Measuring reliability

Die Messung der Zuverlässigkeit

Prof. Dr. Michael Seadle

Berlin School of Library and Information Science

June 2011

Slides in English

Die Vorlesung auf Deutsch.

Reliability testing matters

The second reactor explosion at the Fukushima nuclear power plant.

Source: Law in Action

Public testing should be open and reproducible.

Semi-public testing should be published but may not be fully reproducible because not all data or systems are open.

Private testing is neither published nor is the data available.

Measuring reliability

Two key problems:

- What is reliability in digital archiving?
- How do we measure it?

Measuring and testing are closely related. We cannot measure without testing.

Why bother?

The current discourse about archiving tends to be proscriptive (we ought to do X).

- Standards are proscriptive, for example.

Or the discourse is about marketing claims.

- system Y does it all -- as long as you pay.

We need is evidence in the form of public testing to measure not merely whether an archiving system is (on binary scale) reliable or not reliable, but (on a continuous scale) how reliable.

What to measure?

Key issues in long term digital preservation that need testing are:

- **Integrity** - will be bitstream be unchanged in 100 years?
- **Authenticity** - can the origin and genuineness be shown?
- **Usability** - can migration or emulation be demonstrated?
- **Access** - will the archive system allow access at appropriate times?
- **Financial integrity** - will the support structure be around in 100 years?

No single measurement criteria fits all.

Does auditing suffice?

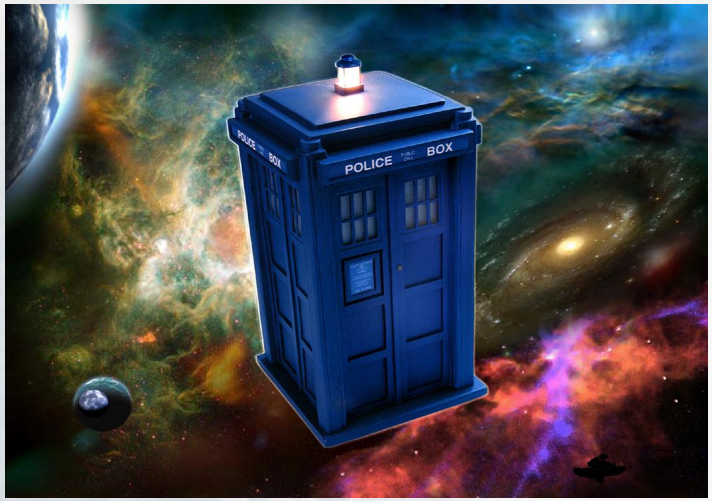
The kind of auditing that the Center for Reserch Libraries does, checks procedures and documentation not performance.

This kind of auditing can flag problems (for example the potential financial problem with Portico), but it uses no active tests to measure performance.

Testing how?

Testing future conditions is hard but not impossible. Engineers stress systems regularly to see how they will perform in the future.

Software engineers also create test data to reflect conditions that they know might occur in the future.



Source: Doctorwho1.com

Dr. Who and his Tardis are not strictly necessary for system testing.

Means for measuring

We have criteria for the reliability of many products and consumer goods.

Stiftung Wahrentest and Consumer Reports use

- explicit tests and
- past experience

to measure reliability.

This testing goes beyond whether a product works. The intent is to measure how it performs.

An example: Bitstream testing

Bitstream integrity testing is arguably the most important kind of testing because without an intact bitstream, authenticity has no meaning and usability may be impaired.

In other words, without bitstream intregity, digital archiving has failed.

Manufacturer claims for product lifetimes should be tested, not believed.

But how can testing be done?

Bitsteam testing 2

"Bit rot" errors can be simulated in a number of ways, both by stressing physical media and with Gedanken-experiments -- known error rates based on specific physical tests can be expanded mathematically.

The key question is: will enough unaltered copies survive in 100 (or more) years?

This will depend on how many copies and how much checking. See David SH Rosenthal's blog for details.

Metrics

In most areas of digital preservation, no reliable metrics exist to measure success.

- How much integrity loss is acceptable?
- How do we measure the authenticity of digital content?
- How do we know when usability has failed?

Until there are tests, there are no benchmarks.

Measuring integrity #1

How much integrity loss is acceptable?

Currently we measure integrity with checksum (hash) calculations. The result is binary: OK or damaged.

The "damage" could be a bit rot problem, a deliberate distortion of the contents, or the equivalent of a marginal comment that adds to but does not change the original.

Measuring integrity #2

With digital forensic tools we can potentially distinguish between these forms of integrity loss -- if enough comparison files exist for a plausible voting process.

Integrity loss of a few flipped bits in the text portion of a document might change only a few letters and be "trivial".

Integrity loss of a few bits in a file header might damage readability and be serious.

Integrity loss due to the addition of a marginal note may not in fact mean any loss to the original content.

Measuring future usability #1

Usability loss can take multiple forms:

- format change
- platform change
- cultural change

Migration and emulation can address the first two and the degree of change there is measurable and testable.

Cultural change is harder.

Measuring future usability #2

Measuring cultural change requires social science tools.

An example of the problem comes from the Michigan/Leeds CAMiLEON project that emulated the 1983 game Chuckie Egg from ZX Spectrum on the BBC Microcomputer.

The end result was no longer especially interesting.

Measuring cultural change requires social anthropology and related methods.

Trust and testing

Traditional physical archiving relies heavily on trusted institutions. This works when the original is well known and not easily changed.

With digital content change is easy. Disgruntled employees represent a major source of computer attacks and damage to data[1]. No single institution is immune.

Distrust, not trust, need to be the basis of digital archiving. Testing plays a key role.

[1]See: Power, R., (2002) "2002 CSI/FBI computer crime and security survey",Computer Security Institute. Available.

Tests & decision-making

With information about:

- the type of storage media,
- the number of copies, and
- the frequency of checking and replacement,

a library could use test results to determine whether an archiving system plausibly can deliver on archiving claims.

Pig in a Poke

Without well-documented, peer-reviewed, publicly available test results, we are buying archiving systems on faith.

That is a poor investment in the future.

Contact information: Michael Seadle: seadle (at) ibi.hu-berlin.de