



*International  
Virtual  
Observatory  
Alliance*

This is  
only an  
outline!  
Every-  
thing is  
still to  
be done.

# Best practices for the creation of and metadata for digital object identifiers in astronomy archives

## Version 1.0

### IVOA Note 2024-01-01

Interest Group

Data Curation and Preservation

This version

<https://www.ivoa.net/documents/DOI4Archives/20240101>

Latest version

<https://www.ivoa.net/documents/DOI4Archives>

Previous versions

This is a draft

Author(s)

August Muench, Gilles Landais, Raffaele D'ABrusco, Anne Raugh

Editor(s)

TBD

## Abstract

Many astronomy archives are producing digital object identifiers (DOI) for datasets and services. This document aims to summarize current workflows for creating and using DOIs, diagnose issues in the metadata of extant DOIs, and develop best practices for workflows and metadata for future DOI deployment. This note is focused on archives in Astronomy, Planetary Science, and Heliophysics. Additional domains may be considered at a later time.

## Status of this document

This is an IVOA Note expressing suggestions from and opinions of the authors. It is intended to share best practices, possible approaches, or other perspectives on interoperability with the Virtual Observatory. It should not be referenced or otherwise interpreted as a standard specification.

A list of current IVOA Recommendations and other technical documents can be found at <https://www.ivoa.net/documents/>.

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Role within the VO Architecture . . . . .	4
<b>2</b>	<b>Extant use of Digital Object Identifiers in Archives</b>	<b>4</b>
2.1	Datasets . . . . .	5
2.2	Collections . . . . .	5
2.3	Services . . . . .	6
2.4	Knowledgebases . . . . .	6
2.5	IVOA Registry . . . . .	7
<b>3</b>	<b>Use cases and current pathologies for Dataset DOIs</b>	<b>7</b>
3.1	Use Cases . . . . .	7
3.2	Pathologies . . . . .	8
<b>4</b>	<b>Best Practices for DOI Workflows</b>	<b>8</b>
4.1	Workflows: Datasets . . . . .	8
4.2	Workflows: Collections . . . . .	9
4.3	Workflows: Knowledgebases . . . . .	9
4.4	Workflows: Other . . . . .	9
<b>5</b>	<b>Best Practices for DOI Metadata</b>	<b>9</b>
5.1	Core Metadata . . . . .	9
5.2	Relationship Metadata . . . . .	11
<b>A</b>	<b>Catalog of Repositories and Archives issuing DOIs</b>	<b>11</b>
A.1	Domain Specific repositories . . . . .	11
A.2	Generalist repositories . . . . .	12
<b>B</b>	<b>Important DataCite Metadata Case Studies</b>	<b>12</b>
B.1	Chandra Data Archive . . . . .	12
B.2	VAMDC & Zenodo workflow . . . . .	12
B.3	Vizier . . . . .	12

<b>C Changes from Previous Versions</b>	<b>12</b>
<b>References</b>	<b>13</b>

## Acknowledgments

???? Or remove the section header ????

## Conformance-related definitions

The words “MUST”, “SHALL”, “SHOULD”, “MAY”, “RECOMMENDED”, and “OPTIONAL” (in upper or lower case) used in this document are to be interpreted as described in IETF standard RFC2119 (Bradner, 1997).

The *Virtual Observatory (VO)* is a general term for a collection of federated resources that can be used to conduct astronomical research, education, and outreach. The *International Virtual Observatory Alliance (IVOA)* is a global collaboration of separately funded projects to develop standards and infrastructure that enable VO applications.

## 1 Introduction

1. General observations. We focus on roles of digital object identifiers (DOIs) rather than the full landscape of FAIR data (e.g., Wilkinson and Dumontier et al., 2016). It is beyond the scope of this document to address all aspects of FAIRness. It is also unreasonable we think for DOIs to be laden with the problems of enabling all aspects of FAIRness. Instead we aim for this document to guide archives with achieve compliance for specific aspects of enabling FAIR data. These specific aspects are F.X, A.Y....
2. Using DataCite Commons<sup>1</sup> and the Crossref API<sup>2</sup>, we find approximate 75000 dataset DOIs in astronomy and related fields (See Table 1).
3. This list is not exhaustive; see Table XXX in Appendix A for a complete listing.
4. Describe use cases and problems for DOIs in astronomy: citation, provenance, ...
5. DOIs do not lead to data and other future problems

<sup>1</sup><https://commons.datacite.org>

<sup>2</sup><https://www.crossref.org/documentation/retrieve-metadata/rest-api/>

Known major missing pieces: Explicit DataCite meta-data keywords; Guidance on DataCite deposit, including JSON markup or Fabrica.

