Directory Interchange Format (DIF) Usability Survey

NASA's Earth Science Data Systems Standards Process Group (SPG) is considering the DIF (Directory Interchange Format) specification, developed by the GCMD (Global Change Master Directory), for adoption as a community standard. Your responses to this survey on the usability of the DIF and the suitability of this specification for Earth science data will be helpful.

Please answer as many of the questions below as you can..

1. Please provide your name, organization and contact information (including email address).

2. Are you answering for your entire organization, for a smaller group, or individually?

__X__ a) Entire organization

_____ b) Smaller group (please specify) _____

_____ c) Individual response

3. Are you a data producer, data consumer, or both?

_____ a) Producer _____ b) Consumer __X_ c) Both

4. How long have you been using the DIF?

Since the mid 1990's

5. Please describe how the DIF is used in your organization.
(1) DIFs are submitted to GCMD for user discovery of datasets.
(2) DIFs are used, with minor modifications via XSLT, to publish data collection metadata to the EOS
(3) DIFs are used to supply collection level metadata for our S4PA archives
(4) DIFs are used to supply collection level metadata for our Mirador search tool
(5) DIFs are in the process of being used as a source for our semantic-web enhancement to Mirador.

For the following, you can answer either about the DIF alone, or relative to other, comparable specifications.

6. What are the strengths of the DIF? How has the use of the DIF helped your organization?

DIFs are reasonably generalized frameworks for storing metadata about data collections. The structure is very stable, yet additional fields can be added by GCMD if sufficient utility is to be had. The XML format makes them eminently machine parseable

7. What are the weaknesses of the DIF? What would you like to change about the DIF or what would make the DIF a better specification?

I don't think the DIF has any significant weaknesses, other than not having been blessed by an official standards organization. (Perhaps this process can fix that.)
Suggested additions:

(1) Add a description field to the Detailed Variable
(2) Add a short description field (e.g. < 256 char) for datasets
(3) Add a "Shortname" field for datasets (e.g. < 12 chars) or else an "alias" field; may datasets are known more commonly by this shortname than by the full descriptive title.

8. How well does the DIF solve your metadata storage, discovery, and/or interchange needs? Are there specific areas it is applicable to vs. areas where it is not applicable or not used?

As seen by the numerous internal applications in which we use DIF, it suits our metadata needs very well. The one area where its use is minimal is in Giovanni, which is predominantly parameter-oriented.

9. How suitable is the DIF for representing your data holdings?

DIF does an excellent job of representing our data holdings.

10. Do you use the DIF to track your own data holdings (i.e. do you use DIF in your own data management activities)?

Yes, extensively. See answer to section 5.

11. What are the limitations of the DIF? Does the DIF prevent you from doing things you would like to do? Does its use make other things more difficult?

More structure in the parameter area would allow us to expand it to serve Giovanni.

12. Do you think ESDS-RFC-012 (and thus the DIF) should be endorsed as a NASA Earth Science Data Systems Standard? Why or why not?

Yes, absolutely: it is very useful for data collection metadata storage and exchange; it is used in the community both as a standard, and as a source for ontology information; it is mappable to similar national (FGDC) and international (ISO 19115) standards; it is supported by a handy web interface for entering and modifying DIFs, as well as machine-to-machine interfaces.