

ALFRED P. SLOAN FOUNDATION

www.sloan.org | [proposal guidelines](#)

PROPOSAL COVER SHEET

Project Information: DuraCloud "Direct-To-Researcher"

Principal Investigator

Michele Kimpton
DuraSpace, PO Box 4177, Ithaca NY
Phone: (415) 845-7588
Email: mkimpton@duraspace.org
(co-PIs: Sandy Payette, Brad McLean,
Jonathan Markow)

Amount Requested: 497,433

Requested Start Date: July 1, 2011

Requested End Date: December 31, 2012

Project Goal: Our goal is to develop and deploy a "Direct-To-Researcher" cloud platform as a safe and reliable way for researchers and scientists to store and manage their research data. We plan to evolve the DuraCloud platform, developed by the DuraSpace not-for-profit organization, to give researchers and data managers a cloud-based solution that is built for their specific needs and that goes beyond what utility cloud vendors provide. We will focus on the challenges that individual researchers confront in facing the "data deluge" by developing an easy-to-use cloud service that directly serves the researcher, without requiring mediation. The researcher can use the same service to delegate to data specialists from supporting institutions to help curate data for preservation, archiving, and access.

Objectives: Our objective is serve researchers by creating cloud-based service for data storage and management, built with the DuraCloud open source software. The service will be positioned as an alternative to institutional systems, elaborate data grids, servers in laboratories, and personal disk drives in researcher's offices. Key research questions include: (1) How can we empower the researcher to have autonomy in using DuraCloud, and also enable an institutional "back door" for curators, data mangers, and archivists?, (2) How should we support institutions in their need for security, data privacy, and other legal aspects of archiving? and (3) What are the economics of a cloud solution for research data?

Proposed Activities: We will engage a diverse set of researchers from our network of established partnerships and enlist a set of pilot participants. Two invitational workshops will be held to define scenarios, use cases, and requirements. We will leverage our extensive experience in creating successful open source software and work with our proven agile development process work collaboratively with researchers while iterating through multiple cycles of design, development, and evaluation.

Expected Products: (1) Reports from two invitational workshops, (2) Documentation of scenarios and use cases, (3) Technical specifications, (4) White paper on the Direct-To-Researcher value proposition, (5) Conference presentations, (6) Public release of new version of DuraCloud open source software, (7) Final report on project, (8) Launch of DuraCloud Direct to Researcher hosted cloud service.

Expected Outcomes: As providers of software technologies, we measure our broader impact in terms of the contributions we make to the greater cause of enabling scientists, researchers, and data curators in managing, accessing, and analyzing data for the purpose of bringing forth new insights. In the longer run, our impact will be measured by the availability of high quality data that has been successfully curated and archived. Long-term impact for data management systems can be achieved with sustainable software, evolvable systems, open standards, and community-driven processes.
