What is Preservation Storage?

Andrea Goethals
Harvard Library
90 Mt. Auburn St.
Cambridge, MA 01970 USA
+1-617-495-3724
andrea_goethals@harvard.edu

Steve Knight
National Library of New Zealand Te
Puna Mātauranga o Aotearoa
Cnr Molesworth & Aitken St
Wellington, New Zealand
+64 21 490 503
steve.knight@dia.govt.nz

Jane Mandelbaum
Library of Congress
101 Independence Ave
Washington, DC 20540 USA
+1-202-707-4894
jman@loc.gov

Nancy McGovern
MIT Libraries
77 Massachusetts Ave (14E-210)
Cambridge, MA 02139 USA
+1-617-253-5664
nancymcg@mit.edu

Gail Truman
Truman Technologies
4096 Piedmont Avenue, Suite 217
Oakland, CA 94611 USA
+1-510-502-6497
gail@trumantechnologies.com

Kate Zwaard
Library of Congress
101 Independence Ave
Washington, DC 20540 USA
+1-202-707-5242
kzwa@loc.gov

Sibyl Schaefer
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093 USA
+1-858-246-0741
sschaefer@ucsd.edu

ABSTRACT

Storage is arguably the most fundamental building block of any technical infrastructure designed to preserve digital content. All institutions responsible for preserving digital content must use storage systems or services, whether they are maintaining their own, sharing infrastructure with other institutions, or relying on a third party to provide this service. There are many decisions to be made as to what constitutes a storage infrastructure good enough to protect the integrity and viability of the content, at a cost affordable to the institution. To date each institution has had to independently make these decisions without guidance from the digital preservation community. This workshop will explore whether or not it is possible for the attendees to write a brief "principles of preservation storage" as a starting point for a larger discussion within the community after the conference.

Keywords
Digital Preservation Storage; Preservation Best Practices

1. PRESERVATION STORAGE
Storage is arguably the most fundamental building block of any technical infrastructure designed to preserve digital content. All institutions responsible for preserving digital content must make use of storage systems or services, whether they are maintaining their own, sharing infrastructure with other institutions, or relying on a third party to provide this service. All storage is not equivalent. There are many decisions to be made as to what constitutes a storage infrastructure good enough to protect the integrity and viability of the content, at a cost affordable to the institution. To date each institution has had to independently make these decisions without the help of guidance from the digital preservation community. This workshop will explore whether or not it is possible for the attendees to write a brief "principles of preservation storage" as a starting point for a larger discussion within the community after the conference.

2. WORKSHOP
Prior to the workshop, the authors of this proposal will use mailing lists and social media to solicit requirements documents, RFPs, or other documents that contain institutional requirements for preservation storage. The authors will aggregate the requirements found in the documents, normalizing and de-duplicating them into a single list of unique requirements that could potentially be applicable to any institution preserving content. This list will be transformed into a short Qualtrics survey asking respondents to categorize each requirement as either very important, somewhat important, somewhat unimportant, or not important. Those who sign up for the workshop when they register for the conference will be asked to fill out this survey. The authors will analyze the results to prepare for the workshop discussion. Some of the key points of the discussion will be:

- Which requirements do most agree are very important?
- Which requirements do most agree are not important?
- Where is there a lot of disagreement, e.g. requirements that many thought were very important and many thought were not important? What are the reasons on both sides?
- For which requirements does the group have vague opinions, e.g. most thought it was somewhat important or somewhat unimportant?
- What additional conditions influence requirements? For example, the extent and nature of content types, capacity/shear volume or size of individual "blobs"/file sizes, country or geographic law or industry regulations?

After the workshop discussion the authors will produce a white paper describing the results of the survey and summarizing the workshop discussion.