Audit & Certification of Digital Repositories: Exposing Risk & Fostering Collaboration
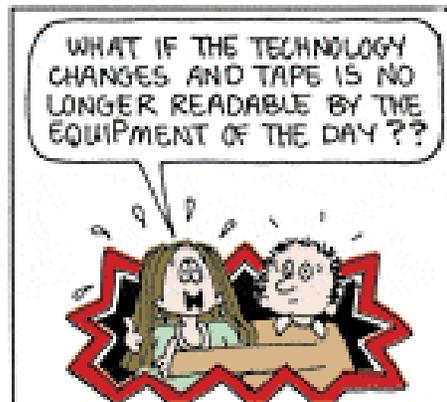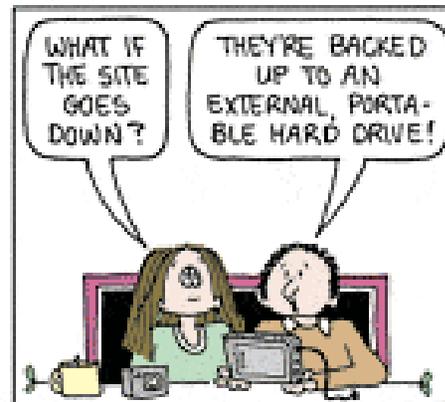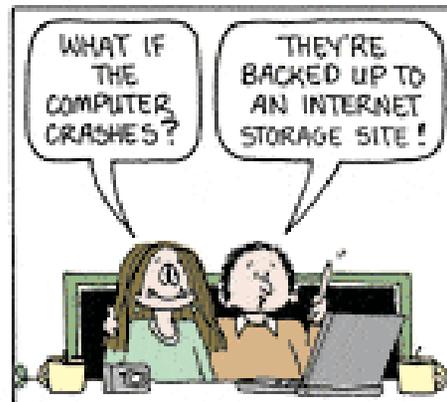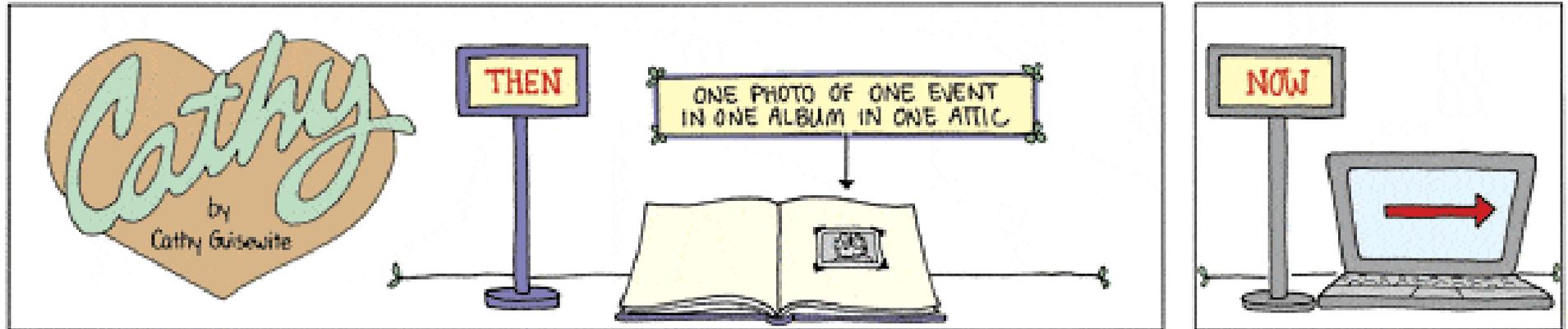
Robin L. Dale

9 October 2006

iPRES 2006: International Conference on the Preservation of Digital Objects

OCLC Research

Overview

- Audit & certification: understanding repositories and assessing risk

- Envisioned uses of a checklist

- Checklist development (RLG-NARA, CRL)

- Emerging audit & certification process

- Potential process outputs, and compliance issues

Understanding & Assessing Risk: a semantically confusing scene

- Digital repositories
- Digital archives
- Data archives
- Institutional repositories
- Digital object management systems
- Digital asset management systems

➢ *de facto* network of varying development, capabilities, content, intent/services, data longevity, vulnerabilities, horizons/sunsets

OCLC Research

Envisioned Uses of a Checklist

- Repository/Archives Planning
  - Set of minimal characteristics for a digital *preservation* repository
- Self Assessment
  - Gap analysis
  - Risk assessment
- Audit
  - Organizations/Institutions
  - Services
- Basis for certification

OCLC Research

RLG-NARA Checklist Development

- **TDR as a *system*, not just technical bits**
  - Organizational infrastructure
  - Digital object management
  - Technical infrastructure

- **Repositories will differ, but good practice and reliable technical infrastructure are universal**

- **Audit as objective evaluation; risk assessment, evidence of capabilities, basis for trust**

- **Focus on transparency**
  - i.e., policies, procedures, content, vulnerabilities

OCLC Research

Moving Theory into Practice: the CRL Project

- Center for Research Libraries (CRL) *Auditing & Certification of Digital Archives* project
- Leveraged RLG-NARA checklist & contributed findings to revisions
- Test audits of three archives/repositories & LOCKSS distributed archiving system
- Audits to determine…
  - Technologies & technical infrastructure
  - Archive processes and procedures
  - Characteristics that affect performance, accountability & business continuity

CRL Project Issues

- Audit process development
  - Additional questions (beyond checklist)?
  - Evidence required
  - Personnel required
  - Institutional investment required

- Reporting Out
  - What constitutes a report?
  - Audience?
  - Who has access to outputs?
    - Information confidentiality issues
    - Community's/depositor's/funder's "need to know"
    - Trusted network partners

Reporting Out: CRL Report Design

- Provost/CIO statement
- Executive Summary
- Full Report
  - Introduction
    - Institutional overview
    - Scoping the audit
  - Findings
    - Organizational Analysis
      - Governance, Staff, Policies & Procedures, Financial, Succession planning

- Technical Analysis
  - Architecture, Data Security, Data Ingest, Archival Storage, Preservation Planning, Data Management, Access Management, Business Continuity, Environmental Management, Disaster Planning
- Content Analysis
- Vulnerabilities
- Observations & Recommendations

OCLC Research

Some CRL Project "Discoveries"

- **Compliance Issues**
  - Is it a true checklist?
  - Levels of compliance are more appropriate
  - Audit compliance versus certification grading
- **Business model(s) for audit & certification**
  - Based on audit audience
  - *Partially* based on geopolitical influences
  - Based upon institutional willingness (investment)
  - Possible model for audit yielding public review

Certification as a Global Effort

- A trustworthy system of repositories, archives, services
  - Data archives, academic content, e-journal archives, government archives, etc
- Common, objective metrics (principles?) for policies, practices, and technical capabilities
- Provisions for evidence, competency, reliability, sustainability
- Collaboration across continents, nations, "specialty" repositories and archives, initiatives

➢ *Cannot rely solely on a de facto network!*

OCLC Research

Audit & Certification "Boiled Down"

- Regularized audits of digital repository capabilities, infrastructure & policies combined with defined levels of compliance
  - Make life predictable
  - Create a community
  - Make it easier to collaborate
- TDR vs. DGR

    Trustworthy Digital Repository

    vs.

    Darn Good Repository

Questions?

Thank you.

Robin_Dale@oclc.org