

BESSIG Meeting Wed, May 28, 3:30 - 5:00, Gondolier

★ We have a new venue! We'll be meeting at the Gondolier, <http://www.gondolierboulder.com>. The restaurant is in the Meadow's Shopping Center, on the southwest corner of Baseline and Foothills. We'll meet in their back room.

The Gondolier is providing us this space for free. In searching for a new venue, I have learned how rare and special it is for a business to provide this. Please come and order something so that the Gondolier can continue to offer us this space.

Who's Afraid of File Format Obsolescence?

Evaluating File Format Endangerment Levels and Factors for the Creation of a File Format Endangerment Index

Heather Ryan, University of Denver Library and Information Science

Much digital preservation research has been built on the assumption that file format obsolescence poses a great risk to the continued access of digital content. In an endeavor to address this risk, a number of researchers created lists of factors that could be used to assess risks associated with digital file formats. My research examines these assumptions about file format obsolescence and file format evaluation factors with the aim of creating a simplified file format endangerment index.

This study examines file format risk under the new lens of 'file format endangerment,' or the possibility that information stored in a particular file format will not be interpretable or renderable in human accessible means within a certain timeframe. Using the Delphi method in two separate studies, this exploratory research collected expert opinion on file format endangerment levels of 50 test file formats; and collected expert opinion on relevance of 28 factors as causal indicators of file format endangerment.

Experts expressed the belief that generally, digital information encoded in the rated file formats will be accessible for 20 years or more. This indicates that file format experts believe that there is not a great deal of short-term risk associated with encoding information in the rated file formats, though this does not preclude continued engagement with preservation activities for these and other file formats. Furthermore, the findings show that only three of the dozens of file format evaluation factors discussed in the literature exceeded an emergent threshold level as causal indicators of file format endangerment: 'Rendering Software Available,' 'Specifications Available,' and 'Community/3rd Party Support.' Consequently, these factors are ideal candidates for use in a simple file format endangerment index that can be used to assess endangerment levels of any file format.

The findings of this study have implications for further exploration of file format endangerment in specific digital information creation domains. In particular, applying this model to file formats created by and used in the Earth and Space Science communities will both strengthen the model and will produce valuable insight into format-centric Earth and Space Science data creation and management practices. This insight can then be applied to risk assessment and subsequent actions to support continued access to datasets over time.

Come join us!