Figure 1: Architecture diagram for this document

6. Outline of IVOA note.
7. A complete listing of all astronomy/heliophysics/planetary science archives issuing dataset (or other forms of) DOIs is given in Appendix A
8. A series of important example case studies in DataCite metadata deposits is given in Appendix B.
9. This document adopts DataCite Metadata Schema 4.5 ([DataCite Metadata Working Group, 2024](#))

Archive	Prefix	ID	Yo	Count
Canadian Astronomy Data Centre	10.11570	nrc.cadc	-	80
Chandra Data Archive	10.25574	si.cda	-	29,449
ESO Science Archive Facility			-	
European Space Agency	10.5270		-	28,858
IPAC (26131 — 26135)	10.2613X	caltech.ipacdoi	-	622
Mikulski Archive for Space Telescopes	10.17909	stsci.mast	-	2,132
NASA Planetary Data System	10.17189	nasapds.nasapds	-	1,976
++Small Bodies Node of PDS	10.26007	sbn.archive	-	4,837
Strasbourg Astronomical Data Center	10.26093	inist.cds	-	17,383

Table 1

### 1.1 Role within the VO Architecture

Fig. 1 shows the role this document plays within the IVOA architecture ([Dowler and Evans et al., 2021](#)). This section could talk about the Registry, Data Origin, BibVO, etc. Someone needs to write it.

## 2 Extant use of Digital Object Identifiers in Archives

We observe four different uses for DOIs in astronomy-related archives. Identifiers are being assigned for individual datasets (at various levels of granularity); for collections of datasets; for services; and for what we will call

---

#### Note

Dataset DOIs do not lead to data, or what they do lead to is widely varied.

---

"knowledgebases" or curated metacollections of data and calculated results. These uses are described in more detail below.

## 2.1 Datasets

Dataset DOIs are the most common usage in astronomy-related repositories, and represent about »XX% of all such DOIs. Individual datasets are being assigned DOIs at various levels of granularity. At the most granular level, every observation in the Chandra Data Archive (CDA) is assigned an individual DOI. Similarly, the European Space Agency assigned DOIs to every individual Herschel, ISO, Planck and XMM-Newton observation. This list also includes every Hubble Space Telescope (HST) observation, which are also archived at the Mikulski Archive for Space Telescopes (MAST) but are not assigned DOIs individually by MAST.

Individual dataset DOIs are implemented mostly consistent with the usage of DOIs by institutional repositories, generalist repositories (e.g., Zenodo), etc. That is, the dataset DOI resolves to a landing page describing the dataset. The landing page provides one or more links to data. The content of the landing page and structure of those data links is not otherwise prescribed. In some cases the landing pages are rich summaries of the dataset, often including much more information than could be gleaned from the DOI metadata alone. In other cases the landing pages pop the user into a web user interface with very little (or almost no) contextual information about the DOI.

Dataset DOIs are also being minted for "High-Level Science Products" (HLSP). Such HLSP datasets are not static and can grow over time, accumulating both data revisions and additions. HLSP landing pages for DOIs have no set structure or content and change regularly as the linked data evolve.

None of these dataset DOIs are versioned. The data discovered by following a dataset DOI to an individual observation will change based upon the version of the archive pipeline used to create it, up real-time or on-demand reprocessing for some archives.

## 2.2 Collections

Collection-type DOIs direct users to a collection of individual data records in a particular archive. Examples of collection-type DOIs include the DOI services provided by the Mikulski Archive for Space Telescopes (MAST)<sup>3</sup> (Novacescu and Peek et al., 2018), Chandra Data Archive (CDA)<sup>4</sup> (Rots and D'Abrusco et al., 2018), and the VAMDC Consortium<sup>5</sup> Query Store

---

<sup>3</sup><https://archive.stsci.edu>

<sup>4</sup><https://cxc.harvard.edu/cda/>

<sup>5</sup><https://vamdc.org>

service (Moreau and Zwolf et al., 2018). There is significant variation in the expected use cases for these Collection DOIs. There are also strong variations in the metadata of Collection-type DOIs.

