

A Preservation Compass

finding digital preservation partners and solutions



Overview

30 min: Context and Survey Results

- Context and definitions
- Survey results

60 min: A Preservation Compass conversation

- Open discussion



Common Threats to Digital Content

Viruses

Accidental erasure

Overwritten files

Theft

File Corruption

**Storage device
malfunctions**

**LOST PASSWORD
OR KEY**

**Physical
disasters**

Hacking

*Malicious
deletion*

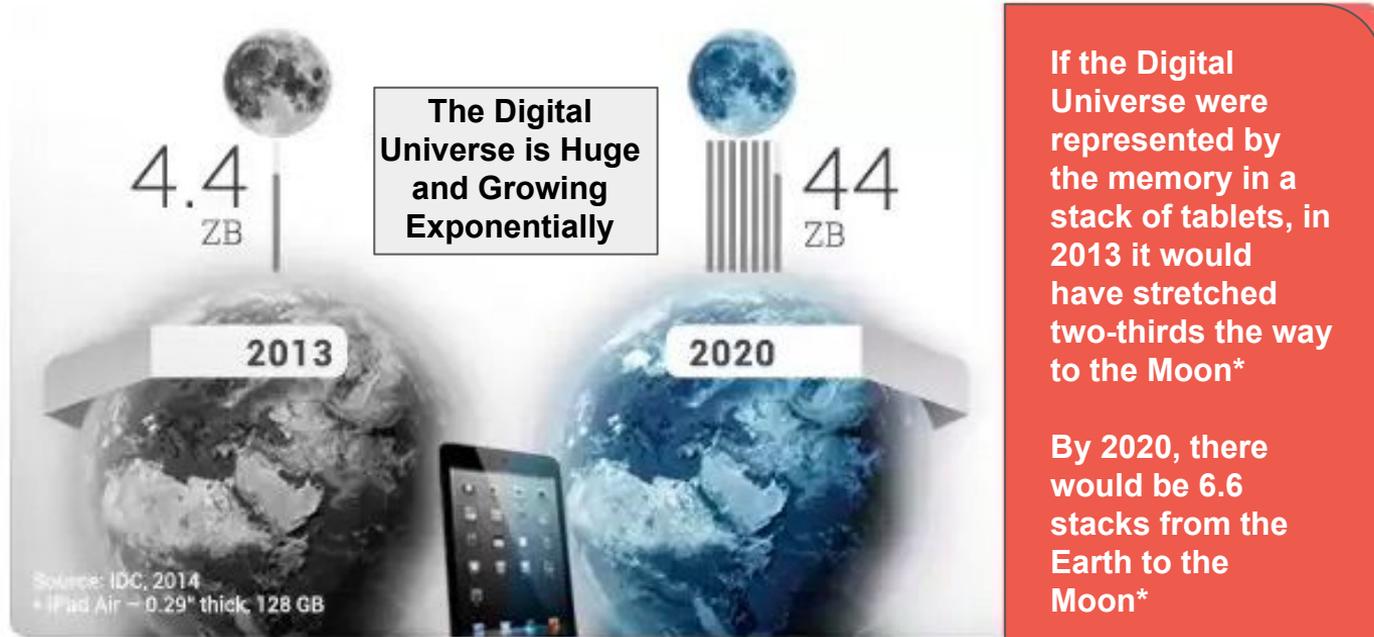


Trio of challenges

- Technical
- Social
- Financial

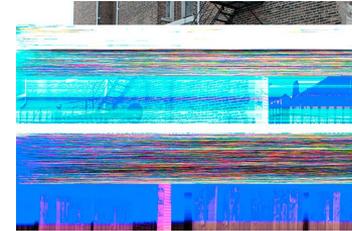
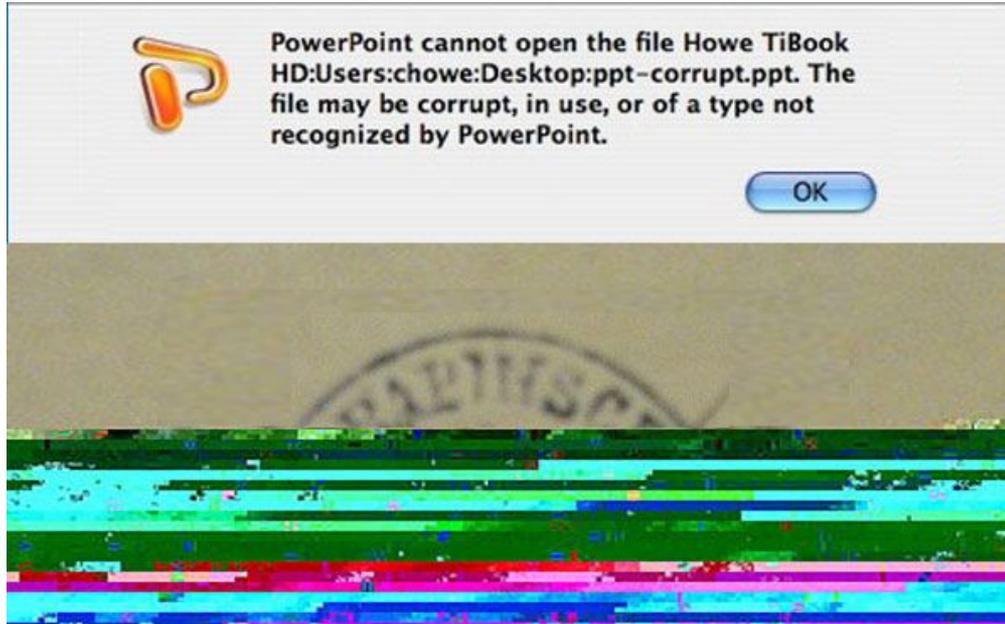


Technical Challenges



Growth of data (and proportional growth of preservation- candidate data)

Technical Challenges (cont'd)



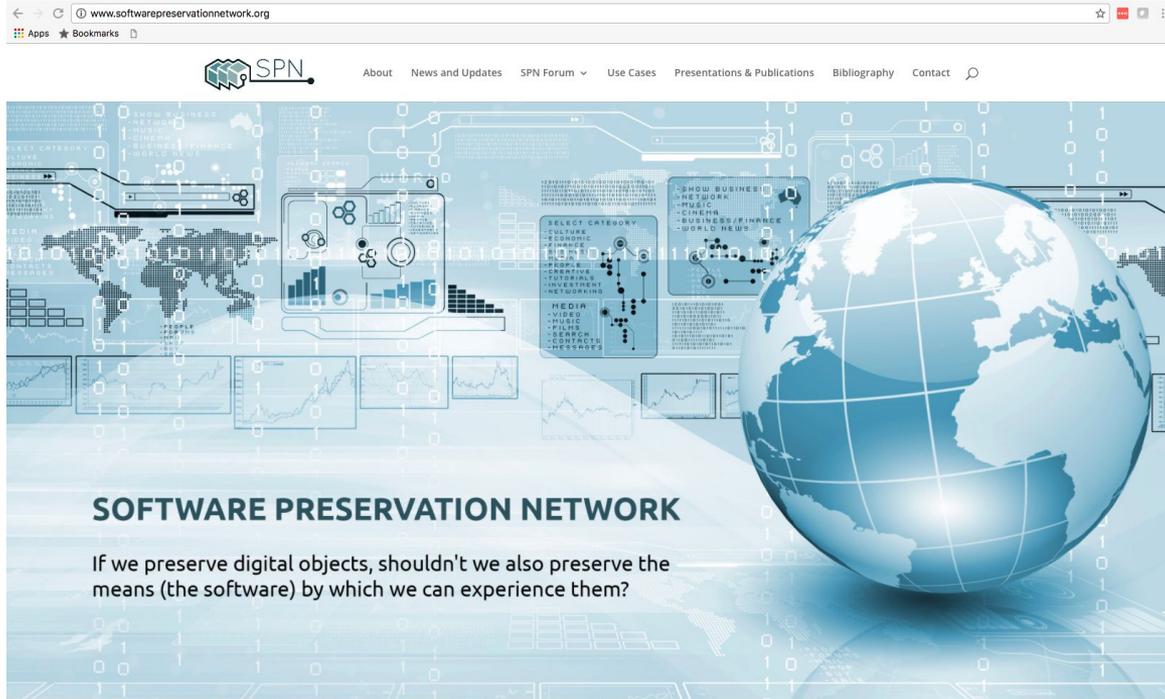
One bit change can render an entire object unintelligible

