Tailoring Certification to User Requirements:

Applications of the **Trustworthy Repositories Audit and Certification Checklist** (TRAC)**, Instrument and Methodology**

Community of Interest

- CRL's 230 members from universities, colleges and independent research libraries are a market for digital content
- Everything from e-Journals to electronic archiving services like Portico and CLOCKS.
- CRL institutions represent researchers and scientists

Librarians and archivists are the watchdogs

They must identify and articulate the needs and interests of future users, such as historians

They advocate for the interests of overlooked market segments, such as citizens of the developing world

TRAC

- Trustworthy Repositories Audit and Certification: Criteria and Checklist
- Final report is available here:
- http://biblpurl.oclc.org/web/16712

Andrew W Mellon Grant

- May 2005, CRL was awarded a grant to create a foundation to develop the processes and activities required to audit and certify digital archives
- The RLG/NARA Task Force on Digital Repository Certification, *Audit Checklist for the Certification of Trusted Digital Repositories*

Repositories Audited

- ICPSR (Inter-university Consortium for Political and Social Research)
- Portico
- Koninklijke Bibliotheek (KB) – Elsevier journals repository

What is TRAC for?

- Auditing and certification of digital repositories
- A template for looking at the organization & practices of digital repositories
- The criteria can be applied to many types of repositories

TRAC users

- Digital Repositories
- Publishers
- Librarians
- Researchers and the general public

TRAC's value is in its flexibility

- The rigor of an audit will vary
- Required documentation will change
- Testing of criteria will be more demanding
- A user group may require greater oversight of a repository

TRAC criteria are not fixed

<TRAC A3.1> The repository has:

a. defined its designated community,

b. [defined its] associated knowledge base's, and

c. has publicly accessible definitions and policies in place to dictate how its preservation service requirements will be met

TRAC asks the repository to identify its own standards

<TRAC B1.1> Repository identifies properties it will preserve for digital objects

In an Audit compliance is relative

The repository is compliant when it defines the needs of its community and then meets those needs

Audit criteria are adjusted to the community's requirements

The audit takes into account the value of the repository to it's user population

Market Forces

The auditing and certification of digital repositories must contend with market forces

The marketplace for auditing and certification of digital archives

- There is no one to pay for audits
- No one is requiring audits
- There is no organizational body to certify the auditors
- There is no demand for audits from university's

We have a number

CRL findings from our Andrew W. Mellon Certification of Digital Archives Grant:

- The average cost for an audit was $65,000
- the costs of an audit will move upward, depending on the level of importance placed on the audit

With auditing not yet required, does TRAC have a role to play in to the Digital Repository World?

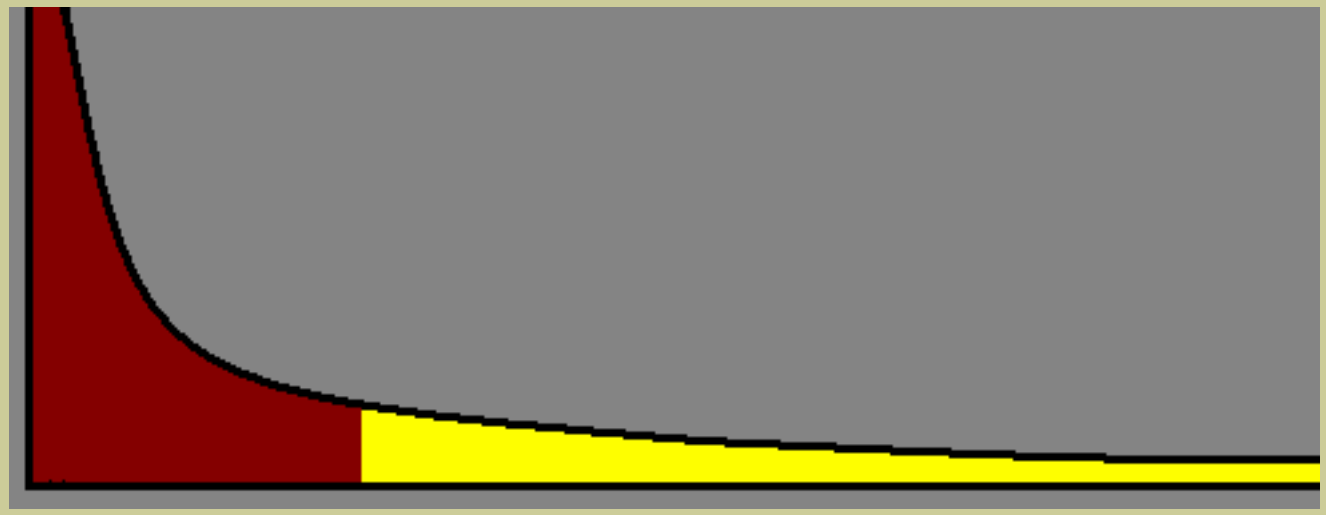
We will look at three different types of repositories for more information

