**Web Archiving Incentive Program  
Call for Proposals**

1. **Program Description**

Thanks to a generous grant from the Andrew W. Mellon Foundation, Columbia University Libraries / Information Services (CUL/IS) will administer a Web Archiving Incentive Program intended to support new developments in the broad domain of web archiving. We now invite proposals from qualifying institutions for projects that improve, enhance or innovate in the areas of web harvesting and use.

* 1. **Types of projects in scope**

The Program welcomes projects that involve analyses, improvements and innovations in areas that include (but are not limited to):

1. Development of new tools for web harvesting, web archive curation and use
2. Improvement/extension of existing web platforms in collaboration with publishers/developers to support archiving

1. Improvement/extension or innovative implementation of existing web archiving tools/services
2. Packaging/bundling of existing tools to assist creation, or use of web archives

Please see the appendix at the end of this call for proposals for examples of projects in scope.

A wide diversity of proposals within the projects in scope is being sought. Projects must focus on automated tools and provide thorough documentation of expected behavior of the tools. Preference will be given to projects that will modularly extend an existing tool, produce a fully functional tool, demo site or working prototype. When appropriate the creation of a full technical specification will be considered a final deliverable (an accompanying product specification is desired but secondary).

A set of sample data from Columbia University’s web archives will be made available for funded projects, but it is not required that the sample data set be used to complete the work funded.

* 1. **Types of projects not in scope**

The current grant will not support proposals for educational or training programs, internships or residencies.

* 1. **Providing access to grant products**

All software developed must be fully documented and made available under one of the open-source software licenses approved by the Open Source Initiative (preferably a GNU General Public License) with code deposited in an open-source repository such as GitHub or SourceForge, or licensed Creative Commons (Attribution-NonCommercial), as applicable. If building on existing open source code, the license already assigned to the code will be applied to additions.

1. **Award Information**

Awards amounts are expected to range from $20,000 to $25,000, but proposals requesting smaller or larger amounts will be considered. Allowable expenses under these awards include salary and fringe for project staff as well as travel. Indirect costs will not be assessed or covered.

All work must be completed and final reports submitted by October 30, 2014. Earlier completion dates are encouraged.

1. **Eligibility**  
     
   Grants must be institutionally sponsored by a college, university or non-profit organization (501(c)3 or 501(c)4) in the United States of America, with final proposals submitted through the institution’s sponsored research process. International applicants are eligible to apply only in conjunction with a qualified institution in the United States. Some work under the award may be subcontracted to third parties.
2. **Application and Submission Information**

An initial one page prospectus is requested. In this prospectus, please include: a brief description of the project and its intended outcomes and benefits, an overview of methodology and standards that will be employed, and a brief summary of the qualifications of the project staff. This prospectus is due September 30, 2013. At this stage, evidence of approval by the institution’s sponsored research (grants) office is not required; however, applicants should be prepared to obtain all necessary approvals should a full proposal be requested.

If the project is approved for further consideration, a full application will be requested by October 30, 2013. The full proposal will be due December 30, 2013 and should include the following:

* 1. Project description -- Abstract
  2. Project description -- Full
     1. Overview and significance of work proposed
     2. Background of applicant(s)
     3. Scope and duration for project
     4. Examples of related past work (if applicable)
     5. Methodology and standards
     6. Work plan
     7. Staff
     8. Budget
     9. Institutional letter of endorsement
     10. Appendices (bibliography, resumes, intellectual property agreement, complete list of participants, consultants & advisers)

Application materials in each round must be submitted in a single .pdf file (please do not send multiple attachments or submit materials in other file formats).

Awards will be announced by February, 2014. A mid-project progress report will be required and the deadline for completion of projects is October 30, 2014.

1. **Application Review Criteria**

Proposals will be evaluated by an Oversight Panel of individuals with expertise in web archiving. Successful proposals will demonstrate the how the project is innovative, feasible and has the potential for wide benefit. Preference will be given to tools that are generalizable such that they will useful to web archiving programs at multiple institutions.

6. **How to Submit**

Applicants should submit the initial one-page prospectus as a pdf email attachment to [culwebgrant@libraries.cul.columbia.edu](mailto:culwebgrant@libraries.cul.columbia.edu) no later than September 30, 2013. Detailed instructions for submitting full proposals will be sent to selected applicants no later than October 30, 2013.

If you have any questions or concerns regarding this call for proposals, please contact the Web Archiving Project Librarian for this grant, Anna Perricci, at [alp2198@columbia.edu](mailto:alp2198@columbia.edu).

**Appendix**

**Examples of types of projects in scope**

The Program welcomes projects that involve analyses, improvements and innovations in areas that include (but are not limited to):

Development of new tools for web harvesting, web archive curation and use

Examples:

* + Scoping crawls, analysis of website archivability
  + Capturing web content difficult to harvest via traditional crawling (rich media, database-driven features, dynamically generated content, changing URIs)
  + Automated QA analysis of harvests
  + Automated metadata extraction/generation
  + Data mining of web archives
  + Data visualization in web archives
  + New APIs for analysis or discovery of web archives
  + Browser extensions for harvesting websites on demand

Improvement/extension of existing web platforms in collaboration with publishers/developers to support archiving

Examples:

* + - * WordPress plugin for archiving
      * Drupal optimization for archiving
      * Solr 4 optimization for indexing contents of .warc files, improving full-text search relevance, result clustering, multi-language support
      * APIs (i.e. Twitter)
      * Browser configuration to improve navigation of archived websites (i.e. Chrome)
      * Promoting native compliance with Memento
      * Integration of discovery services (such as widely adopted metasearch products)

Improvement/extension or innovative implementation of existing web archiving tools/services

Examples:

* + Memento
  + SiteStory
  + WAIL
  + WebCite
  + ArchiveReady

Packaging/bundling of existing tools to assist creation, or use of web archives

Examples:

* + - * Supplementing crawlers with headless browsers or modules for capturing special content, e.g. streaming media
      * Providing a suite of data mining and/or data visualization options

A wide diversity of proposals within the projects in scope is being sought. Projects must focus on automated tools and provide thorough documentation of expected behavior of the tools. Preference will be given to projects that will modularly extend an existing tool, produce a fully functional tool, demo site or working prototype. When appropriate the creation of a full technical specification will be considered a final deliverable (an accompanying product specification is desired but secondary).