Technical Challenges (cont'd)



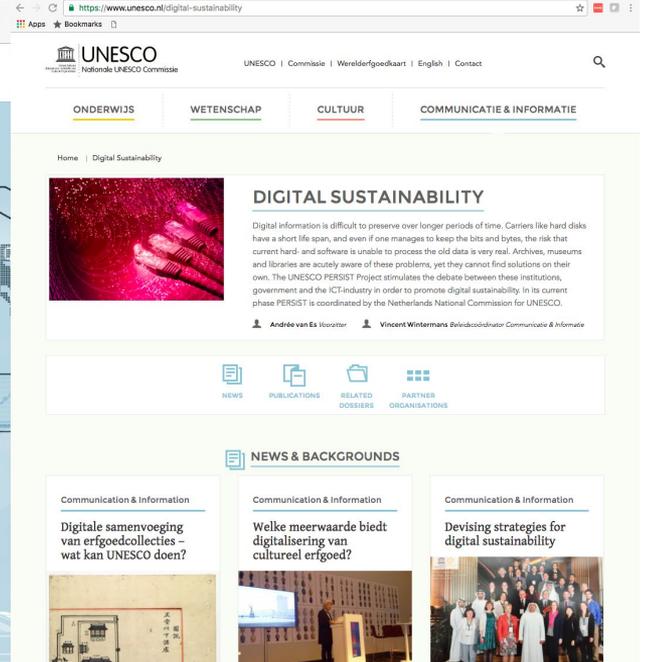
Obsolescence is swift and painful

Technical Challenges (cont'd)



SOFTWARE PRESERVATION NETWORK

If we preserve digital objects, shouldn't we also preserve the means (the software) by which we can experience them?



UNESCO
Nederlandse UNESCO Commissie

UNESCO | Commissie | Werelderfgoedkaart | English | Contact

ONDERWIJS | WETENSCHAP | CULTUUR | COMMUNICATIE & INFORMATIE

Home | Digital Sustainability

DIGITAL SUSTAINABILITY

Digital information is difficult to preserve over longer periods of time. Carriers like hard disks have a short life span, and even if one manages to keep the bits and bytes, the risk that current hard- and software is unable to process the old data is very real. Archives, museums and libraries are acutely aware of these problems, yet they cannot find solutions on their own. The UNESCO PERSIST Project stimulates the debate between these institutions, government and the ICT-industry in order to promote digital sustainability. In its current phase PERSIST is coordinated by the Netherlands National Commission for UNESCO.

André van Es *hoofdstor* | Vincent Wintermans *Beleidscoördinator Communicatie & Informatie*

NEWS | PUBLICATIONS | RELATED DOSSIER | PARTNER ORGANISATIONS

NEWS & BACKGROUNDS

Communication & Information

Digitale samenvoeging van erfgoedcollecties – wat kan UNESCO doen?

Communication & Information

Welke meerwaarde biedt digitalisering van cultureel erfgoed?

Communication & Information

Devising strategies for digital sustainability

The software that renders digital objects is in need of preservation attention, too!

Social Challenges (cont'd)



We cling to those that look like us (sometimes to our detriment).

Financial Challenges



Digital preservation is a new cost, not counterbalanced by reduction in any other cost

One institution's view: MIT Library Planning

“Sustainable digital preservation is an important problem shared by the entire scholarly community, and MIT should continue to build partnerships and provide leadership to help solve the inherent technical, social, and economic issues in this area. Digital stewardship is a collaborative, inclusive effort that includes creators, users, champions, content managers, and preservers... Any sustainable vision for the future of research libraries must include a commitment to the long-term stewardship of the scholarly record.”

