Digital Preservation: The Next Library Frontier

Presented at:

IGeLU 2009

Helsinki, Finland
8 September 2009

igelu
International
Group of Ex Libris Users

Edward M. Corrado

Head of Library Technology
Binghamton University
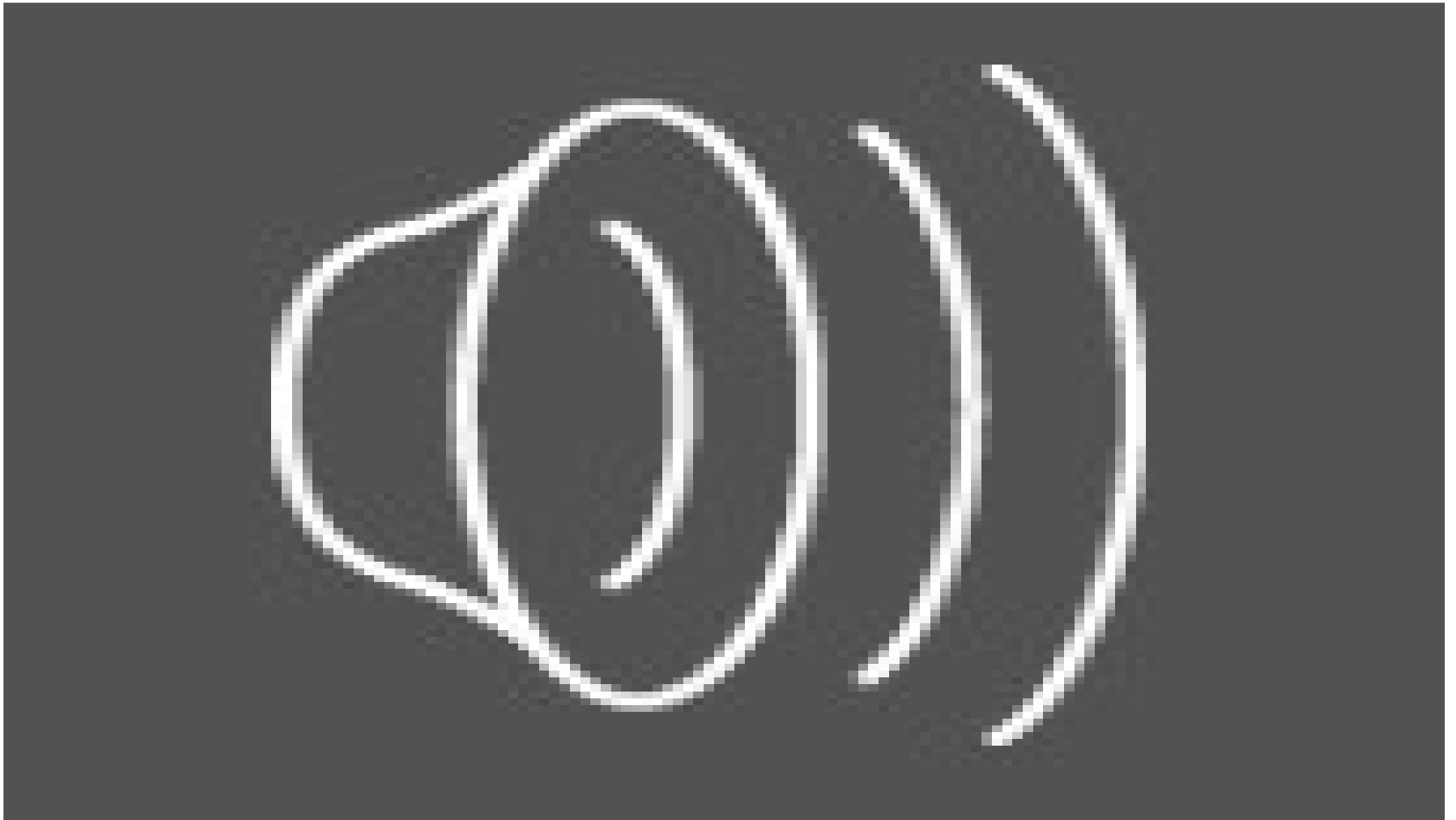Binghamton, NY, USA
ecorrado@binghamton.edu

Libraries & Preservation

Libraries (esp. Academic & National Libraries) have traditional played important roles in preserving information

Manuscripts, Papers, Archives...

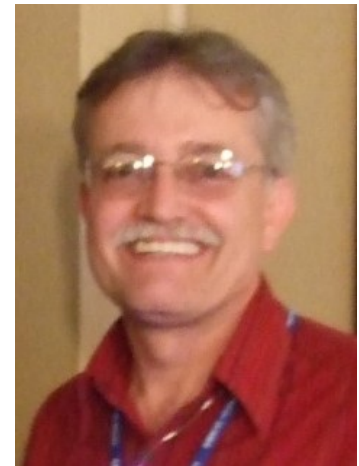# Digital Preservation & Nuclear Disaster: An Animation

What is Digital Preservation?