Considering use-cases, both MAST and Chandra Collection DOIs collect dataset identifiers within their respective databases. In both cases there is no expectation that these Collection DOIs would ever themselves collect attribution (i.e., be cited in the reference list of a corresponding Journal article). However, VAMDC Collection DOIs, are intended to collect and distribute credit to the collection of database identifiers they contain. They attempt to distribute this citation/credit to the collection of resources by "citing" all the related resources in the saved VAMDC query record. Because this depends upon the capabilities of both the chosen DOI minting service (Zenodo) and the DataCite Schema, additional notes on the details and the outcomes of this effort is provided in one of our case studies in Appendix B.

Similarly the metadata of Collection DOIs vary between MAST and CDA examples. As described in D'Abrusco and DPAOps Team (2023), Chandra Data Collection DOIs create complete records of individual using `relatedIdentifier` tags and predicates. The metadata of MAST Collection DOIs do not provide detailed information on the individual MAST records collated in the collection.

## 2.3 Services

Examples of DOIs for services include IRSA (DUST). Service DOIs lead to query tools. They may lead to query results – I would need examples.

Sometimes Collection DOIs act like Service DOIs but they are not. Collection DOIs may result from queries performed at a Service.

**Warning: weeds.**

## 2.4 Knowledgebases

A knowledgebase is a collection of material collated from many discrete sources. All of the values contained in a knowledgebase have a provenance traced to other resources and have been curated into a single database for reuse. Examples include: Simbad (as originally described in, Wenger and Ochsenbein et al., 2000), NASA Exoplanet Archive (NEA) (NASA Exoplanet Science Institute, 2020) (as originally described in, Akesson and Chen et al., 2013), NASA Extragalactic Database (NED) (NASA/IPAC Extragalactic Database (NED), 2019) (as originally described in, Helou and Madore et al., 1991).

Current observations about DOIs for knowledgebases include:

- DOIs for knowledgebases lead users to interstitial landing pages rather than directly to the collated, curated resources.

- DOIs for knowledgebases do not lead to individual values, e.g., the results of a query against that knowledgebase.
- DOIs for knowledgebases never provide information in their *metadata* about the *state* of a knowledgebase: its current version; last update; etc. Nor is this information on the interstitial landing pages of knowledgebases.
- Services that return DOIs for queries against knowledgebase are considered Service DOIs (See Section 2.3).

**Warning: weeds.**

## 2.5 IVOA Registry

DOIs are being used in some manner by the Registry. This includes both the indexing of DOIs related to Registry records and the minting of DOIs for Registry entries. *Until I understand this better, I consider this category distinct from any of the others. One reason is that I fear that there is a potential mis-management of DOI services here. The publisher of a DOI is responsible for the upkeep of the digital resource. If IVOA DOI minting is providing DOIs for resources it does not manage, then this is potential mis-management of DOI services. It is not unlike the ill-begotten aims of similar 3rd party DOI minting by [Re3data.Org](#) and [FAIRsharing.org](#) that have, for example, created DOIs for Simbad ([Re3data.Org, 2012](#); [FAIRsharing Team, 2018](#)) that really ought never be used by anyone aiming to cite Simbad for the source of their results.*

## 3 Use cases and current pathologies for Dataset DOIs

### 3.1 Use Cases

Present use cases comprise:

1. Providing a means for citation & attribution
2. Tracking reuse and discovering reuse modes
3. Enabling open data (and reuse) (whatever FAIR means)

Additional use cases that can be discussed as needing dataset DOIs include:

1. Provenance: following DOI relationships to procure the provenance of a dataset.

2. IVOA specific desires (if different from above): reuse Service DOIs to enable queries?
3. Reproducibility (but not really until dataset DOIs actually go to data there's a gaping chasm between the resolved DOI and whatever you want to get out from it).

### 3.2 Pathologies

The primary pathology evident in DOI creation today is a mismatch between the metadata created by an archive and the use case (intended to be) implemented by that DOI. Succinctly, the metadata supplied by astronomy archives is often insufficient to ensure the accurate citation of the datasets.

