

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?

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 c) Both

4. How long have you been using the DIF?

We have been building and using DIF records since at least 1995.

5. Please describe how the DIF is used in your organization.

For each product in our archive we create, deliver and maintain a DIF record at the GCMD. The maintenance was somewhat haphazard in the past. We have recently designed a system to automatically generate DIF records from a spreadsheet database. We plan to maintain our DIF fields via this method and repopulate modified DIF records to the GCMD on an annual basis.

We have also adapted DIF on a granule basis for a certain suite of products

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?

The DIF record is relatively straightforward. Overall it is simple to create a DIF record although maintenance and changes can take a lot of time if there are many products in GCMD. This is not the fault of DIF though.

The DIF is very good at representing gridded satellite products and flexible for other large earth science data holdings.

It has helped our organization for the perspective of data discovery via the GCMD.

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?

DIF is tuned towards describing product collections. It is less suited to describe metadata at the granule (file) level although it can be adapted. It also has no provisions for systematically describing quality and product provenance accurately. Just loosely in a summary section.

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?

Yes, as mentioned above, the DIF is not that suited for the granule metadata and this inhibits discovery of granules. Overall, for discovery of a large collection of data for the same product it works well.

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

In the past, it was very suitable at the collection level (i.e., the product level itself). However, as remote sensing archives increase in size and resolution, and users want to get closer to the native processing level (i.e., level 1 and 2) it becomes more of an issue with regards to representing metadata at the granule level.

In our organization, it will always have a place for representing the collection metadata.

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

We have adapted DIF to work with granule metadata for a certain suite of products. So the answer is yes.

Separately, an internal product database and spreadsheet is compatible and cross-mapped to DIF.

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?

Some of them mentioned above. Actually none of the metadata formats such as FGDC or ISO 19115 are more elegant with respect to describing granule metadata. But ISO has a well defined quality and lineage section. In the future as data holdings becomes even larger and more complex (i.e., more derived, interdisciplinary products) this will be increasingly important.

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?

I agree from the heritage perspective that it should be endorsed. But I look forward to seeing NASA metadata distributed in more modern formats such as FGDC and ISO 19115. These formats seem to have a large community behind them.

Ideally there should be tools to crosswalk between all these formats or the GCMD should build capability to receive and deliver any of these metadata formats.