- News Repositories
- Electronic Journal Repositories
- Science Data Repositories

News has a value in the new Internet economy

The Long Tail business model –

There's business potential in the yellow area



The cost for news is no longer stable and is rising for many libraries.

…In fact, many library's can no longer afford Thompson Gale's *World News Access* or Proquest's *Historic Newspapers*

News Repositories

Libraries are no longer archiving news

Today news organizations maintain their non-current materials

BBC, Frankfurter Allgemeine Zeitung Electronic Media, and the AP distribute their news content via the Web

News is old after 90 days

Here are our news archives:

NEWSBANK

LEXIS/NEXIS

FACTIVA

(Is Google far behind?)

Look at the results

.

Electronic Journal Repositories

The emergence of the third-party electronic journal repository

- E-Depot
- Portico
- Ohiolink

Electronic journal repositories are growing fast

Portico has 341 member libraries and 33 Publishers depositing content

Ohiolink just added its 8 millionth electronic journal article

Users are not requiring audits

Maybe this is because there is little to no competition among electronic journal repositories

Certification is not a consideration

Libraries and publishers are currently investing in repositories, regardless of their certification status

When electronic journals proliferate, and there is a choice as to where a publisher might deposit content, then we will see an incentive for auditing and certification.

Science Repositories

Despite large investments from both public and private sectors, not all science data is surviving.

Science Repositories

- Many science data collections are maintained by their community of interest
  - Worldwide Protein Data Bank (wwPDB), Medical community
  - Sloan Digital Sky Survey, Astrophysical Research Consortiums
  - Luxembourg Income Study, participating government agencies

Community ownership does not guarantee sustainability

The Archaeological Data Archive Project (1993-2002)

- Insufficient funding
- Insufficient participation

Science repositories are going to be huge

We’ve only begun to see the long-term effects of the "data deluge" which the scientific community is producing.

Protein Bank

The Protein Data Bank began in 1970 with less than a dozen structures, today there are 42,752 Structures.

Today there is an average of 130,000 downloads of this information a day.

Looking to the Future

- The market affects the scale and nature of auditing and certification
- Auditing and certification can be an important lever in the marketplace

TRAC can assist in identifying needs

TRAC goes beyond the basic data security needs, which today satisfy most publishers, and looks at the organizational and preservation needs of libraries

TRAC can help in repository planning

TRAC can define an organizational framework for a repository of any size.

TRAC can assist in communication between interest groups

TRAC criteria give users a framework for communicating their requirement to the repository

TRAC can help identify the costs of archiving data

News agencies, publishers and any community with a stake in archiving information need to understand the cost of good archiving practices

TRAC can address some of the concerns we have expressed regarding digital repositories

- For news, TRAC exposes concerns regarding degraded graphics and text
- For Science data it’s origins in the OAIS reference model are

Further Suggestions

Librarians worldwide can negotiate licensing terms using TRAC and other criteria to articulate preservation standards to publishers

Publishers can negotiate with repositories with whom they contract using TRAC and other criteria to communicate their needs

Wrapping it up

- What do we think?

  For the moment, TRAC is an important tool for guiding librarians, publishers and repositories in the requirements and criteria for digital repositories.

  And it has bright future.

Contact information

- Bernard Reilly
  President, CRL
  reilly@crl.edu

- Marie-Elise Waltz
  Special Projects Librarian, CRL
  waltz@crl.edu