Geophysical Research Abstracts
Vol. 12, EGU2010-12092, 2010
EGU General Assembly 2010

Data Discovery, Exploration, Integration and Delivery - a practical experience

Peter Kirsch, Tim Barnes, and Paul Breen
British Antarctic Survey, Physical Sciences Division, Cambridge, United Kingdom (pjki@bas.ac.uk)

To fully address the questions and issues arising within Earth Systems Science; the discovery, exploration, integration, delivery and sharing of data, metadata and services across potentially many disciplines and areas of expertise is fundamental.

British Antarctic Survey (BAS) collects, manages and curates data across many fields of the geophysical and biological sciences (including upper atmospheric physics, atmospheric chemistry, meteorology, glaciology, oceanography, Polar ecology and biology). BAS, through its Polar Data Centre has an interest to construct and deliver a user-friendly, informative, and administratively low overhead interface onto these data holdings.

Designing effective interfaces and frameworks onto the heterogeneous datasets described above is non-trivial. We will discuss some of our approaches and implementations; particularly those addressing the following issues:

How to aid and guide the user to accurate discovery of data?

Many portals do not inform users clearly enough about the datasets they actually hold. As a result the search interface by which a user is meant to discover information is often inadequate and assumes prior knowledge (for example, that the dataset you are looking for actually exists; that a particular event, campaign, research cruise took place; and that you have a specialist knowledge of the terminology in a particular field), assumptions that cannot be made in multi-disciplinary topic areas.

How easily is provenance, quality, and metadata information displayed and accessed?

Once informed through the portal that data is available it is often extremely difficult to assess its provenance and quality information and broader documentation (including field reports, notebooks and software repositories). We shall demonstrate some simple methodologies.

Can the user access summary data or visualizations of the dataset?

It may be that the user is interested in some event, feature or threshold within the dataset; mechanisms need to be provided to allow a user to browse the data (or at least a summary of the data in the most appropriate form be it a plot, table, video etc) prior to making the decision to download or request data. A framework should be flexible enough to allow several methods of visualization.

Can datasets be compared and or integrated?

By allowing the inclusion of open, 3rd party, standards compliant utilities (e.g. Open Geo-Spatial Consortium WxS clients) into the framework, the utility of a data access system can be made more valuable.

Is accessing the actual data straightforward?

The process of accessing the data should follow naturally from the data discovery and browsing stages. The user should be made aware of terms and conditions of access. Access restrictions (if applicable) and security should be made as unobtrusive as possible.

How is user feedback and comment monitored and acted upon?

In general these systems exist to serve science communities, appropriate notice and acknowledgement of their needs and requirements must be taken into account when designing and developing these systems if they are to be of continued use in the future.