Project Proposal

For more information or guidance please see the host requirements, project examples and project characteristics as found on the NDSR website.

c/o George Coulbourne
ndsr@loc.gov | www.digitalpreservation.gov/ndsr

Date Submitted

February 9, 2016

Submitted by:

[redacted]

Project Proposal Title

Data Stewardship and Preservation Program

Project Summary
(provide 2-3 sentences summarizing the project)

Create the foundation for a digital data stewardship and preservation program at the World Bank. Research and design: an assessment framework for appraisal of datasets, guidelines for documenting data transformation, and plans for linking datasets to documents and analytical reports that use them.

Specific Goals and Objectives
(identify goals and objectives to be completed by end of residency)

1. Develop an assessment framework for the identification and appraisal of high value datasets for preservation.
   a. Validate the assessment against datasets that are frequently updated.
   b. Test this framework on recent sample datasets provided to the resident by World Bank Staff.
   c. Test this framework on data stored on legacy media in custody of the Archives.

2. Create guidelines for a process to document how to reproduce results of data transformation (such as for a poverty assessment). The output of this process will be an 'analytic process document' – similar to what is required by journals as part of the peer review process. This will require research into how other organizations perform this activity. The guidelines will be tested with at least one example, provided to the resident by World Bank Staff.

3. Research and document best practices and existing guidelines (such as the IHSN Principles and Good Practice for Preserving Data and the Data Seal of Approval).

4. Create guidelines for linking datasets to documents and analytical reports that use the respective datasets.

Timeframe and Deliverables
(provide a quarterly breakdown of assignments and deliverables necessary to complete objectives)

Overall: 12 months
Months 1-3:
• Learn about existing data initiatives
• Learn about Digital Vault and digital preservation strategies in place in the World Bank Group Archives
• Research and document best practices and existing guidelines (such as the IHSN Principles and Good Practice for Preserving Data and the Data Seal of Approval).
• Background research for assessment framework for the identification and appraisal of high value datasets for preservation

Months 4-6:
• Draft and document assessment framework for the identification and appraisal of high value datasets for preservation
• Test assessment framework against varied datasets provided by World Bank Staff and mentors
• Present results of assessment framework tests

Resources Required
(identify any resources necessary to complete the project)

• Mentors: [redacted] 1 resident
• Access to background information, material and data gathered by mentors and supporting teams
• Access to on-site library and research resources
• Access to resources compiled by the WBG Archives on related subject matter

Context
(provide a narrative statement explaining the project – no more than 500 words)

The World Bank Group (WBG) creates and uses a wide array of data, across the organization. Currently, much of the creation, curation and analysis of datasets is done by individual teams and team members with little centralized organization. The focus is on results and the creation of final reports that leverage the data analysis. There is an ongoing WBG initiative "Development Data Hub (DDH)" that aims to address these issues and improve the Bank's capacity to capture, curate and archive development datasets. It builds on work already done by the Open data Catalog and Microdata library initiatives. This is one of the key priorities endorsed by the World Bank's Data Council, chaired by the WBG's Managing Director and Chief Operating Officer, and the WBG's Chief Economist. One of the council's objectives is to oversee the development and implementation of relevant data policies, standards and protocols, including procedures for managing data acquired or developed by Bank staff in the course of their operational and research work.

Data preservation is one of the focus areas of the DDH initiative. While some high profile data sets are more carefully preserved, there is major opportunity to create a digital data stewardship and preservation program at the Bank.

In addition, the WBG Archives is currently designing and implementing a digital preservation strategy and a preservation platform called the Digital Vault.

This project will be executed as a separate project but will link to both the broader DDH initiative and the

Required Knowledge and Skills of Resident
(identify requirements necessary to successfully complete the project – technical skills, educational background, specialized experience, etc.)

Graduate degree in Library and/or Information Sciences, Information Technologies or equivalent from an accredited institution.
Experience within an archives or library setting
General knowledge in digital data management principles
Experience with using Microsoft Windows computers, Microsoft Office and Adobe Acrobat software
Some experience with using graphical design or imaging systems
Knowledge of basic digital preservation practice

Preferred Knowledge or Experience of Resident
(identify preferred knowledge or experience – technical skills, educational background, specialized experience, etc.)

Basic understanding of data stewardship and digital data preservation
Knowledge of appraisal and data assessment
Knowledge of business process analysis
Experience creating process flows