*--Preliminary Report of the Institute-Wide Task Force on the Future of Libraries
Published October 24, 2016, captured from <https://www.pubpub.org/pub/future-of-libraries>, November 23, 2016*

What is Digital Preservation?

Digital Preservation is the active management of digital content over time to ensure ongoing access.

Library of Congress, <http://www.digitalpreservation.gov/about/index.html>,
(November 26, 2016)

Ingest Format Validation Audit Storage Access

Repair Data Wrangling Fixity Checking Trust

Geographic Replication Metadata Testing Rights (etc...)

What is Digital Preservation? (cont'd)

Distributed Digital Preservation emphasizes the importance of such factors as content **replication**, **independence**, and **coordination** for ensuring the longevity of digital objects

Key: geographic distribution,
infrastructure heterogeneity,
organizational diversity



Survey says...

[Survey-results](#) may suggest topics for further discussion

Discussion

Katherine Skinner, katherine@educopia.org

Chip German, cgerman@virginia.edu

END of ASERL Presentation

Additional information related to ASERL discussion
(next)

Tools and Strategies

Emergence of definitions, concepts, best practices, standards (last 20 years)

- NDSA levels of preservation
 - http://www.digitalpreservation.gov:8081/ndsa/working_groups/documents/Levels_v1.pdf
 - Updates needed on format migration, software preservation and other components
- Getting to the Bottom Line: 20 Cost Questions
 - <http://metaarchive.org/cost-questions>
 - Provides guidance on what questions need to be asked of potential service providers or collaborative partners

Tools and Strategies (cont'd)

Pillar 1: Accept the inevitability of needing multiple solutions, and favor the collaborative ones

- Collaborations to avoid duplication of effort in preservation of published digital material
- Collaborations to achieve economies of scale in preservation of unique digital material
- Engagement in collaborative projects creates knowledgeable staff members rapidly and ensures up-to-date skills and understanding
- Collaborations should embrace all memory organizations
 - Many may have limited resources (long-tail dilemma)

Tools and Strategies (cont'd)

Pillar II: develop clear preservation policies/strategies that are connected to general stewardship planning unique to the institution

- Example: evolution of strategy at UVA from rigorously selective to more permissive, especially including material not fully processed

Tools and Strategies (cont'd)

Risk Factor Examples (based on Kara McClurken presentation from UVA)