By detailing these pathologies and triaging their less-than-desirable outcomes we can aim to develop empirically-defined best practices to guide repositories forward with the use of identifiers. Here is a topical list of pathologies:

1. Incomplete metadata: missing authors, generic or misstated titles, misunderstood dates;
2. Inconsistent metadata: transmutations of metadata between systems lead to inconsistent metadata deposits. Example: transformation from schema.org to Crossref left ESA DOIs in a nasty state;
3. (Un)versioned data: versioning is mostly nonexistent and when provided it is ill defined and often opaquely transmitted;
4. Misconceptions: DOIs do not lead to data, or what they do lead to is widely varied.

## 4 Best Practices for DOI Workflows

We consider the choices made by existing astronomy archives when enabling different types of DOI services. Using that information, we make recommendations for best practices in the workflows used to set up such DOIs. This section could also include advice on how to submit metadata to DataCite, e.g.,

### 4.1 Workflows: Datasets

Include related identifiers for other sources of the data (e.g., references for catalogs merged into yours), for other digital objects that describe or contextualize the data (e.g., Journal articles that describe the data), and for software related to the creation of this data.



## 4.2 Workflows: Collections

Include related identifiers linking the collection DOI to internal and if possible external identifiers for the dataset(s). This will make the DOIs broadly more useful and interoperable. Other systems will be able to use the alternate identifiers to correctly link their records to yours.

## 4.3 Workflows: Knowledgebases

Maybe don't use DOIs or figure out how to use them effectively to transmit *state* or *version* information? This section may be

## 4.4 Workflows: Other

People have been making Service DOIs, but why?  
Broad discussion of Query DOIs?

# 5 Best Practices for DOI Metadata

## 5.1 Core Metadata

**Authorship** Archives must decide on what mechanism will be used to transmit authorship of data records. The choices include (a) full itemization of all contributions to a dataset and (b) formal use of "collaboration/institute" names, e.g., The CANDELS collaboration. And stick to it.

If an itemized author list is used then the authorship list must include all contributors to the creation of a resource, individually and properly fielded. If all contributors cannot be provided then none must be.

It maybe that some use cases do not need complete, individual author lists. Collection DOIs, which may never be indexed or cited again, might use truncated author lists. However, truncated author lists must be a rare exception and never used when attribution is expected for the reuse of that material.

When choosing a collaboration/institute name please consider the following important ideas:

- Make a decision about the "The";
- Keep the same institution name for all related datasets. If you use "NASA Exoplanet Archive" or "Kepler Mission" for some records, don't switch to the "NASA Exoplanet Science Institute" or "Kepler Project" for others.
- 

Some kind of table/graph relating meta to use case would be nice

**Title** The title of a deposit must be unique and should describe the contents of the resource. Duplicating the title of a related article or dataset is easy but is not unique and almost certainly does not describe the contents of the deposit. Common choices are to prefix the title of a related Journal article with phrases like, "Data for...".

However this doesn't really describe the contents of the deposit, does it?

**Version** heaven help us. A version is not mandatory, but without it most use cases fall apart.

Discuss the pros and cons of versioned DOIs, explicit versioning models (SemVer ([Preston-Werner, 2023](#)), CalVer ([Hashemi, 2016](#))), implicit versioning using dates on the landing page.

Discuss just having a date anywhere on the landing page would be nice. You could even have two dates! One for date the publication of the original resource. And one date for the most recent data update Oh, and maybe one more for the most recent *metadata* update!

**Dates** This section talks more specifically about the dates submitted to DataCite for a dataset.

The date describes when the resource was made available on line, not when the DOI was minted. The date of the most recent update to the DOI contents and/or metadata is a different date. You can provide different kinds of dates to DataCite.

**Description** A description of the the resources is mandatory.

Consider if ADS indexes your resource. If you do not provide a description then the resource will be indexed w/o any text/abstract! Even a minimal summary is an improvement over no description/abstract.

There are a couple "fields" that are free text but that might be better if they were structured. Description might be one of those fields.

And yet maybe it is better to talk about embedding metadata in landing pages, e.g., schema.org instead of trying to dictate what people put in the description field in DataCite.

**Licensing and reuse permissions** Mandatory

Pretty important that the license is provided and that the license matches any other license that has been applied to these results.

In other words, if you copy data from a resources then you can't give it a new license. NASA wants you to use CC0.

**Transmutations of metadata** Consider how your user is going to cite these records. Will they find them in ADS or in Google Scholar or will they try to cite the DOI directly?

If you are supplying metadata in multiple formats (on landing pages, in DataCite XML, in schema.org tags) then ensure that all forms are consistently transformed. There are crosswalks between these metadata schema (e.g. between DataCite and Schema.org Dataset). Using such crosswalks can ensure that the metadata is expressed consistently to all such services.

