Repositories Group: Best Practices on Artifact Integration

Simon Adar, Dirk Beyer (moderator), Patricia Cruse,
Gustavo Durand, Wayne Graves, Christopher Heid,
Lundon Holmes, Chuck Koscher, Meredith Morovatis,
Joshua Pyle, Bernard Rous, Wes Royer, Dan Valen

Agenda

- Identification, Immutability, Long-Term Availability
- Meta Data
- Referencing
- Access
- Functionality

Identifiability

Best Practice 1:
Assign a DOI to be able to uniquely identify.

At time of submission:

- Assign a DOI to the artifact

Immutability

Best Practice 2:
Ensure that the contents does not change.
Repeatability and Replicability require the exact contents that the authors used in the paper.

- ▶ DOIs (technically and by definition) allow the objects they identify to change.
- ▶ Do not use that feature.
- ▶ Hard requirement: Providers of DOIs guarantee immutability of the content they point to.

Best Practice 3:
Add SHA2 hash to the meta data of the artifact.
(Cf. Software Heritage Foundation: Lookup via SHA hash.)

Long-Term Availability

Was not discussed, but solutions were provided in the afternoon session Dec. 7.

Meta Data

Best Practice 4:
Each artifact should be accompanied by a set of self-descriptive fields.
Each kind of artifact requires different meta data,
but there are standards
(CrossRef, DataCite, Research Data Alliance).

At time of submission:

- Use self-descriptive meta data
- Standard for meta data: to be defined

Referencing

Best Practice 5:
There should be no reference without a DOI.
For papers:

- Don't use URLs! They disappear, content changes.
- If used in references, then only with DOI!

Best Practice 6:
DOI of artifact has to occur in text and meta data (!)
of paper that is based on it.

Best Practice 7:
New version of article and new version of artifact result
in possibly 2 new DOIs.

Access

For reviewing, it is required to provide access to an artifact without the DOI being published already.

Best Practice 8:
Provider needs to offer a URL with access token.

Access with username and password is sometimes asked for.

Functionality (Minimal Set)

Best Practice 9:
Repository needs to provide at least the following functions:

- ▶ Unique Identification
- ▶ Immutability
- ▶ Long-Term Availability
- ▶ Open Access
- ▶ Download
- ▶ Meta Data