Digital Object Identifiers and the ESA Science Archives

Deborah Baines

ESAC Science Data Centre (ESDC), European Space Agency

Christophe Arviset, Guido de Marchi & Bruno Merín

Summary

- ESA has DOI capability, through agreement with CrossRef: prefix 10.5270 = ESA
- Already used for publications and Earth Observation data, now also starting with Space Science data
- Goals: help scientists track and reuse data
  - Immediate: simplify curation and long-term preservation of space science data
  - Ultimate: simplify linking papers to data

Rational

- DOI is a tool that can help advancement of science, but avoid unnecessary complications
- Same solution across all archive areas
- Same solution applies to collections defined by missions and by users in the community
- ESA DOIs use format like ESA–h2o7nvx as we want them to be opaque, no second-guessing
- ESA DOIs always refer to current version of calibrated data, no new DOI if recalibrated

Granularity

- DOIs are assigned at collection level
- DOIs will be available in a portal to point to data
- Collections are defined depending on area of research (see next slide)
- Collections can contain data from more than one mission
- Different collections can contain (part of) the same data

Collection Levels - Missions

- Astronomical observatories: DOI assigned at proposal level. Currently about 20,000.
- Other astronomy missions: DOI assigned at level of catalogue (e.g. Gaia Data Release) or map (e.g. Planck). Currently about 1000s.
- Planetary science missions: DOI assigned to each dataset (usually a few per instrument/experiment). Currently about 500.
- Heliophysics missions: DOI to each dataset (usually instrument / experiment / campaign). About 3000.
- Missions to generate about 1000 new DOIs per year.

Collection Levels - Users

- For ESA, ultimate goal of DOI is to help scientists track and reuse data discussed in papers
- Most practical way to do this: let scientists assign DOI to collections of data used in their papers
- Paper defines data used, scientists will select them from archive and mint DOI on the fly (STScI already offer this)
- Also community-generated Higher Level Science Products (e.g. FP7, H2020, etc.) delivered to ESA will receive DOI in the same way

Summary

- ESA has DOI capability, through agreement with CrossRef: prefix 10.5270 = ESA
- Already used for publications and Earth Observation data, now also starting with Space Science data
- Goals: help scientists track and reuse data
  - Immediate: simplify curation and long-term preservation of space science data
  - Ultimate: simplify linking papers to data