Abstract

Funder mandates, publisher expectations, and open data initiatives are motivating factors for researchers to identify data sharing infrastructures that will provide permanent access to and long-term preservation of their research data. In addition, there is great interest in library science and information science communities to use shared research data in other contexts. Where might institutional data repositories fit in this picture? A use case involves the Purdue University Research Repository (PURR). Research datasets produced through collaborative project spaces within PURR are published with a permanent Digital Object Identifier (DOI) and become available for discovery and citation world-wide. PURR is now developing capacity for other discipline content service providers to access PURR data to include in their content management systems. The use case involves two archeological content systems, Open Context and Visible Past.

Open Context publishes archeological data contributed by researchers through an editorial review process and aligns data to standards, such as Open Linked Data. Visible Past is an interactive historical atlas enhanced with narratives, pictures, video clips, and audio files. It is also a site where researchers can collaboratively work to produce scholarly products, such as monographs and papers. Visible Past and Open Context are linked via Open Linked Data. Visible Past harvests records from Open Context to include in its content management system.

PURR is preparing to publish archeological datasets from a Rough Cilicia Archeological Survey, 2003 Pottery Sherd Study. Each dataset will include the image(s) of a pottery sherd and a MODS metadata file serialized in XML that contains all the data about the pottery sherd. The MODS file will be accessible for Open Context to harvest, programatically manipulate, and map to their systems’ CIDOC-CRM schemata. The processes to be implemented will enable Open Context to open PURR’s downstream humanities data flow to Visible Past.

Methods

Open Context used ArcheoML to serialize object metadata in XML. ArcheoML is no longer maintained. Open Context switched to a conceptual reference model, CIDOC-CRM.

Result

Align PURR’s pottery sherd data to CIDOC-CRM.

Conclusion

PURR currently makes descriptive Dublin Core metadata available via OAI-PMH. However, we are interested in exposing more than descriptive metadata. We will develop more robust mechanisms to make available serialized metadata for ingestion in service providers’ content management systems.

Future Directions

PURR Archeological Data Serialized in MODS Metadata XML File

Resources

Purdue University Research Repository (PURR)
https://purr.purdue.edu/
Open Context
https://opencontext.org/
Visible Past
http://visiblepast.net/see/
CIDOC-CRM
http://www.cidoc-crm.org/
MODS Metadata
http://www.loc.gov/standards/mods/