IDCC17 | *Poster*

The Curating for Reproducibility (CURE) Consortium

Florio Arguillas
Cornell University

Limor Peer
Yale University

Thu-Mai Christian
Sophia Lafferty-Hess
University of North Carolina at Chapel Hill

Abstract

In July 2016, the Institution for Social and Policy Studies (ISPS) at Yale University, the Cornell Institute for Social and Economic Research (CISER), and the Odum Institute for Research in Social Science at the University of North Carolina at Chapel Hill formed the Curating for Reproducibility (CURE) Consortium. These academic institutions all maintain data archives that have been involved in implementing workflows that put into practice data quality review (Peer, Green, & Stephenson, 2014), a framework that includes research data curation and code review. This framework helps to ensure that research data are well documented and usable, and that code executes properly to reproduce analytic results.

The poster will outline the goals of the consortium and provide examples of how these institutions have integrated data quality review into workflows, tools, and protocols.

CURE Goals

The main goal of the consortium is to work with a variety of stakeholders to establish the standards, share the practices, and promote the philosophy of data quality review.

1. Establish standards: Building on existing and evolving standards (e.g., data citation, TOP guidelines) CURE is dedicated to establishing and communicating the driving principles and criteria for proper curation for reproducibility.
2. Share practices: Curating for reproducibility involves multiple tasks and several stakeholders. A primary goal of CURE is to map the vital elements of the workflow, and to share the best practices that have emerged within each organization.

Submitted 18 November 2016

Correspondence should be addressed to Thu-Mai Christian, CB# 3355, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3355. Email: thumai@email.unc.edu

The 12th International Digital Curation Conference takes place on 20–23 February 2017 in Edinburgh. URL: http://www.dcc.ac.uk/events/idcc17/

3. Promote data quality review: Curating for reproducibility involves multiple tasks and several stakeholders. A primary goal of CURE is to map the vital elements of the workflow, and to share the best practices that have emerged within each organization.

Three Distinctive Use Cases

While the scope and scale of data quality review vary across organizations, each of the three founding CURE consortium members have demonstrated its commitment to reproducible research and open science.

Institution for Social and Policy Studies, Yale University

The Institution for Social and Policy Studies (ISPS) originally developed the data quality review framework and has implemented this framework for datasets deposited within the ISPS Data Archive. ISPS has also worked in collaboration with Colectica to develop a data curation software that facilitates the management of the numerous tasks involved in curating data for reproducibility.

Cornell Institute for Social and Economic Research, Cornell University

The Cornell Institute for Social and Economic Research (CISER) provides a research reproducibility service (R^2) to Cornell University researchers that uses standardized protocols for depositing data into the archive. CISER then curates the data and code to ensure they meet reproducibility requirements.

Odum Institute, University of North Carolina at Chapel Hill

The Odum Institute has developed workflows that integrate data curation and verification within the publication process. Odum currently works with two political science journals, *AJPS* and *SPPQ*, to enforce their data replication policies by verifying that submitted data and code properly reproduces the findings presented within a conditionally accepted manuscript.

References

[journal article] Peer, L., Green, A., Stephenson, E. (2014) Committing to data quality review. *International Journal of Digital Curation*, *9*(1), 263–291. doi:10.2218/ijdc.v9i1.317