

COLLABORATIVE DATA PUBLICATION UTILIZING THE OPEN DATA REPOSITORY'S DATA PUBLISHER. N.

Stone¹, B. Lafuente³, T. Bristow³, R. M. Keller³, R. T. Downs², D. Blake³, M. Fonda³, C. Dateo³, and A. Pires²,¹ Open Data Repository, Gray, ME² University of Arizona, Tucson, AZ³ NASA Ames Research Center, Mountain View, CA.

Introduction: For small communities in multi-disciplinary fields such as astrobiology, publishing and sharing data can be challenging. While large, homogenous fields often have repositories and existing data standards, small groups of independent researchers have few options for publishing data that can be utilized within their community.

In conjunction with teams at NASA Ames and the University of Arizona, a number of pilot studies are being conducted to assess the needs of these research groups and to guide the software development so that it allows them to publish and share their data collaboratively.

Objectives: Data Publisher aims to provide an easy-to-use software tool that will allow researchers to create and publish database templates and related data. The end product will facilitate both human-readable interfaces (web-based with embedded images, files, and charts) and machine-readable interfaces utilizing semantic standards.

Characteristics: The Data Publisher software runs on the standard LAMP (Linux, MySQL, Apache, PHP) stack to provide the widest server base available. The software is based on Symfony (www.symfony.com) which provides a robust framework for creating extensible, object-oriented software in PHP.

The software interface consists of a template designer where master database templates can be created and customized (see Fig. 1). A master database template can be shared by many researchers to provide a common metadata standard that will set a baseline for all derivative databases. Individual researchers can then customize their instance of the master template with specialized fields, file storage, or visualizations that may be unique to their studies. This allows groups to create compatible databases for data discovery and sharing purposes while still providing the flexibility needed to meet the needs of scientists in rapidly evolving areas of research.

Research: As part of this effort, a number of ongoing pilot and test projects are currently in progress. The Astrobiology Habitable Environments Database (AHED) Working Group is developing a shared database standard, and will be using Data Publisher templates to make data available from a number of sample astrobiology databases[1]. Work

with this group helps determine what data different researchers in astrobiology need to share and archive. Additionally, this pilot helps determine what standards are suitable for sharing these types of data – from internally developed standards to existing open standards such as the Dublin Core (<http://dublincore.org>) and Darwin Core (<http://rs.twdg.org>) metadata standards[2].

Further studies are ongoing with the University of Arizona Department of Geosciences where a number of mineralogy databases are being constructed within the Data Publisher system.

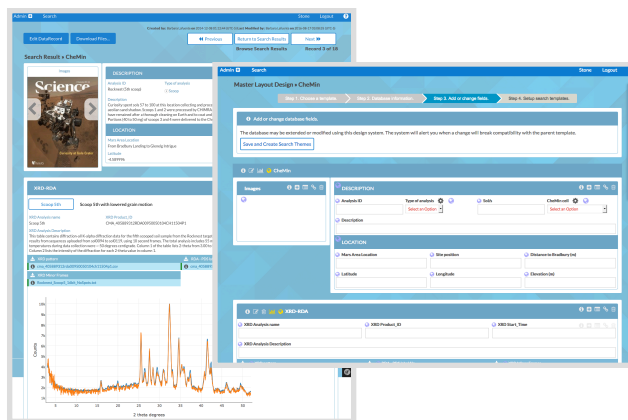


Figure 1. Examples of a data record view (left) and design of a database template (right).

Conclusions: Through the ongoing pilots and discussions with individual researchers and small research teams, a definition of the tools desired by these groups is coming into focus. As the software development moves forward, the goal is to meet the publication and collaboration needs of these scientists in an unobtrusive and functional way.

References: [1] Lafuente B. et al. (2016) AGU 2016, San Francisco, CA 12-16 Dec. [2] Lafuente B. et al. (2017) Abscicon 2017, submitted.

Acknowledgment: We gratefully acknowledge the support for this study by the Science-Enabling Research Activity (SERA) and NASA NNX11AP82A, Mars Science Laboratory Investigations, and University of Arizona Geosciences.