

## **Data Management Plan**

**Project Title:** Bridging communities of practice: Emerging technologies for content-centered linking

**Project Director:** Douglas W. Oard

**Beginning:** 05/01/2013 **Ending:** 4/30/2014 **Duration:** 12 months

This data management plan was created on September 24, 2012 for submission to the Office of Digital Humanities (ODH), National Endowment for the Humanities, as required by ODH Guidelines in order to request funding for this project. This is the first version.

**Types of Data:** The project will generate three types of data: (1) administrative data, such as registration information for each workshop, (2) work-in-progress data, such as meeting notes and draft reports, and (3) records of continuing value. Only the records of continuing value will be managed under this plan. Two types of records of continuing value are expected: (3a) records with value principally over a limited period, and (3b) records of enduring value. Records in group (3a) will include the project's Web site, audio and video recordings of workshop presentations, powerpoint slides from workshop presentations, and Twitter posts made during the workshop. Records in group (3b) will include the formal work products of the project, including the final report, and additional records necessary to provide context for those formal work products (e.g., agendas and lists of participants in the two workshops).

**Data Standards and Capture:** The project Web site will contain materials originally generated for the project, along with links to pre-existing external resources. Participants will be invited to upload materials, and we will make a concerted effort to obtain presentation materials and audio and video recordings of formal presentations. Presentation materials and audio and video will be hosted in widely-adopted formats such as pdf, mp3 and mp4 that are easily stored and shared. Additional formats offered by presenters (e.g., powerpoint) may also be hosted to facilitate reuse. Social media output will be captured using application programming interfaces (APIs) that these services provide. For materials in group (3b) the project will use naming conventions created in accordance with University of Maryland (UMD) Libraries policies. Group (3b) materials will be provided with digital identifiers to facilitate persistent citation and deposited with UMD digital collections.

**Metadata:** The Dublin Core Metadata scheme, as used by UMD libraries, will be used during the process of depositing group (3b) data. Metadata will be manually supplied by project personnel, and UMD library staff will ensure consistent application of standards.

**Legal Policy:** While participants of the workshop will retain the copyright and other intellectual rights over the material they contribute to the project, they will be asked to agree that such material will be licensed under an open source license (such as Creative Commons) according to UMD policies. This will also be the case for materials generated by the project. For audio and video recordings, participants will be asked to sign waivers consenting to their capture and sharing.

**Data Storage, Security, and Backup:** Group (1) materials will be stored in password protected servers with appropriate disaster recovery backup services. Group (2) materials will be managed using operational information systems such as enterprise computing and email systems, with appropriate disaster recovery backup services. Group (3a) materials will be hosted in the University of Maryland Institute for Advanced Computer Studies (UMIACS) research computing infrastructure for 10 years from the date of completion of the project. Both local and remote disaster recovery backup services are provided for this data. Data in group (3b) will be stored in the UMD digital collections infrastructure, and with copies maintained both onside and offside for disaster recovery. Business continuity failover services are not required for data in any group for this project.

**Access, Sharing & Re-use:** All group (3a) data from this project will be available for browsing or download from the project's Web site on a best effort basis (which historically has better than 99% availability). All group (3b) data from this project will be available for browsing or download from both the project Web site (during the 10-year term for which continued access is planned) and from UMD digital collections (permanently). Data will be available from the project Web site immediately upon release, and group (3b) data will be available from UMD digital collections within 6 months of release. No additional permissions will be required to download or reuse data.

**Long-Term Preservation:** The University of Maryland is invested in the management and curation of data created by researchers affiliated with the University, and in providing digital preservation services. Long-term preservation services for group (3b) data, including file validation, integrity checks, and (when required) format conversion, will be provided by UMD Libraries.