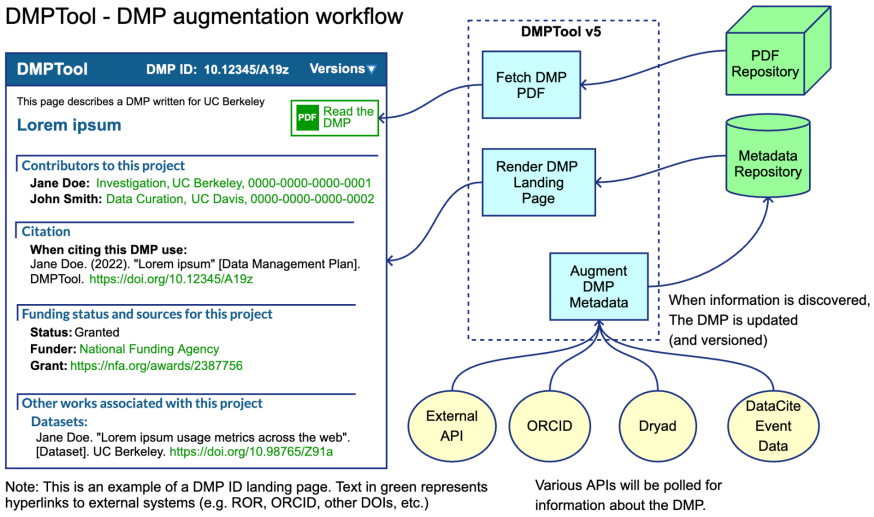

Figure 2: Sample workflow of programmatically augmenting DMPs with information pertaining to research outputs via API connections to openly available external systems.

3. **Administrator dashboards (**available to all DMPTool administrative users) will offer summary reports of DMPs, including a view of the current project status and any associated project outputs, including data repositories utilized. In addition, the dashboard will include reporting and tracking features designed for institutions seeking to document compliance with data-sharing mandates. The Project Manager will coordinate user testing for feedback on iterative dashboard design and what features are required to optimize the utility of the dashboards for administrators.

Figure 3: Iterative technical development will allow new features to continually improve and evolve based on user feedback, resulting in a more robust and user-friendly solution.

Pilot Partners Test Technical Enhancements

By working directly with four institutions on campus implementations of machine-actionable DMPs, we will have feedback to iterate on the technical development of new features and workflows to maximize the capacity of institutions to support researchers with data sharing; and to track the research outputs emerging from their institutions. By selecting a diverse representation of institutions in the pilot and communicating incremental findings regularly with the broader community, project findings will be reusable and scalable.

Project pilot partners will be asked to test new maDMP features as they are added to the DMPTool. We expect pilot partners to involve multiple stakeholders at their institution in all aspects of the program, including requirements gathering, testing, feedback, and implementation of maDMP features and workflows. All information acquired from pilot partners will be fed back to the CDL Project Developer and used to inform subsequent iterations. Features built in this iterative process will benefit from hands-on implementation in institutional workflows and will be available to all DMPTool partner organizations.

The DMPTool is intentionally community-maintained and community-informed in its technical development and functionality. When the DMPTool was launched in 2011, CDL established an advisory board to obtain feedback from the library community and others to guide the development and prioritize features. Building upon this successful model for community engagement, the project team proposes a multi-phased engagement and feedback workflow that prioritizes engagement with a diverse set of institutions.

A Project Manager at ARL will oversee all aspects of community recruitment, engagement, and outreach for the pilot. They will also oversee the project timeline and serve as the primary point of contact for Pilot Partners. The Project Developer at CDL will design, build and test all new features within the DMPTool open-source codebase.

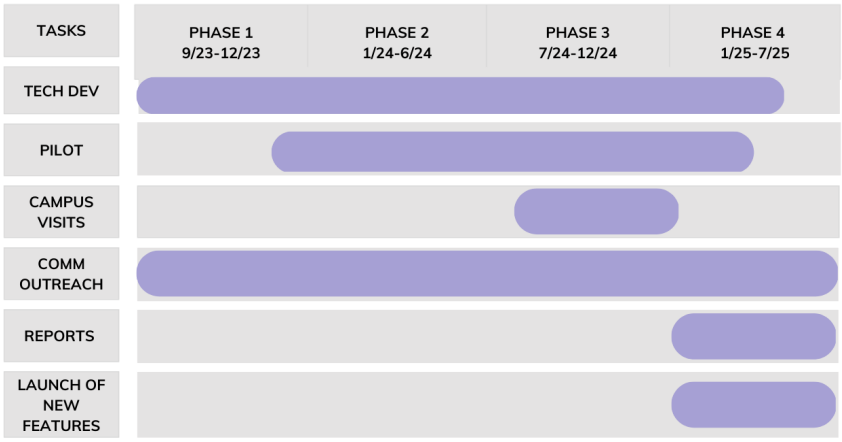

Figure 4: The project work plan is divided into four phases and includes the iterative technical development and pilot testing of features and implementing maDMP workflows.

