

Project Title: Data Management Training Clearinghouse**SECTION 1. PROJECT ADMINISTRATIVE INFORMATION****CDI SSF elements:** Science Data Lifecycle Model, Computational Tools and Services**Name of lead USGS cost center:** Upper Midwest Environmental Sciences Center**Name of Principal Investigator:**JC Nelson, USGS Upper Midwest Environmental Science Center, 2630 Fanta Reed Road LA Crosse WI 54603, 608-781-6370, Fax 608-783-6066, jcnelson@usgs.gov**Additional PIs:**Nancy Hoebelheinrich, Knowledge Motifs LLC, San Mateo, CA 94401, 650-302-4493, Fax 650-745-3333, nhoebel@kmotifs.com.Tamar Norkin, USGS Core Science Analytics, Synthesis and Libraries, Denver Federal Center, Bldg. 810, MS 302, Lakewood, CO 80225, 303-202-4229, Fax 303-202-4229, tnorkin@usgs.gov**Collaborators:**Amber Budden (DataONE, aebudden@dataone.unm.edu), Robert R. Downs (CIESIN, Columbia University, rdowns@ciesin.columbia.edu), John Faundeen (USGS EROS, faundeen@usgs.gov), Sophie Hou (University of Illinois, Urbana-Champaign, hous@illinois.edu), Shelley Knuth (University of Colorado, Boulder, Shelley.Knuth@colorado.edu), Matt Mayernik (NCAR, mayernik@ucar.edu), Erin Robinson (Federation of Earth Science Information Partners, erinrobinson@esipfed.org)**Short description of the project:** The team will create an online clearinghouse of data management training resources. The clearinghouse will provide a searchable and browseable inventory of training resources related to data management in the Earth sciences. To assist with discoverability and scope definition, training resources will be associated with elements of the [USGS Science Data Lifecycle Model](#) and the [Learning Resource Metadata Initiative \(LRMI\)](#) specification. The team will develop a sustainable mechanism for adding and describing new resources through crowdsourcing.**List of anticipated deliverables:** Online clearinghouse of data management training resources, template and online help guide to facilitate the description of training resources**SECTION 2. ESTIMATED BUDGET**

Budget Category	Federal Funding "Requested"	Matching Funds "Proposed"
1. PERSONNEL (SALARIES including benefits):		
Federal Personnel Total:	\$10,000	\$5,000
Contract/Collaborator Personnel Total:	\$20,000	\$5,000
Total Salaries:	\$30,000	\$10,000
2. TRAVEL EXPENSES:		
Travel (Per Diem, Airfare, Mileage/Shuttle) x # Trips:	\$2,000	\$2,000
Other Expenses (e.g. Registration Fees):	\$0	\$830
Total Travel Expenses:	\$2,000	\$2,830
3. OTHER DIRECT COSTS: (itemize)		
Equipment: hosting and hardware allocation	\$4,000	\$0
Publication Costs:	\$0	\$0
Office Supplies, Training, Other Expenses (specify):	\$0	\$0
Total Other Direct Costs:	\$4,000	\$0
Total Direct Costs:	\$36,000	\$12,830
Indirect Costs (18.62%):	\$6,703.20	\$2,388.95
GRAND TOTAL:	\$42,703.20	\$15,218.95

SECTION 3. PROJECT SUMMARY

Many organizations, government agencies and academic institutions in the U.S. and abroad have developed excellent educational resources to inform and train scientists and their information partners about best practices in data management. Some of these resources are focused on data management within specific domains. It has been our experience that educational resources are often difficult to locate, as they are spread out across numerous websites. In addition, resources are often difficult to evaluate with respect to their effectiveness and their relationship to a specific domain or framework. To address this need, a group of organizations that have experience creating, presenting and distributing data management training resources will collaborate to establish a clearinghouse for data management training in the Earth sciences.

This project will build on previous CDI-funded projects, including the [Science Data Lifecycle Model](#) (part of the [CDI Science Support Framework](#)) and the [USGS Data Management website](#). Although both of these were focused within the USGS, the proposed project will have a broad scope and will involve collaboration with academic institutions and non-governmental organizations and communities. An advisory group of interested and committed organizations has been formed to guide and support the project; represented organizations include the Federation of Earth Sciences, DataONE, USGS Core Science Analytics, Synthesis and Libraries, and Knowledge Motifs LLC. Members of the advisory group have experience creating and presenting training resources (for example, the [DataONE education modules](#) and the [ESIP Data Management Short Course for Scientists](#)). The advisory group will continue its work beyond the grant period to sustain the effort.

Funds allocated for salaries would cover tasks including project management, template creation and developer support. Travel expenses are included for 2 people to present at ESIP meetings and a CDI event. Equipment costs are included for hosting the clearinghouse and hardware allocation.

	Objectives	Timeline
1	Develop an inventory of Data Management Training (DMT) resources, collected from a search of existing online resources.	March 2016
2	Identify and adopt a hosting environment that provides mechanisms for collecting, storing, describing, browsing, and searching resources. In addition, it will allow curation and description by means of crowdsourcing with appropriate vetting. Current candidates for the hosting environment include ESIP Federation's Commons (a content management system-supported portal using Drupal), the Open Science Framework, and GitHub.	Mar–Apr 2016
3	Create a template and online help guide to facilitate and promote consistency in the description of DMT resources. The template will associate the resources with the USGS Science Data Lifecycle model, a key component of the CDI Science Support Framework. If technically possible, the template will also incorporate relevant tags from the Learning Resource Metadata Initiative (LRMI) specification, which is designed to improve discoverability of learning resources through Google and other search engines that use Schema.org's metadata tags.	Apr 2016
4	Using the template, incorporate the descriptions of the DMT resources into the host environment and make them available for searching and browsing.	May-Aug 2016
5	Assess and report on the feasibility of establishing an evaluation framework and sustainable review process for the DMT resources and potential features such as faceted searching.	Sept 2016

CDI15-NJ3655

Community for Data Integration Statement of Interest for Fiscal Year 2016