

Proposed book chapter title: Data Curation for Small Databases; Research Data Management Solutions for Dinky Databases (Proposed for inclusion in *The New Librarianship*)

Authors:

Laurie N. Taylor, Laurien@ufl.edu

Mark V. Sullivan, marsull@uflib.ufl.edu

Val Minson, vdavis@uflib.ufl.edu

[More may be added from the Data Management/Curation Task Force, if the proposal is accepted. Limited group initially for the proposal, given the timeline for proposal submission.]

Abstract:

Researcher needs, funding agencies, and governmental requirements have made Research Data Management a critical concern. Academic research libraries are uniquely positioned to collaborate with other core institutional groups to address research data management needs, building from existing roles with librarians as liaisons and collaborators with many different researchers and research fields. Much of the research to date has focused on information gathering (e.g., data curation profiles and surveys) and information sharing (e.g., new training institutes and programs on data, data management, data curation, e-science, etc.). Additional work is needed on data to support librarians in terms of new roles, responsibilities, skills, and activities in the data age.

This chapter focuses on work by the University of Florida Smathers Libraries in collaboration with Research Computing and the Division of Sponsored Programs in building support for small or “dinky” databases as a supporting part of a larger socio-technical infrastructure for data management. This chapter includes an overview on the problem of dinky databases as small datasets requiring operational database functions for sorting and searching, normally supported by individual researchers and thus at risk for data loss and without the benefits of standardized support. Dinky databases are both a simple problem in terms of data size and normalization, and a complex problem with needed operational database functions where simply downloadable data is an insufficient solution for researchers to change current, problematic data practices. The chapter covers work to date and future plans by the UF Libraries for the dinky database problem, including gathering information on the many UF dinky databases, developing a database module for the SobekCM Open Source Digital Content Management System (which powers the institutional repository and digital collections at UF), and collaborating with Research Computing to develop additional computational supports by connecting the SobekCM interface frontend with advanced data hosting and computational backend supports by Research Computing. The chapter specifically addresses the work by librarians for the problem of dinky databases to show the work as a case study for how to build capacity for data management as a way to grow the overall culture of data management beginning with librarians as critical, expert collaborators.