

Curating a Community Registry of Research Organizations

California Digital Library (CDL) requests \$249,747.11 for a National Leadership Grant for Libraries (NLG-L) Project Grant in the National Digital Infrastructures and Initiatives category to establish a Community Curation Advisory Board (CCAB) for the [Research Organization Registry \(ROR\)](#).

Statement of National Need

[ROR launched](#) in January 2019 with a mission to provide open, trusted, and noncommercial infrastructure for research organization metadata. ROR is a comprehensive and authoritative registry of organizations (close to 100,000 at present), all with unique IDs. ROR's open API can be integrated into any system that handles affiliations (journal workflows, [data repositories](#), institutional repositories). [DataCite](#) and [Crossref](#), which collectively register [DOIs](#) for the majority of scholarly outputs globally, are supporting ROR IDs in their metadata, which enables powerful and clean affiliation data for every object with a DOI.

ROR fills a crucial gap in scholarly infrastructure. Persistent identifiers (PIDs) are increasingly central to the [research landscape](#) in the United States and beyond. Researchers are using [ORCID](#) IDs, and DOIs are assigned to articles, datasets, dissertations, and other outputs. PIDs facilitate access to research as well as the tracking of research use and impact. But without an identifier for the *institutions* affiliated with authors and outputs, ORCIDs and DOIs only take us so far. ROR is the missing piece of the PID puzzle.

Identifying all outputs from an organization should be simple, but this vital information is often locked behind a paywall, or may only exist as unstructured text, even for the same affiliation (“University of California-Los Angeles” vs. “UCLA”). This results in inefficient efforts by institutions to gather and reconcile various data sources and the high cost of paying commercial vendors for their own information.

ROR is uniquely focused on accessible infrastructure by and for the scholarly community. Its scope encompasses all research organizations and its aim is explicitly toward and driven by the needs of research stakeholders. Librarians and academic administrators increasingly need access to data on their institutions' publications and research outputs to support reporting requirements, funder and federal mandates, institutional Open Access policies, and library collection development and licensing negotiations.

Project Design

To maintain the ROR registry, ROR will establish a Community Curation Advisory Board (CCAB). Research organizations are not static, and ROR must respond to inevitable changes when new organizations form and existing organizations change names, merge, or close. We believe academic institutions and especially librarians are needed to develop curation policies and processes for ROR, as they ensure the integrity and longevity of registry data in line with community-based values.

The CCAB will bring together library and PID experts—two groups that do not always intersect, but whom ROR has already connected through a [Community Advisory Group](#) formed earlier this year. Community group members have engaged in discussions and preliminary planning about curation policies and processes, and have reinforced the import of involving institutions, librarians, metadata experts, and research administrators in curation. Work with this group has revealed an existing base of experts willing to

be part of crucial infrastructure initiatives within the community, and we plan to transform and expand this base into a formal curation board.

ROR aims to embody the heterogeneity of research organizations within a controlled framework for managing organizational metadata. Therefore, we seek CCAB members experienced in non-English languages, non-Western institutions, and scholarly communications in the Global South, and who represent different institution types and geographic regions. We will solicit CCAB nominations in Fall 2019 from groups at all career stages and across library, scholarly communication, digital curation, open infrastructure, and PID communities. The ROR project team will lead the selection process.

The CCAB will work with a curation coordinator from the ROR project team to develop policies to guide curation practices, including criteria for adding organizations to the registry, how ROR will identify trusted representatives for organizations (i.e., those authorized to request metadata changes on behalf of an organization), data modeling (i.e., adding new metadata fields according to community priorities), and processing changes (i.e., what information to log when a record is changed, and how records might be versioned to capture changes over time). We will also finalize workflows for how requests are submitted to ROR, how they are triaged and reviewed by the CCAB, and how curation decisions will be made and communicated. While the policy work is underway, the curation coordinator will oversee the development of tools needed to support curation, including the functionality to edit and version records and to support searches and queries in ROR's database.

Requests to update ROR are currently handled via email by the ROR project lead. The CCAB and parallel technical development to support curation will allow us to move beyond this pilot phase. The CCAB will be operational by the time we would be eligible for a full proposal. If funded, we will be scaling curation workflows by the time the grant period begins.

The CCAB will collaborate primarily via email, Google Docs, conference calls, and Github. Given the importance of developing appropriate and effective policies, face-to-face meetings are needed for in-depth discussions and collaborative work sessions. Meetings will be colocated with major conferences to facilitate travel logistics, and we will provide travel funding to ensure all curators are able to attend.

National Impact

Open infrastructure for identifying research outputs by institution gives librarians and administrators an opportunity to control and leverage their own data, tracking and capturing key insights such as which of their researchers' outputs are supported by a particular funder, or available Open Access, or which of their publications might be affected by a high-profile license negotiation like the University of California's recent experience with Elsevier. The ROR CCAB is an opportunity to bring academic librarians front and center to open infrastructure initiatives that increase public access. Their expertise is crucial for ROR to be a trusted source of information and to be seen, valued, and used as a community good.

Budget Summary

CDL is requesting \$249,747.11 in IMLS funds for a two-year project to establish the CCAB. This funds a PI and Curation Coordinator (\$145,509.43: salary and benefits) and in-person meetings (\$66,500: travel and meeting expenses) to ensure CCAB participation by a diverse group of stakeholders regardless of financial need. Indirect costs are calculated at our Federally negotiated rate 17.8% of MTDC (\$37,737.68).