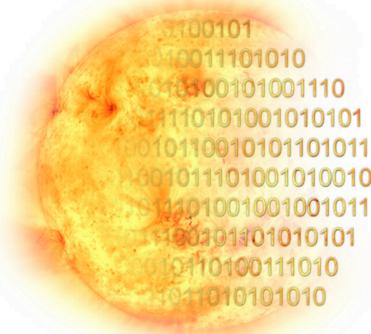
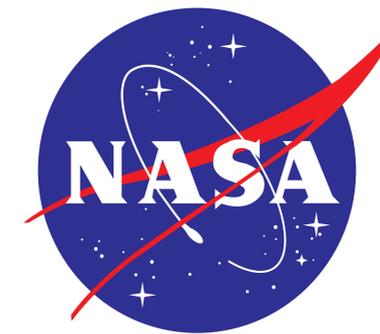


# Recommendations for Data & Software Citation in Solar Physics



J.A. Hourclé  
NASA-GSFC (Wyle)  
joseph.a.hourcle@nasa.gov

Virtual Solar Observatory  
<http://virtualsolar.org/>

We present a series of recommendations to improve the citation of solar physics data to ensure validation and reproducibility.

We include recommendations for data providers to make their data more easily cited by solar physicists and the wider scientific community, as well as recommendations for authors who are using the data.

We hope to improve the acknowledgement rate of not only solar physics data but also the tools used by our community, so that we can ensure continued maintenance and availability of the infrastructure used by the science community.

We also hope to establish guidelines for describing the history of the data that may be necessary for verification of the research, including the original source, and the methods and tools used in processing the data for analysis.

## For Data Providers:

### 1. Data Distributors should define logical groupings for the data they provide.

We treat each group as a separate 'publication', so that scientists can concretely talk about a specific set of data, as there may be multiple processed forms (raw, calibrated, averages, etc) in multiple packaged forms (FITS, ASCII tables, IDL save files).

They need not be mutually exclusive (a data product may be in more than one group); the groupings should reflect the different forms in which the data is distributed and how other scientists will likely subset the data for use in research.

For example:  
SDO/AIA level 1

SDO/AIA level 1, full resolution  
SDO/AIA level1, full res. 94 Ångstrom  
SDO/AIA level1, full res. 131 Ångstrom

SDO/AIA level 1.5

SDO/AIA level 1.5, 4x4 binned  
SDO/AIA level1, 4x4 binned, 94 Ångstrom  
SDO/AIA level1, 4x4 binned, 131 Ångstrom

### 2. Each grouping should be given a title, creator, and other information necessary for citation.

Creator : NASA/SDO and the AIA Science Team  
Title : SDO/AIA 171 Angstrom Level 1 Intensity Images  
Publisher : SDO JSOC  
Publication Year : 2010

See DataCite for generally recommended attributes; specific guidelines for solar physics will be posted to Solar News in the near future.

### 3. Create a web page with this information, and links to relevant documentation.

See "Linking Articles to Data" handout for an explanation of the advantages of this over proxies for data citation.

### 4. Assign a DOI to each information page.

Digital Object Identifiers (DOIs) are persistent identifiers commonly used in the publishing industry for identifying books, journal articles and other items. DOI resolvers can obtain a URL for the item, but the URL can be changed should the content be moved.

Becoming a DOI registrant require ensuring that the documents are going to be preserved for the long-term under stricter guidelines than most scientific data archives. It may be necessary to work with your organization's library or institutional repository.

### 5. Include that DOI in the FITS header.

```

NAME$C = 1024 / length of data axis 2
COMMENT ---Documentation & Contact Information---
COMMENT This is a level-1 SWAP FITS file produced by p2sw_prep v1.1 at the Royal
COMMENT Observatory of Belgium. If you have difficulty with this file or wish
COMMENT to make suggestions for improvements, please contact the SWAP
COMMENT Instrument Team via email at swap_lyna@oma.be.
COMMENT For information on data rights, keyword definitions, citing this data
COMMENT and up-to-date reports on known problems and data quality, see:
COMMENT http://dx.doi.org/10.5067/example/PROB22.SWAP_Level1
COMMENT ---Observation Identification---
FILENAME= 'swap_lv1_20110806_000614.fits' / FITS filename
  
```

## For Software Developers

### 1. Create a similar page with information on how to cite the software for each major version of the software.

### 2. Assign a DOI to those information pages

(see notes re: DOIs under Data Providers)

### 3. Provide an 'about screen' in the software that links to the description page.

### 4. When appropriate, generate a 'receipt' that lists the data processed and the processing applied.

For example, from the VSO Web Interface, a VSO Cart ID stores:

What query parameters were used  
When the query was run  
What data products were selected

Ideally, this should provide the information that authors need to properly cite the data, and provide sufficient provenance information to allow the researcher to re-apply the processing if the input changes (eg, calibration updates) or others to re-run the processing for validation (peer review).

## For Authors

### 1. Cite the data used as you would any other article, with a link to the data info page.

NASA/SDO and the AIA Science Team, (2010). *SDO/AIA 171 Angstrom Level 1 Intensity Images*, SDO JSOC. <http://dx.doi.org/10.1234/example.vso.sdo.aia.lev1.171>

### 2. Cite any software used as you would any other article, with a link to the info page.

The VSO Team, (2004). *The Virtual Solar Observatory*, Solar Data Analysis Center. <http://virtualsolar.org/>

Alternate form, including an identifier for a 'receipt', like the VSO Cart ID:

The VSO Team, (2004). *The Virtual Solar Observatory*, Solar Data Analysis Center <http://virtualsolar.org/>. VSO-SDAC-120604-001638

### 3. Include the list of data and software used & processing applied in the article's 'extended methods' supplement.

The exact details will need to be fleshed out with use and as formal standards are developed. For the time being, a text summary explaining the range of the data used, any filtering applied, and the general processing applied is an easy first step.

A VSO Cart ID can be used to link to the list of data used.

## References

- Ball, A. & Duke, M., (2011). "Cite Datasets and Link to Publications". <http://www.dcc.ac.uk/resources/how-guides>
- DataCite, (2011). "DataCite Metadata Schema for the Publication and Citation of Research Data". <http://dx.doi.org/10.5438/0005>
- Hourclé, Chang, Linares, Palanisamy and Wilson, (2012). "Linking ArticlestoData". <http://docs.virtualsolar.org/wiki/Citation>
- Hourclé, (2012). "Advancing the Practice of Data Citation", ASIS&T Bulletin, accepted



Copy of this poster, handout and additional information is available at :  
<http://docs.virtualsolar.org/wiki/Citation>