## 5.2 Relationship Metadata

1. Related Identifiers
2. Predicates (which to use, which to ignore)
3. Add References as Related Identifiers.

## A Catalog of Repositories and Archives issuing DOIs

This is not an update of the [Chen and Schmitz et al. \(2022\)](#) appendix. This is meant to also provide additional information about what kinds of DOIs are minted and what types of deposits the archive will accept.

Needs  
a lot of  
work  
to doc-  
ument  
these

### A.1 Domain Specific repositories

1. CADC
2. CDA
3. China-VO
4. ESA
5. ESO
6. IPAC
7. MAST
8. NEA
9. PDS

## **A.2 Generalist repositories**

1. Dataverse
2. Dryad
3. Zenodo
4. Other?

## **B Important DataCite Metadata Case Studies**

This section describes case studies around the metadata created by three archives. These case studies let us break out specific DataCite metadata recommendations and JSON snippets

### **B.1 Chandra Data Archive**

CDA is doing cool things with related identifiers.

### **B.2 VAMDC & Zenodo workflow**

The good and the bad of transitory data citation.

### **B.3 Vizier**

Vizier produces some awesome DataCite metadata.

## **C Changes from Previous Versions**

No previous versions yet.

---

## Note

All dataset citations in this document were created using BibTeX. The BibTeX for the data/service/collection DOIs is based upon the `verbatim` metadata supplied by the archive to their choice of DOI registration services as expressed by the doi2bib service<sup>a</sup>. The doi2bib service was selected because it does the best job protecting unfielded "author" names from complete BibTeX annihilation. Every other reference BibTeX came from NASA ADS<sup>b</sup>.

---

<sup>a</sup><https://doi2bib.org>

<sup>b</sup><https://ui.adsabs.harvard.edu/>

---

## References

- Akeson, R. L., Chen, X., Ciardi, D., Crane, M., Good, J., Harbut, M., Jackson, E., Kane, S. R., Laity, A. C., Leifer, S., Lynn, M., McElroy, D. L., Papin, M., Plavchan, P., Ramírez, S. V., Rey, R., von Braun, K., Wittman, M., Abajian, M., Ali, B., Beichman, C., Beekley, A., Berriman, G. B., Berukoff, S., Bryden, G., Chan, B., Groom, S., Lau, C., Payne, A. N., Regelson, M., Saucedo, M., Schmitz, M., Stauffer, J., Wyatt, P. and Zhang, A. (2013), ‘The NASA Exoplanet Archive: Data and Tools for Exoplanet Research’, *PASP* **125**(930), 989, arXiv:1307.2944. doi:10.1086/672273, <https://ui.adsabs.harvard.edu/abs/2013PASP..125..989A>.
- Bradner, S. (1997), ‘Key words for use in RFCs to indicate requirement levels’, RFC 2119. <http://www.ietf.org/rfc/rfc2119.txt>.
- Chen, T. X., Schmitz, M., Mazzarella, J. M., Wu, X., van Eyken, J. C., Accomazzi, A., Akeson, R. L., Allen, M., Beaton, R., Berriman, G. B., Boyle, A. W., Brouty, M., Chan, B. H. P., Christiansen, J. L., Ciardi, D. R., Cook, D., D’Abrusco, R., Ebert, R., Frayer, C., Fulton, B. J., Gelino, C., Helou, G., Henderson, C. B., Howell, J., Kim, J., Landais, G., Lo, T., Loup, C., Madore, B., Monari, G., Muench, A., Oberto, A., Ocvirk, P., Peek, J. E. G., Perret, E., Pevunova, O., Ramirez, S. V., Rebull, L., Shemmer, O., Smale, A., Tam, R., Terek, S., Van Orsow, D., Vannier, P. and Wang, S.-Y. (2022), ‘Best Practices for Data Publication in the Astronomical Literature’, *ApJS* **260**(1), 5, arXiv:2106.01477. doi:10.3847/1538-4365/ac6268, <https://ui.adsabs.harvard.edu/abs/2022ApJS..260....5C>.
- D’Abrusco, R. and DPAOps Team (2023), ‘A New Way to Cite Chandra Data’, *Chandra News* **34**, 5. <https://ui.adsabs.harvard.edu/abs/2023ChNew..34....5D>.