"Digital preservation combines policies, strategies and actions that ensure access to digital content over time." ALA (2007; Short definition)

"All activities concerning the maintenance and care for/curation of digital or electronic objects, in relation to both storage and access." (Research Councils UK 2008: 6)

"I think this episode also illustrates, once again, that preservation is not about technology at all, it's about *institutional commitment*. ... Without institutional commitment, no technology on earth can save you." - Roy Tennant

Digital Preservation is not:

An Institutional Repository

Computer Storage / Backups

Short term commitment

Only a technical problem

Cheap!

Amazon S3 Storage Costs

$1.80 * 20000   storage
$2.04 * 20000   data transfer

$36,000 year 1 storage (20TB)
$40,800 year 1 data transfer (20TB)
$46,800 year 2 storage (26TB)
$12,240 year 2 data transfer (6TB)
$61,200 year 3 storage (34TB)
$16,320 year 3 data transfer (8TB)

$213,360 over 3 years

Ongoing Costs

Many other ongoing costs. Can't just fund once....

- Electricity

- Environmental

- Format conversion

- Staff

- Conformation (can data still be read?)

"Digital information lasts forever -- or five years, whichever comes first." - Jeff Rothenberg (as quoted in Holtz (2009))
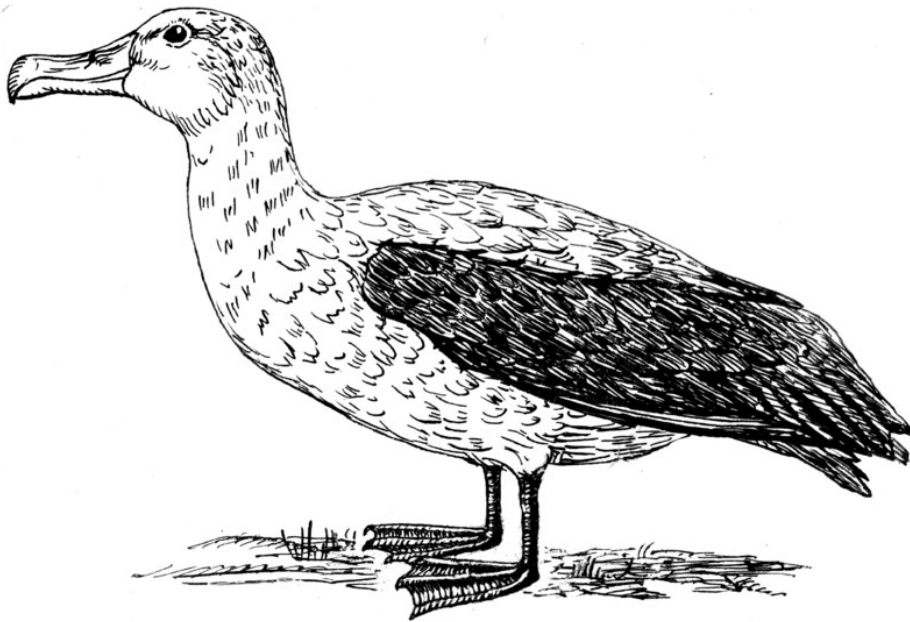
"Digital Preservation as an Albatross"

