Bibcodes in the 21st Century: an Identity Crisis

Alberto Accomazzi (ADS)
Harvard-Smithsonian Center for Astrophysics

IVOA
Data Curation & Preservation
14 May 2013

Bibcode Format

YYYYJJJJJVVVVMPPPPA

Year
Journal
Volume
Qualifier
Page
Author Initial

History and use

- Bibcodes were born by agreement between NED and CDS as de-facto unique, persistent identifiers for bibliographic records in Astronomy. Spec available at http://adsabs.harvard.edu/abs/1995ASSL..203..259S

- Bibcodes are easily computable so that different parties can create them from the basic article metadata (a "citation string") since nobody "owns" their minting rights

- Bibcodes enable linking of resources across archives without the development of major system infrastructure

- Note that some coordination is still needed to mind journal abbreviations and to deal with non-standard cases

DOIs

- Spec defined in the late 90's, contains some minimal structure and no built-in semantics:

  ```
  10.PPPP/whatever

  PPPP: provider code
  whatever: opaque string
  ```

- For example, the article mentioned earlier has the DOI: `10.1007/978-94-011-0397-8_24`

- DOIs can easily be turned into actionable URIs: http://dx.doi.org/10.1007/978-94-011-0397-8_24

DOI properties

- DOIs are persistent identifiers which are (usually) minted by the publisher when the article is created, then deposited in a global registry with associated metadata

- However, in general, DOIs cannot (and should not) be computed from a citation to a paper found in the literature, as they are void of semantics by design

- To compensate for the shortcoming above, infrastructure has been developed so that a global registry and services built on top of it can be used to:

  ▸ find a DOI from a citation string

  ▸ find the metadata and fulltext of a paper from a DOI

Do we need both?

- Bibcodes are widely used in astronomy and provide a basic layer of interoperability across archives

- Not all entries in ADS have DOIs, but all have bibcodes so we still need identifiers for those

- DOIs are widely deployed and are used across disciplines, so we may as well embrace them as standards

- The good news: ADS allows you to look up and discover metadata via bibcodes, DOIs, arXiv ids:
  http://adsabs.harvard.edu/abs/2000A%26AS..143...85A
  http://adsabs.harvard.edu/doi/10.1051/aas:2000172
  http://adsabs.harvard.edu/arXiv/astro-ph/0002105

So what's the problem?

- The model upon which bibcode is built is breaking down in the electronic publication age (no more "pages")

- Field length for journal/volume/{page|eid} not sufficient

  ▸ PhRv 6-digit ids, Hindawi 7-digit article ids, GCN 5-digit volumes

- Basic metadata not available at publication time

  ▸ "online early" article only have a DOI, no volume/page

- Metadata model mismatch

  ▸ Vizier catalogs, observing proposals, ASCL software entries, PhD thesis

But I like bibcodes!

- So do I. Once a bibcode, always a bibcode, so no broken links/resources

- For Astro and other "well behaved" journals, we'll be able to compute bibcodes for a long time

- ADS will continue to provide robust services for correlating identifiers, merging metrics, etc.

- Proposals to "stretch" the bibcode would buy us a little time, but at a significant cost, and without long-term guarantees. Probably not worth the effort

So what should I do?

- Don't panic. Nothing is changing in the short term

- We (ADS, NED, SIMBAD) started talking about these issues and will continue to do so

- If you are developing software or maintaining a system, think of literature linking as a process which will require a discovery/verification step

- ADS will continue to provide services and API to support discovery and linking of its resources

- We think we know what the community needs, but feel free to provide us with use cases and feedback at any time

Useful Links

- Notes about AAS2013 bibcode discussion: https://docs.google.com/document/d/128gdglvL6N1vFvClpKzMMRv-gPyufiEq51yR2ptcEPY/edit#

- ADS help page on bibcodes: http://doc.adsabs.harvard.edu/abs_doc/help_pages/data.html#Bibliographic_Identifiers

- The ADS bibliographic reference resolver: http://ads.harvard.edu/pubs/resolver/

- Crossref DOI lookup: http://www.crossref.org/guestquery/

- For our semantic friends: DOIs as linked data post by Ed Summers: http://inkdroid.org/journal/2011/04/25/dois-as-linked-data/