Phase 1 August- December 2023 Recruitment and Initial Wireframing

- Develop criteria: Develop a recruitment process to enlist four institutions to participate in the pilot project. Develop criteria to ensure a representative group of institutions (balancing expertise and diversity) to participate in the pilot program. See Appendix Item 2 for a draft questionnaire.
- Recruit institutions: To ensure a diverse representation of institutions (including non-R1 academic organizations) and individuals, the project team will leverage the wide communication networks of the DMPTool and ARL to promote participation in the institutional pilots. We will also leverage professional organizations to communicate about the opportunity (such as the Coalition for Networked Information, the Data Curation Network, COGR, and DataCite). As an incentive for participation and sustained engagement, we are offering institutional teams stipends for their engagement throughout the pilot period to offset costs such as graduate student assistants. The ARL Project Manager, in partnership with the project team, will develop and deliver this multifaceted institutional maDMP pilot program.
- User stories and profiles: The ARL Project Manager, in partnership with the CDL Developer, will create user stories and profiles for the principal stakeholders utilizing the DMPTool to collect functional requirements from pilot institutions.
- Wireframing: The CDL Project Developer and ARL Project Manager will design initial wireframes detailing maDMP workflows and share these with potential pilot partners to collect feedback. We will continue to iterate on workflows throughout the project; however, this phase of work will provide important guidance to inform the course of our initial technical development.
- Webinars: The ARL Project Manager and team leaders will offer at least three open webinars offering a background to maDMPs, an overview of this project, and ways for institutions to get involved. These webinars will be open to the larger library community and serve as the first public introductions to our work.

Phase 2 January - July 2024 Institutional Pilot Launch

- Institutional Pilot Launch: The pilot program will begin in January 2024 and run until December 2024. The pilot partners will review project workflows and features and provide insight into how new tooling could overlay with their internal workflows to support pre and post-award tracking.
- Kick-off meeting: the project team will host a half-day virtual kickoff meeting with the institutional pilot teams. This meeting will introduce the pilot program to the institutions and discuss the 12-month pilot program. We expect all members of the institutional teams to attend this kick-off call. The kickoff meeting will include representatives from the library, research office, and IT. See the Draft Agenda in Item 3 of the Appendix.
- Testing and feedback: The CDL Developer will deploy the first round of new features to track DMPs and connect plans to research outputs. The Project Manager will facilitate testing and feedback sessions with all pilot partners to gather responses to the features and communicate ideas to the Developer to inform feature iterations.
- Monthly check-ins: Following the kickoff meeting, the Project Manager will conduct monthly check-ins with the institutional teams to provide technical and functional updates on the maDMP development and get feedback on requirements and maDMP features.
- Campus guidelines: The ARL Project Manager will begin to draft campus guidelines and best practices and engage the broader library community in feedback and reviews. The ARL Project Manager will also blog about the ongoing project on the CDL and ARL websites.

Phase 3 August - December 2024 Campus visits with each pilot institution

- Campus visits: To test new DMPTool features, discuss institutional needs, and create a plan for campus coordination, project team members will schedule in-person campus visits with each institutional pilot group. These one and half day visits will allow institutions to leverage the expertise and knowledge of the project team around maDMPs, while socializing and building local support for their workflows. The campus visit will include, a minimum, the following conversations, meetings, and presentations:
    - Institution-wide: Introduction to maDMPs and the goals of more easily facilitating campus coordination for research data services
    - Administrator-focused: Vision and institutional commitments, including representatives from the research office, university library, research computing, and IT, among others
    - Institutional team: Models for operationalizing the maDMP workflow at your institution
    - Workshop on drafting a work plan and administrative report for the next months to move towards coordinated maDMP workflow
- Synthesizing feedback: Following the campus visits, the ARL Project Manager will assemble feedback, outstanding feature requirements, and areas for future development and produce a report synthesizing the current state of the project. We will share this report via the extensive CDL and ARL communication channels.

