Authors
Stephanie Wykstra
Research Manager

IASSIST Quarterly

New Curation Software:

Step-by-Step Preparation of Social Science Data and Code for Publication and Preservation

by Limor Peer[1] and Stephanie Wykstra[2]

Abstract

research will be more credible if others can have full access to all aspects of scholarly work

New Curation Software: Step-by-Step Preparation of Social Science Data and Code for Publication and Preservation

As data-sharing becomes more prevalent throughout the natural and social sciences, the research community is working to meet the demands of managing and publishing data in ways that facilitate sharing. Despite the availability of repositories and research data management plans, fundamental concerns remain about how to best manage and curate data for long-term usability. The value of shared data is very much linked to its usability, and a big question remains: What tools support the preparation and review of research materials for replication, reproducibility, repurposing, and reuse? This paper describes key curation tasks and new data curation software designed specifically for reviewing and enhancing research data. It is being developed by two research groups, the Institution for Social and

Policy Studies at Yale University and Innovations for Poverty Action, in collaboration with Colectica. The software includes curation steps designed to improve the research materials and thus to enable users to derive greater value from the data: Checking variable-level and study-level metadata, verifying that code can reproduce published results, and ensuring that PII is removed. The tool is based upon the best practices of data archives and fits into repository and research workflows. It is open-source, extensible, and will help ensure that shared data can be used.

April 18, 2016