Topic: Citation Breakout/Preservation Ontology
Date: Thursday, July 14, 2011
Time: 1:30 pm, Mountain Daylight Time (Denver, GMT-06:00)

Mark Alan Parsons: Presenting information on Citation Form and Content

Citations are the information that allows a user to locate and characterize the data. It isn't a part of discovery...

Roles to be cited:
        -Authors
        -Editors
        -Archivist or Distributor(publisher-term to be avoided)

Distinction made between and Identifier and a Locator
        -Identifiers are things that are concrete and don't change
        -Locators are things that places data somewhere, e.g. structure, hierarchy level...

There should only be one locator for a data set.

Most issues with locators can be solved by a good versioning scheme, tracking and documentation, and due diligence.
        -Versions should be well documented, past versions of data(even if not available) should be documented. This is in case it is cited and a user needs to understand what the data set was.
        -How to best create locators/citations for parts of a dataset? Are there structural indices that are common to earth sciences?
        -The Archive Information Unit or structural index should be devised from the structure/elements in the data.

A citation must include: Author(s), release date, title, version, archive and/or distributor, Persistent location service. Other optional parts: Editor, Archive/Distributor Place, other institutional role, identifier of subsets used.

<Discussion>
Terminology about locator vs location service vs identifiers resulted in the conclusions that locator and location aren't the same. A web link is not a locator or identifier, because it can change.

Citations need to have human readability, they can't simply all be like a doi.

How often is the data cited, how are subsets determined? What about operational vs archived vs archived and operational data? If data is constantly updated is it's citation constantly

changing.

Space and Time are likely enough for most remotely sensed data to identify them in a citation. it is up to the data producer to determine the structural index for a citation, e.g. the chapter and verse.

Citations can only answer so much, the remainder is left to the user to determine through other sources, e.g. description of dataset in text of the paper.

What is a 'version' change? Is it a temporal change, is it a break in back compatibility or change in science contained in the dataset?

 What is necessary to determine what dataset a user has conducted their work on? Date assessed, citation, and description of dataset in text by user, help answer that.

Curt offered a proposal to have a journal or editor to send proposed papers are sent to a list of volunteers to analyze the data citations prior to publication.
        -Plan B: Send citation requirements to editors to incorporate into their guidelines.

It was offered that members bring Data Citation up during their AGU committees.

Mark referred to this document:
https://docs.google.com/document/d/1mHYlu3IU3q6RaJfkxyNZQcykYMB-iak4N2Gd3fNgd2Q/edit?hl=en_US
<End Discussion>