- DataCite Metadata Working Group (2024), ‘DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs v4.5’. Publisher: [object Object] Version Number: 4.5. doi:10.14454/G8E5-6293, <https://datacite-metadata-schema.readthedocs.io/en/4.5/>.
- Dowler, P., Evans, J., Arviset, C., Gaudet, S. and Technical Coordination Group (2021), ‘IVOA Architecture Version 2.0’, IVOA Endorsed Note 01 November 2021. doi:10.5479/ADS/bib/2021ivoa.spec.1101D, <https://ui.adsabs.harvard.edu/abs/2021ivoa.spec.1101D>.
- FAIRsharing Team (2018), ‘Fairsharing record for: Simbad astronomical database’. doi:10.25504/FAIRSHARING.RD6GXR, <https://fairsharing.org/FAIRsharing.rd6gxr>.
- Hashemi, M. (2016), ‘Calendar Versioning CalVer’. <https://calver.org/>.
- Helou, G., Madore, B. F., Schmitz, M., Bica, M. D., Wu, X. and Bennett, J. (1991), The NASA/IPAC extragalactic database., *in* M. A. Albrecht and D. Egret, eds, ‘Databases and On-line Data in Astronomy’, Vol. 171 of *Astrophysics and Space Science Library*, Springer, Dordrecht, pp. 89–106. doi:10.1007/978-94-011-3250-3\_10, <https://ui.adsabs.harvard.edu/abs/1991ASSL..171...89H>.
- Moreau, N., Zwolf, C.-M., Ba, Y.-A., Richard, C., Boudon, V. and Dubernet, M.-L. (2018), ‘The VAMDC Portal as a Major Enabler of Atomic and Molecular Data Citation’, *Galaxies* **6**(4), 105. doi:10.3390/galaxies6040105, <https://ui.adsabs.harvard.edu/abs/2018Galax...6.105M>.
- NASA Exoplanet Science Institute (2020), ‘Planetary systems table’. doi:10.26133/NEA12, <https://catcopy.ipac.caltech.edu/doi/doi.php?id=10.26133/NEA12>.
- NASA/IPAC Extragalactic Database (NED) (2019), ‘Nasa/ipac extragalactic database (ned)’. doi:10.26132/NED1, <https://catcopy.ipac.caltech.edu/doi/doi.php?id=10.26132/NED1>.
- Novacescu, J., Peek, J. E. G., Weissman, S., Fleming, S. W., Levay, K. and Fraser, E. (2018), ‘A Model for Data Citation in Astronomical Research Using Digital Object Identifiers (DOIs)’, *ApJS* **236**(1), 20, arXiv:1801.00004. doi:10.3847/1538-4365/aab76a, <https://ui.adsabs.harvard.edu/abs/2018ApJS..236...20N>.
- Preston-Werner, T. (2023), ‘Semantic Versioning 2.0.0’. <https://semver.org/>.
- Re3data.Org (2012), ‘Simbad astronomical database’. doi:10.17616/R39W29, <https://www.re3data.org/repository/r3d100010163>.

- Rots, A., D’Abrusco, R. and Winkelman, S. (2018), A Model for Using DataCite DOIs in Observatory Bibliographies, *in* ‘European Physical Journal Web of Conferences’, Vol. 186 of *European Physical Journal Web of Conferences*, p. 12011. doi:[10.1051/epjconf/201818612011](https://doi.org/10.1051/epjconf/201818612011), <https://ui.adsabs.harvard.edu/abs/2018EPJWC.18612011R>.
- Wenger, M., Ochsenbein, F., Egret, D., Dubois, P., Bonnarel, F., Borde, S., Genova, F., Jasiewicz, G., Laloë, S., Lesteven, S. and Monier, R. (2000), ‘The SIMBAD astronomical database. The CDS reference database for astronomical objects’, *A&AS* **143**, 9–22, arXiv:astro-ph/0002110. doi:[10.1051/aas:2000332](https://doi.org/10.1051/aas:2000332), <https://ui.adsabs.harvard.edu/abs/2000A&AS.143....9W>.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J., Groth, P., Goble, C., Grethe, J. S., Heringa, J., ’t Hoen, P. A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. and Mons, B. (2016), ‘The fair guiding principles for scientific data management and stewardship’, *Scientific Data* **3**(1), 160018. doi:[10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18).