

# Workset Creation for Scholarly Analysis: Prototyping Project

The [HathiTrust Research Center](#) is seeking **proposals for prototyping projects** to define and implement a tool or service that will help scholars better identify and select relevant resources at scale from the [HathiTrust](#) corpus and/or facilitate the construction of large-scale worksets useful for scholarly analyses. Grants of \$40,000 will be offered to each of four successful respondents to be conducted over a nine-month period beginning April 2014. Workset Creation for Scholarly Analysis: Prototyping Project (WCSA) is generously funded by the Andrew W. Mellon Foundation.

A complete copy of the RFP is available online at: [http://worksets.htrc.illinois.edu/worksets/?page\\_id=20](http://worksets.htrc.illinois.edu/worksets/?page_id=20).

## RFP Schedule:

RFP Available: 22 November 2013  
Letters of Intent Due (preferred): 16 December 2013  
Final Proposals Due: 13 January 2014  
Shortlist Meeting Invitations Issued: 20 January 2014  
Shortlist Meeting: 20 February 2014  
Award Notification: No later than 15 March 2014

## Program Description (see the [full RFP](#) for more detail):

The HathiTrust (HT) is a large digitized-text corpus (> 10 million volumes) of keen interest to researchers working in a wide range of scholarly disciplines. To tap the analytic potential of this large and diverse corpus, to tame it and make it useful to them, many researchers need the wherewithal to gather together, into a kind of personal digital carrel, cohesive and coherent subsets of HT texts (potentially tens or hundreds of thousands of volumes or parts of volumes) amenable to the in depth forms of analysis they want to do. The attributes on which they seek to collocate digitized texts are not always recorded in standard bibliographic descriptions.

The HTRC will collaborate with four independent sub-awardees in conducting individual prototyping projects to develop and validate the potential of specific algorithms, services and/or tools that can enable the creation of large and small scale worksets of digitized texts and parts of digitized texts for scholarly analysis in ways not currently feasible. We are seeking proposals from engaged teams of digital humanists, librarians and computer scientists. We anticipate that the proposals received will approach the problem in a variety of different and complementary ways. Proposed prototype experiments must respond to real scholar needs and requirements.

Respondents are urged to contact [htrc.wcsa@gmail.com](mailto:htrc.wcsa@gmail.com), in advance of proposal submission to discuss eligibility, project details, prerequisites, and HTRC support with a member of the project team. Prime award project PIs are:

J. Stephen Downie, Graduate School of Library and Information Science, University of Illinois  
Tim Cole, University Library, University of Illinois  
Beth Plale, Data to Insight Center, Indiana University