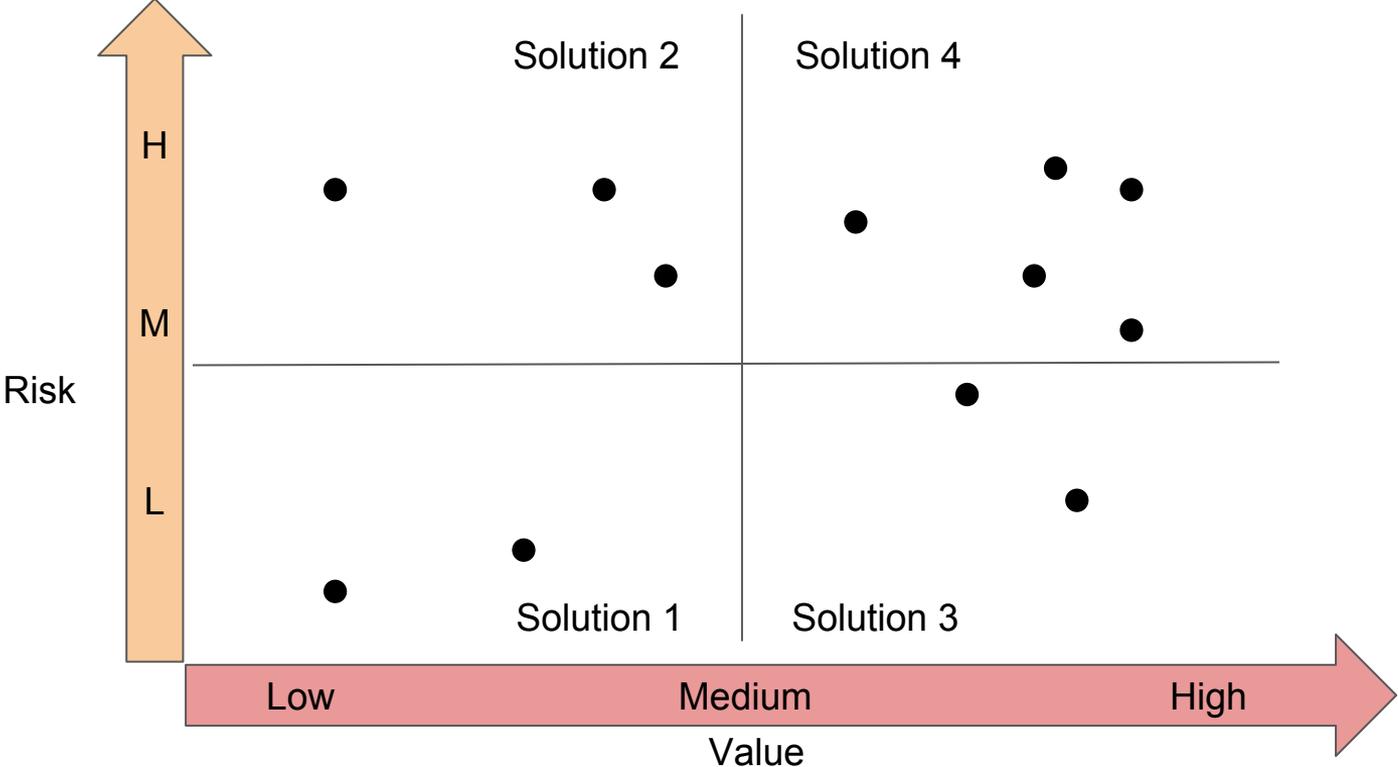
- Uniqueness or rarity
- Not preserved (and not just stored) elsewhere
- Difficulty of recreating (if digitized material, could it be re-digitized?)
- Threat of loss of access
 - Website coming down
 - Company or institution dissolving

Tools and Strategies (cont'd)

Value Factor Examples

- Research value (high weight)
- Importance as cultural heritage (high weight)
- Institution-committed responsibility for content (high weight--examples include promise to donor)
- Documentation of rights status

Preservation Decision Tool



Tools and Strategies...

Pillar III: maintain continuing management of your content in all preservation locations, while anticipating your institution's disappearance

- Desirable capability: revise/update metadata without having to reingest the entire package of preserved material
- Ensure that your metadata includes your understanding of copyright status (<http://rightsstatements.org/en/>)
- Retain the capacity to delete

Hot Topics and Open Questions...

- Funding (if we don't figure this out, who will?)
 - RCM budgeting sometimes exposes true cost
 - The only way to use margin strategies is through multi-institutional collaborations to produce economies of scale (note “capacity” funding vs “content” funding)

| Funding Models for Non-profits | |
|--------------------------------|-----------------------------|
| Heartfelt connector | Policy innovator |
| Beneficiary builder | Beneficiary broker |
| Member motivator | Resource recycler |
| Big bettor | Market maker |
| Public provider | Local nationalizer |
| | Taxer (additional category) |

Hot Topics and Open Questions...

- Funding
 - Selling general concept of digital preservation as good societal practice won't stimulate funding at scale
 - Pushing digital preservation to audiences by making it about content for which they care deeply may
 - Still must make provisions for longtail (culturally important material with no current natural audience)
- Capturing ephemera
 - Web archives
 - Scale challenges description other than chronological
- Explosion of 3D data
- Urgent need of software preservation

To do, or not to do...decision making and partnerships

- The paralysis of uncertainty
 - Still in the earliest stages of development of strategies for digital preservation
 - Ethical bond between higher education-driven digital preservation initiatives means that each is alert to risks to the content of another
 - Likely to result in a self-healing system of higher education projects that assists in relocating content from any one that fails or dissolves
 - Once your data-driven analysis is complete, you will still have choices

Choose the collaborative projects best suited to your institution and circumstances, including in “personality.” Your choice will not be wrong