Digital Preservation is the mechanics (how)
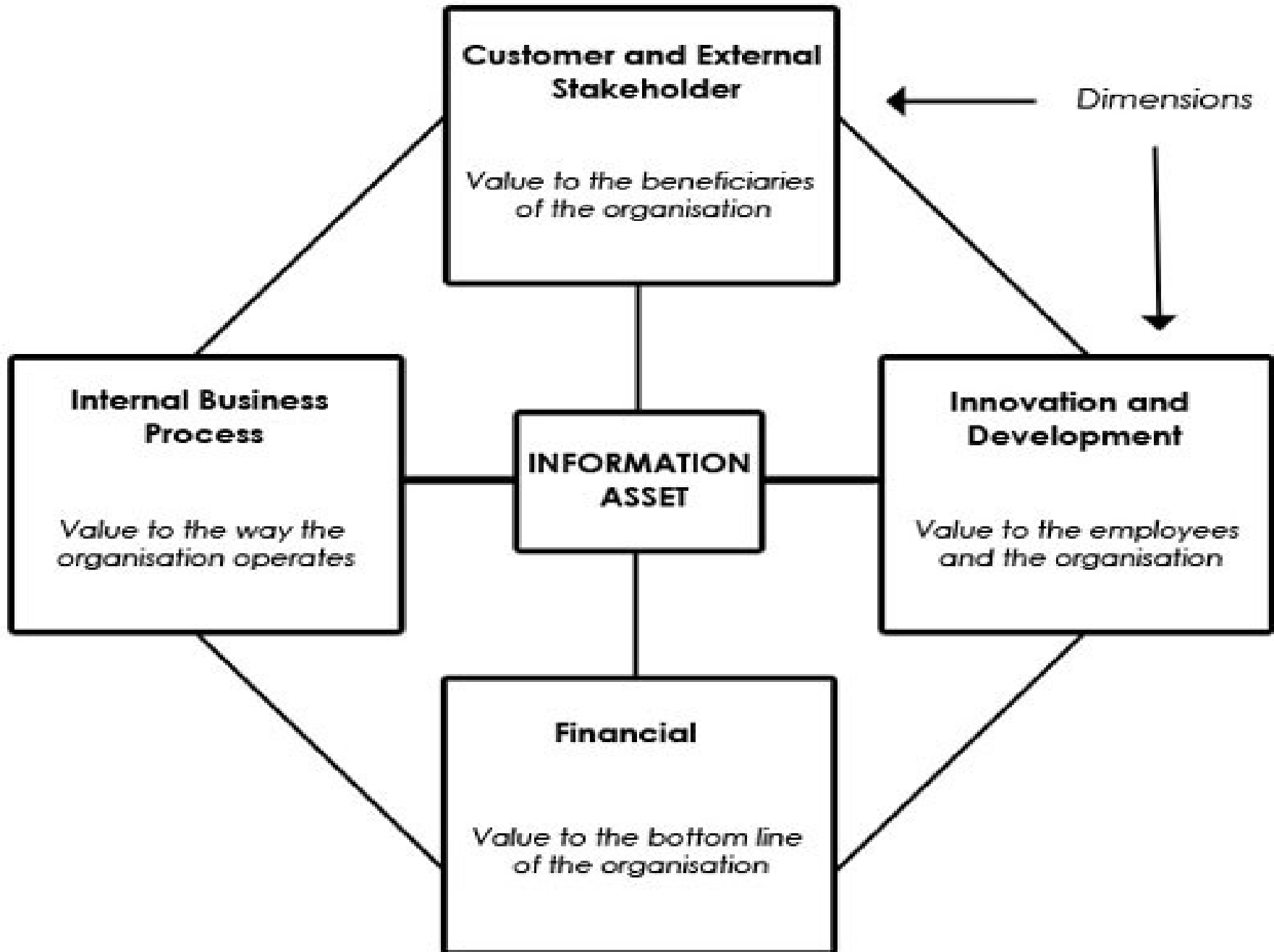
Not which is really important (what)

Not the purpose (why)

Business Case needs to be made

Currall, J., P. McKinely, and C. Johnson (2006)

espida Balanced Score Card

Who needs to Preserve Stuff

Governments (Country, States, Cities, etc.)

Institutions trusted to create and maintain knowledge (e.g. Universities)

Why Libraries?

Expertise in [Print] Preservation

Expertise in Information

What needs to be preserved?

It is easy to say everything, but that is not practical

"Our ability to collect data now outstrips our ability to maintain it for the long run" - William Michener (as quoted in Holtz (2009))

Libraries don't preserve everything in print

"It would be tragic if there were no record of lives that were so influential" – J. L. John (as quoted in Holtz (2009))

What needs to be preserved? (cont'd)

Government Mandates

Unique Items (last copy; only copy)

Institutional memory

Collection Development policies

In other words, not very different than print.

