



Metadata Group Working to Serve Members

August 24, 2016

By Drew Krewer

The Preservation Metadata Standards Working Group (PMSWG) is working to outline the metadata standards that DPN will follow for preservation. Their primary deliverables are to define a set of fields that will remain constant and available with all deposits made into DPN and to document the process leading to the selection of these standards. Thus far, the PMSWG's work has centered around three areas: refining the DPN BagIt Specification, identifying an initial use case, and creating a working concept of the "core" metadata record.

At the beginning of our work, we were informed that a specific institution was storing vital ingest metadata in tag files. The BagIt File Packaging Format Specifications (V0.97) detail that non-required tag files do not have to be listed in a bag's tag manifest, and content of non-declared tag files will be ignored. However, if the tag files are included in the tag manifest, these "optional" tag files are able to undergo checksum creation.

Due to the clear use case for ensuring the fixity of ingest metadata stored as tag files, the PMSWG made the recommendation that metadata stored as tag files must be announced in the tag manifest to ensure that checksums are created for this important content.

After making the decision regarding the BagIt Specification, we looked toward use cases and metadata fields. Of the eight working group members, Jennifer Mullins of Dartmouth College had a specific use case available for analysis. The guiding principle of the case is that each ingest package must be packaged in a way that the contents can be retrieved in one, twenty, or fifty years' time, and that the contents of the package can be understood, verified, and used. At this juncture, the working group collaborated to create initial "core records," based on the metadata fields currently in use at the six educational institutions involved with the PMSWG. The group then applied the use case to the fields employed at the six institutions, noting the most commonly used fields, as well as fields that were missing but would be valuable in fulfilling the requirements of the Dartmouth College use case. From this exercise, we have an idea of the extent of a core record and how it may look.

To move further in our work and to ensure its relevance, the PMSWG recognizes the need to expand our involvement with the DPN community. To that end, the DPN PMSWG is preparing to make a call for members to participate in a focus group designed to help us understand the diverse use cases of DPN members and to identify how the PMSWG can best support preservation efforts. A formal announcement with more details is forthcoming. Please contact Drew Krewer for more information or to get involved. The working group is excited about the tasks ahead and is looking forward to collaborating with members to better serve the DPN community!

Drew Krewer is the Digitization Services Coordinator at the University of Houston and serves as the Chair for DPN's Preservation Metadata Standards Working Group. He may be reached at ajkrewer@uh.edu.