Phase 4 January - July 2025 Final deployment of new DMPTool features with corresponding campus guidelines and best practices

- Final focus groups: After the pilot program, the project team will conduct brief focus groups or interviews with each institutional team to understand best practices, key challenges, and opportunities for the widespread adoption and use of maDMPs. This information will culminate in a set of campus-based guidelines and best practices for other institutions to use. Given the small number of institutions involved in the maDMP workflow pilot, the project team members will offer six $500 stipends for individuals from a diverse set of institutions to review and provide feedback on the guidelines. This review aims to ensure significant community engagement and a broad representation of institutions in the best practices.
- Production maDMP release: Finalize the deployment of the DMPTool maDMP feature sets, including new features for administrative reporting and research output tracking. All workflows and features will be available to all DMPTool participating members, and all software will be shared as non-proprietary and open-source with a fully permissive MIT License.
- Final documentation and communication: Publish wikis, readmes, and other documentation to support additional deployments of the software and workflows. The Developer and Project Manager will produce this documentation in partnership to ensure the usability of the new functionalities and workflows developed in this project with the larger DMPTool user community.
- Write and disseminate guidelines: Finalize campus guidelines and best practices and engage the broader library community in feedback and reviews. The project team will disseminate these guidelines and best practices with the larger community through a series of 5 open webinars featuring lessons learned through the pilot project, challenges, opportunities, and ideas for the next steps in continuing to expand the capabilities of maDMPs.

Diversity Plan: The DMPTool has widespread community adoption with participating organizations in forty-nine states and significant usage in all geographical regions of the US (see Appendix Item 4 for a geographical breakdown of members). Partnering institutions range from R1 research universities to smaller research institutions, laboratories, and field stations (see Appendix Item 5 for more details on the classification of members). A quarter of all academic institutions partnering with the DMPTool are designated as Minority Serving Institutions.

At least two of the four pilot partners in this project will be designated Minority Serving Institutions. Additionally, the project will give half of the stipends available in the project's final phase to individuals from non-R1 institutions or Minority Serving Institutions. CDL and ARL will use well-established policies to ensure diversity, equity, and inclusion principles inform recruiting and hiring the Developer and Project Manager, respectively.

This project includes outreach to smaller, less-resourced institutions that do not have the current infrastructure or staffing levels to adapt to the rapidly changing environment of data management and sharing mandates. Administration of the DMPTool is generally the responsibility of library staff, often a data librarian or the equivalent. As data-sharing mandates have increased, researchers are increasingly tasked with new research data management

responsibilities, and the role of the data librarian is expanding. However, thus far, this growing load has yet to be met with increasing staff levels. This project seeks to bolster open infrastructure in the scholarly communications space to level the field of quality information and resources for all. The project will meet current needs by expanding libraries' capacity to scale services and respond to a changing landscape by offering new, freely available tools and workflows.

Project Results: Our work will showcase enriched information regarding research outputs and how that information is collected and disseminated. The long-term effect will be a simplified approach for the entire community to track information about research project outputs from inception through publication and sharing results. The results generated from this project will facilitate the streamlined and effective coordination of research data services within institutions, providing institutions with a means to meet the new federal data sharing requirements.

Intended results:

1. Development of a community-designed and tested system for the users to query and access information about research outputs. This system will collect and display grant-related information and research outputs to track connections between DMPs and related research outputs throughout the lifecycle. For example, connecting funding and grant information with associated preprints, datasets, and software packages that have been published and linking these to the DMP.
2. Series of public webinars updating the community on our work and showcasing applied use cases of maDMP workflows.
3. Publication of reports, use cases, and campus-based strategies from pilot implementations.

All technologies developed in this work will be shared as non-proprietary and open-source with a fully permissive MIT License. All workflows and resources that facilitate tracking research data outputs will be publicly available for all communities to participate in, thus allowing all to use and benefit from this project. These new features and workflows will be built into the DMPTool, which already has widespread adoption by a diverse and growing community of institutions.

In addition to the large DMPTool community, the resources will be available for implementation by the global community of service providers who utilize the shared, open-source codebase. There are many implementations of the codebase worldwide, including national instances in Brazil, Hong Kong, the United Kingdom, France, Canada, South Africa, Spain, Belgium, Germany, Denmark, Australia, Finland, the Netherlands, and Japan.

Rather than building new systems which add an increased burden for the long-term sustainability of the project outputs, this project builds on existing open infrastructure to increase the capacity of institutions to respond to expanding federal requirements for data sharing. CDL has demonstrated its commitment to supporting the DMPTool for the past thirteen years and will continue to provide this support as the tool continues to mature into a robust tool supporting effective data management practices at scale.

