Pre-Accession Preservation Appraisal Report

About:

The pre-accession preservation appraisal report is intended to offer guidance to curators on digital preservation concerns prior to acquiring content. The Digital Preservation Coordinator (DPC) may be asked to assess a collection of born-digital content and/or reformat storage so that a curator has additional information which may influence the decision to acquire the content.

The report will be written at a high level, highlighting some of the observable preservation concerns. Note that this report should not be considered a preservation or project plan, but may inform future plans should the content be acquired. The DPC will not analyze the intellectual content of the files or assign research value to the collection.

Sampling:

The report may be conducted on a sample of the content.

A sample will likely be used in cases where:

1.) Vended support is required to access the media

2.) The number of pieces within the collection is greater than ten

3.) There are several media formats included in the collection

3.) The media is in poor condition

4.) Early attempts at accessing media prove excessively difficult and time-consuming

Digital Preservation Assessment:

Content will be assessed based on:

1.) Identifying long-term preservation challenges including:

1. Digital storage space required, including provisions for redundancy while processing the collection
2. Intellectual property rights concerns. A high-level analysis of observable possible rights concerns. An item-level assessment will not be performed.
3. Sensitive data concerns. A high-level review for sensitive data may be conducted.
4. Impediments to accessing the content. This may include:
    1. Content stored on media that is not accessible by the Born Digital Media Reformatting lab
    2. Content stored on media that is damaged or appears to be in poor physical condition

3. Content locked in proprietary or unusual file formats (file format will weigh heavily in the assessment; see item #2.)
4. Unusual file systems/operating system environments/data encoding
5. Little to no information about the environment in which the content was created including:
    1. computing environment
    2. software and specific software versions used to create the content
    3. in the case of data or complex digital objects, lack of codebooks or descriptive information about the data/object

2. **File format analysis**

Content rendering challenges are often more resource-intensive and costly than recovering content from external media for bit-level preservation. High-level analysis of file formats can be indicative of the need for greater resources in the future to render the files. File formats will be interrogated at a high-level based on some of the questions below:

1. Is the content stored in proprietary file formats which may be difficult to access?
2. Are there many files in unusual formats?
3. Are there many files without easily identifiable file extensions?
4. What is the date range of the files being assessed?
5. What steps might be necessary to access the content? This may include:
    1. Review of a sample of the content in a hex editor
    2. Acquiring software
    3. Purchasing software
    4. Migrating content from its original file format to one deemed acceptable within our repository for long-term preservation
    5. Emulating the content

Document Version History

Date	Version	Author	Change Notes
2014-04-10	1	Popp, Tracy	First draft
2020-04-15	1.25	Popp, Tracy	Edited for clarity; added text about sampling.