

The California Digital Library (CDL) and The Association of Research Libraries (ARL) request \$668,047 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 the research library community to accelerate its 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, such as the DMPTool and persistent identifier (PID) registries, to focus on meeting funding agency requirements for creating machine-readable 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 and, 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 within research institutions, enabling institutions of varying sizes and research budgets to meet a range of 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 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 that will have data sharing requirements, including IMLS and the National Endowment for the Humanities among others. The next steps in supporting these new requirements is to perform technical work to enable implementation and compliance by re-tooling existing research data management infrastructure; and within libraries, developing or enhancing institutional research data services strategies that incorporate technology-agnostic tools and services. 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 inter-institutional research support coordination.

IMLS was one of the original funders of the DMPTool with a 2012 grant to support training in research data management for librarians. 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 and after several years of consultation with the broader scholarly infrastructure community and within the Research Data Alliance, CDL identified the need for a digital center—or hub—to collect and disseminate updates about research. In response, CDL built a prototype version, currently referred to as the DMPHub, to generate PIDs for DMPs created via the DMPTool and to record updates to the DMP that occur over time. Our current proposal builds on this experience with the DMPTool and the DMPHub to create a fully scalable workflow that will enable institutions seeking to demonstrate compliance with new data sharing mandates.

CDL and ARL have a history of collaboration on the application of PIDs and DMPs to support data sharing and reuse. The two organizations were awarded an NSF grant (#[1945938](#)) in 2019, which resulted in published recommendations to support the broad adoption of PIDs and maDMPs. This proposed project will put these recommendations into practice by combining infrastructure and strategic planning to address the needs of the research library community as it continues to evolve services to meet faculty needs and to support institutional compliance for research data management and sharing policies.

The Project Work Plan is designed to include two work packages over the course of two years. The work packages are carried out in parallel, with each package informing the work of the other as they progress.

Work Package 1: Campus strategy and guidance

- Partner with at least three ARL member institutions to pilot the new DMPHub workflows utilizing information about research outputs that are made available through the maDMP (implemented in Work Package 2)
 - Develop criteria to ensure a representative group of institutions (balancing expertise and diversity) to participate in the pilot program.
- Conduct research, user studies and profiles, functional requirements testing, and other work informing the development of new tools, features, and workflows to support the expansion of the DMPHub.
 - Draft campus guidelines and lessons learned in the pilot program, which will guide iterations.
- Engage the broader library community (including non-R1 institutions). Provide honoraria for institutional teams to review the guidelines or workflows for the adoption of maDMP's.
- Activate an outreach campaign to publish, adopt, educate, and share project outputs with the larger community through CDL and ARL websites, webinars, conference presentations, etc.

Work Package 2: Infrastructure development

- Design, build, and deploy features to expand the capacity of the machine-actionable maDMP (features researched and developed in our previous NSF award).
- Iterate on the technical features of the DMPHub based on pilot feedback (from Work Package 1).
- Develop open source software extending the reach of the DMPHub platform.

Project Results: Our work will showcase enriched information regarding research outputs and how it 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 the publication and sharing of results.

1. Creation of a new open-source tool that tracks connections between DMPs and related research outputs throughout the research lifecycle.
2. Development of usable interfaces for the users to query and access information about research outputs.
 - a. Collect and display grant-related information and research outputs. Connect funding and grant information with associated preprints, datasets, and software packages that have been published and link these to the DMP.
3. Publication of reports, use cases, and campus-based strategies from pilot implementations.

Budget: Total Funds requested: \$668,047. Direct costs: \$594,741 and Indirect costs: \$73,306. **California Digital Library:** Direct cost included: (1) \$376,067 salary and wages for a Project Developer and (2) \$10,765 for travel and meetings. **Association of Research Libraries:** (a) \$98,640 for salary and wages for a Project Manager, (b) \$20,000 for travel and meetings, (c) \$24,000 for Participant Support and \$3,000 for honoraria for non-R1 workflow consultation, (d) \$30,500 for other direct costs and (e) total subcontractor IDC: \$31,776. CDL and ARL will provide a \$668,047 cost share for their respective budgets.