Schedule of Completion: DMP Tool: Building a scalable data management infrastructure for strategic institutional coordination

	Year 1				Year 2			
DMP Tool: Building a scalable data management infrastructure for strategic institutional coordination	**2023 Q3**	**2023 Q4**	**2024 Q1**	**2024 Q2**	**2024 Q3**	**2024 Q4**	**2025 Q1**	**2025 Q2**
Develop a recruitment process to enlist four institutions to participate in the pilot project								
Institutional pilot launch								
Establish partnerships with partner institutions to pilot the new DMPHub workflows								
Conduct research, user studies and profiles, functional requirements testing								
Kick-off meeting between project team and pilot campuses								
Testing and feedback of technical features to support desired workflows								
Campus visits / in-person consultations								
Monthly check-ins with pilot campuses								
Iterate on the technical features of the DMPHub based on pilot feedback								
Ensure open source software is openly available for others to leverage								
Webinars to inform community of our work								
Document and update campus guidelines and lessons learned in the pilot program								
Final focus groups								
Production maDMP release								
Final documentation and communication								

Color Key
General Planning & Coordination
Technical Work
Outreach & Communications
Meetings

Digital Products Plan

What digital products will you create?

1. Open-source software: This project will develop a series of new machine-actionable features within the existing DMPTool codebase. All workflows and resources that facilitate tracking research data outputs will be freely and publicly available for all communities to participate in, thus allowing all to use and benefit from this project.

2. Publication of reports, use cases, and campus-based strategies from pilot implementations.

3. Series of public webinars updating the community on our work and showcasing applied use cases of maDMP workflows.

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

1. Open source software: All technologies developed in this work will be shared as non-proprietary and open-source with a fully permissive MIT License and accessible via open Github repositories.

2. Reports: All reports will be openly available and given persistent identifiers. All project reports will be available via the Zenodo repository. Zenodo data files are stored in CERN Data Centres, primarily in Geneva, with replicas in Budapest. Data files are kept in multiple replicas in a distributed file system, which is backed up to tape nightly.

3. Public webinars will be openly available to all (not just pilot participants) and be recorded for future viewing. Recorded versions of the webinars will be posted on the ARL and CDL communication channels and available via YouTube. All materials produced from the webinars will be made publicly available via Zenodo and designed to be reproducible for further events.

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?

All data and products developed for the project will be made available for review and reuse during the course of the project and following its completion. All data produced by consuming partners will be freely available to the public; we anticipate no sensitive or confidential data.

How will you address the sustainability of your digital products?

All new features and workflows of this project will be built into the DMPTool, which already has widespread adoption by a diverse and growing community of higher education institutions. The DMPTool is built on a widely used open-source codebase and will be available for implementation by the global community of service providers who utilize the shared, open-source codebase. There are many implementations of the codebase worldwide, including national instances in Brazil, Hong Kong, the United Kingdom, France, Canada, South Africa, Spain, Belgium, Germany, Denmark, Australia, Finland, the Netherlands, and Japan. Rather than building new systems which add an increased burden for the long-term sustainability of the project outputs, this project builds on existing open infrastructure to increase the capacity of institutions to respond to expanding federal requirements for data sharing. CDL has demonstrated its commitment to supporting the DMPTool for the past thirteen years and will continue to provide this support as it matures into a robust tool supporting effective data management practices at scale.

Organizational Profile: California Digital Library (CDL), the University of California Office of the President

California Digital Library (CDL) exists to support the University of California (UC) community's pursuit of scholarship and to extend the University's public service mission. CDL is a unit in the department of Academic Planning, Programs, and Coordination, under Vice Provost Douglas Haynes at the UC Office of the President. Guenter Waibel is the Executive Director responsible for the direction, development, and management. CDL's primary advisory body is the UC Systemwide Library and Scholarly Information Advisory Committee, representing the faculty senate, information and educational technology, university libraries, and scholarly publishing.

The University of California founded CDL in 1997 to take advantage of emerging technologies that were transforming how digital information was published and accessed. As a result, CDL's diverse and talented staff have assembled one of the world's largest digital research libraries and have changed how faculty, students, and researchers discover and access information. Our vision is to elevate the digital library for UC to become "expansively global and deeply local."

California Digital Library is uniquely positioned within the community as a library, service provider, publisher, data repository, and community advocate. CDL is a long-established Open Science infrastructure provider with extensive experience running cross-cutting services. These services include work in data publishing (including a partnership with the open data repository, Dryad), persistent identifiers (including as project lead for ARKs: Archival Record Keys and ROR: the Research Organization Registry), research data management (including as the provider of the data management planning tool, DMPTool), and digital preservation (including the Core Trust Seal-certified Merritt repository). CDL also convenes global discussions on topics important to the research data community, such as Make Data Count, PIDapalooza, CSVconf, and the Support Your Data matrix. As part of UCOP, CDL is also a participating member in the HELIOS project, a cohort of colleges and universities formed in partnership with the National Academies of Sciences, Engineering and Medicine's Roundtable on Aligning Incentives for Open Science to create collective action to advance open scholarship. In this project, CDL will lead all aspects of the technical development and implementation.