Recommendations for Data & Software Citation in Solar Physics

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 level 1, full res. 94 Ångstrom
   - SDO/AIA level 1.5
     - SDO/AIA level 1.5, 4x4 binned
   - SDO/AIA level 1, 4x4 binned, 94 Ångstrom
   - SDO/AIA level 1, 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 Ångstrom Level 1 Intensity Images
   - Publisher: SDO/SSC
   - 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. DOIs can be assigned to data items, making it easier for users to find the data they need.
   - We recommend using DataCite for DOI registration and management.

5. Include that DOI in the FITS header.
   - This ensures that the data can be easily referenced and cited by others.

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

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

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

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

References
- DataCite, (2011). “DataCite Metadata Schema for the Publication and Citation of Research Data”. http://dx.doi.org/10.5067/R885
- 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