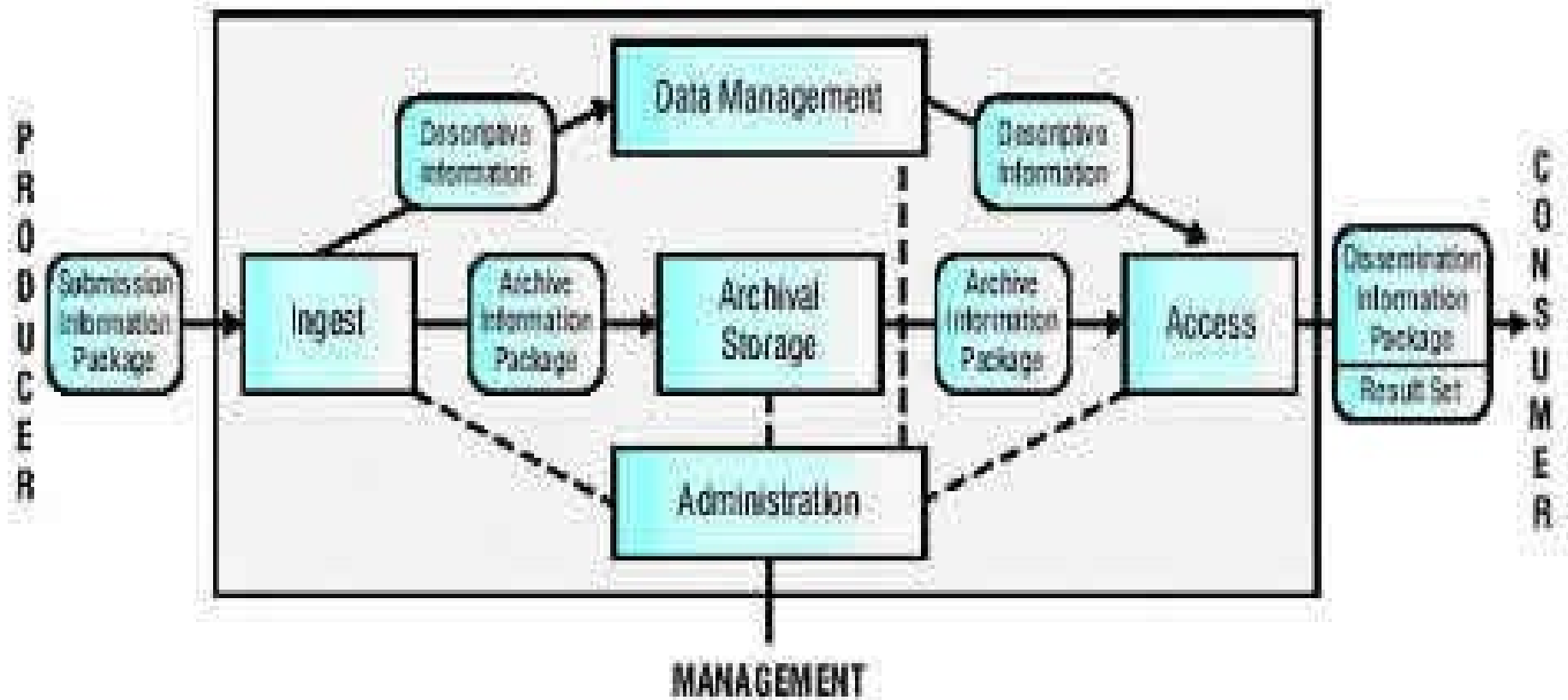# Open Archival Information System (OAIS)

Figure 3

Lavoie, 2000

PREMIS (Preservation Metadata: Implementation Strategies)

Objectives:

Define an implementable set of "core" preservation metadata elements...

Draft a data dictionary to support the core preservation metadata element set

Caplan. P. and Guenther, R. 2005

PREMIS (Objectives Con't)

Examine & evaluate alternative strategies for the encoding, storage, and management of preservation metadata ... as well as for the exchange of preservation metadata among systems

Caplan. P. and Guenther, R. 2005

PREMIS (Objectives Con't)

Conduct pilot programs for testing the group's recommendations and best practices in a variety of systems settings

Explore opportunities for the cooperative creation and sharing of preservation metadata.

Caplan. P. and Guenther, R. 2005

Resources

ALA. (2007. Definitions of digital preservation. Chicago: American Library Association. Available at: http://www.ala.org/ala/mgrps/divs/alcts/resources/preserv/defdigpres0408.pdf

Caplan, P. and R. Guenther, 2005 "Practical Preservation: the PREMIS Experience" Library Trends: 54 (1) Summer 2005 http://www.loc.gov/standards/premis/caplan_guenther-librarytrends.pdf

CCSDS (Consultative Committee for Space Data Systems) (2002). Reference Model for an Open Archival Information System (OAIS). Blue Book, Issue 1. Washington, DC (US): CCSDS Secretariat, January 2002. Technical report. CCSDS 650.0-B-1. Recommendation for Space Data System Standards. Available at: http://public.ccsds.org/publications/archive/650x0b1.pdf.

Currall, J., P. McKinney, and C. Johnson (2006) Digital preservation as an albatross. Archiving 2006 - Final Program and Proceedings, pp. 75-78. http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-36049044810&partnerID=40

Digital Preservation Europe (DPE): http://www.digitalpreservationeurope.eu/

Hotz, R. L. (28 Aug 2009) "A Data Deluge Swamps Science Historians." Wall Street Journal. http://online.wsj.com/article/SB125139942345664387.html

Lavoie, B. (2000) Meeting the challenges of digital preservation: The OAIS reference model. http://www.oclc.org/research/publications/archive/2000/lavoie/

Tennant, R. (1 Sept 2009). GPO Purls (Code4lib e-mail list message): http://serials.infomotions.com/code4lib/archive/2009/200909/1443.html