

CASRAI Reconnect UK 2018

Workshop notes: vocabularies and definitions in Digital Preservation

General consensus among participants that technical terms (in any field) can be a barrier to communication and fostering a clear understanding of issues. In digital preservation, some terms can be particularly problematic e.g. digital curation, to the extent that some people avoid using them.

Ensuring terms are contextualised and explaining that context are key components to clear understanding.

Digital Preservation Coalition (DPC) glossary and handbook

Observations on arrangement of terms in glossary:

- Useful to put glossary terms together in families e.g. all the acronyms for organisations; dark archive, dim archive; preservation tools/processes etc. It is easier to define terms in groups rather than each one individually.
- Absence of high-level terms can be a barrier for novices, who then think the handbook is not for them.
- Context for understanding terms important e.g. 'perpetual access' to librarians means ongoing access to content you've subscribed to. To an IT professional, it might mean preservation of bits.
- Easier to contribute to development of terms if terms are presented in a more layered/tree arrangement e.g. wiki where you can set disambiguation.

Note:

The DPC glossary sits behind the handbook, which explains the terms, so looking at the glossary in isolation does not give the whole picture. Could link terms in glossary to relevant section in handbook, so that handbook entries become scope notes for glossary and users can easily navigate to more in-depth information.

Can DPC handbook be opened up to invite contributions?

Answer – this functionality already exists, but is not generally used, probably because it is not evident to users.

CASRAI approach to creating definitions

CASRAI defines the root term, then presents context-specific uses (root term with qualifying adjective) as separate entries, rather than trying to make the root definition cover all uses.

Example:

Root term: curation

Extended term 1: Museum curation

Extended term 2: Digital curation

This approach could be a useful way to develop the glossary/dictionary.

CASRAI dictionary

- Happy to have multiple contexts defined for an entry in CASRAI dictionary.
- Maintain structured process to ensure terms are changed through agreement and due governance process.
- Ideally should have pipeline for users to submit new terms and meanings for dictionary. Terms would then be reviewed/discussed before being adopted and published. Could work jointly with DPC on this.

Suggested that, if the DPC can put take forward some work around developing/refining terms, then CASRAI can put in place process to incorporate terms into their resources.

Next steps

Several participants agreed that it would be useful to have a list of 10-20 really basic digital preservation terms with their definitions. Each term could have links to related terms for users to explore if they want to know more.

1. Get together list of key terms which would form basic list.
2. Information and calls for contributions could be shared through DPC's jiscmail listserve dc-preservation@jiscmail.ac.uk , CASRAI or similar.
3. Is there a way to use web analytics to identify which terms are searched for/clicked on most often as a way to work out which terms should be prioritised to ensure that they are clearly and accurately defined.
4. Consider where/how the terms can be made available e.g. through DPC handbook or CASRAI as appropriate.

CASRAI uses business scoping processes to drive and define development e.g. what information is needed to enable researchers in different fields to work with funders and develop funding applications. How can we add value to business case use through development on vocabularies?

Value in adding homonyms and synonyms to support interoperability.

CASRAI Reconnect UK 2018: Digital Preservation terms/vocabularies workshop

Digital Preservation Coalition glossary: terms asterisked - people unsure of meaning

- Authentication – suggestion to broaden to include users, not just digital objects.
- Authenticity
- Dark archive
- Digital archiving
- Dim archive
- Digital preservation
- DIP
- Documentation
- Escrow
- File format
- Information assurance
- Lifecycle management
- LOTAR
- Metadata
- Perpetual access
- PREMIS
- Three-legged stool
- TRAC
- Trigger event

DPC glossary: terms proposed as new terms

- ANDS – Australian National Data Service
- Archive
- Authority
- Authority list
- Authorship
- Bit rot
- Characterisation
- CLOCKSS, Controlled LOCKSS
- Graphical user interface (GUI)
- Handle (alternative to DOI)
- Integrity
- JISC
- Language
- Licence
- Metadata schema
- Normalisation
- Preservation
- Provenance
- Qbits
- Reproducibility
- Schema
- Taxonomy
- User interface (UI)
- User experience (UX)
- Zettabyte