DLESE and NSDL

The role of the
Digital Library for Earth System Education* (DLESE)
in the National SMETE Digital Library

Presented by Dave Fulker
Director of the Unidata Program Center
University Corporation for Atmospheric Research
Boulder, Colorado

*DLESE is funded jointly by the GEO & EHR Directorates of the NSF

Two Emphases for this Talk

1. Sketching one of DLESE's contributions to the nascent NSDL initiative so far:
   - ***Building community foundations***
   - ***(Skipping other efforts by DLESE & partners)***
     - Cataloging methodologies & tools
     - Sharable middleware to support interoperability...
2. Sketching a role DLESE is positioned to play: enabling data access & discovery in NSDL
   - ***Relevant data should be easy to find***
   - ***Data found should be easy to use & understand***

Community Foundations for NSDL

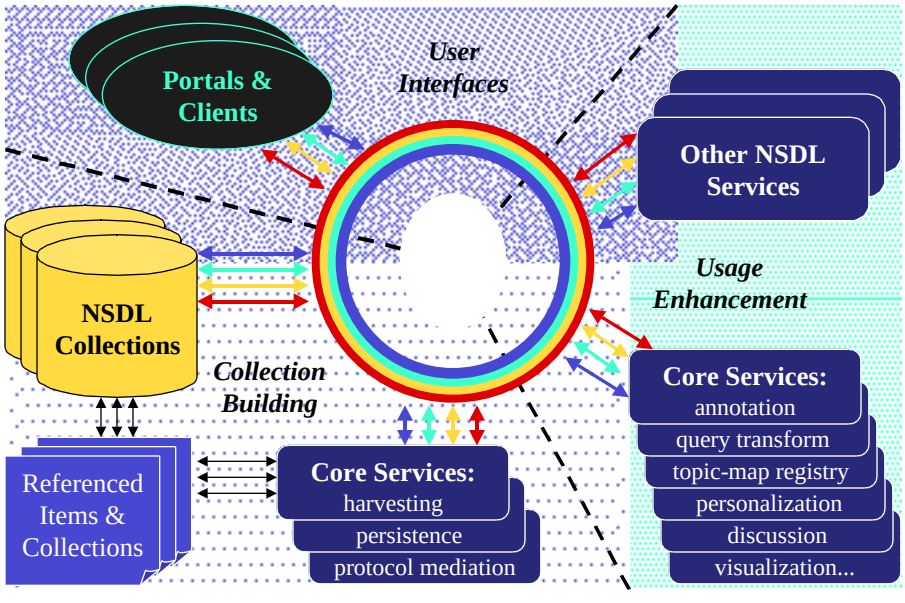
- All hands meeting
  - ***Attendees from all 42 of the funded projects***
  - ***Posters, brainstorming...***
- Election of an NSDL Coordinating Committee
- Formation of workgroups, active in planning NSDL
  - ***Individual workgroup reports***
  - ***Guidance on Core Integration functionality for NSDL***
  - ***Overarching white paper, see:***
    http://www.smete.org/nsdl/workgroups/coordcomm/

7 Workgroups

- Governance
- Community, Education, & Pedagogy
- Collections
- Standards
- Services
- Evaluation
- Technical Infrastructure

NSDL Information Architecture

Developed by the Technical Infrastructure Workgroup



Enabling Data Access & Discovery in NSDL
(reflecting the convergence of needs from two worlds)

- In the digital libraries world, science educators expect to find & to use (unfamiliar) data/images meaningfully
  - ***Tools are envisioned that will help***
    - Integrate/visualize complex data/images from diverse sources
    - Handle geo-referencing, resampling, units of measure...
  - ***But data/tool classification is an unsolved problem***
- In the scientific data management world
  - ***One can publish data/images almost as easily as text, & one can access remote data as easily as local data***
  - ***But it is hard to find data & match them to the tool at hand...***

Å Common need: discovery systems & rich metadata

- ***Embracing discipline-specific data semantics***
- ***Supporting the use of powerful software tools (GIS...)***

A Dilemma: Who Performs the Work?

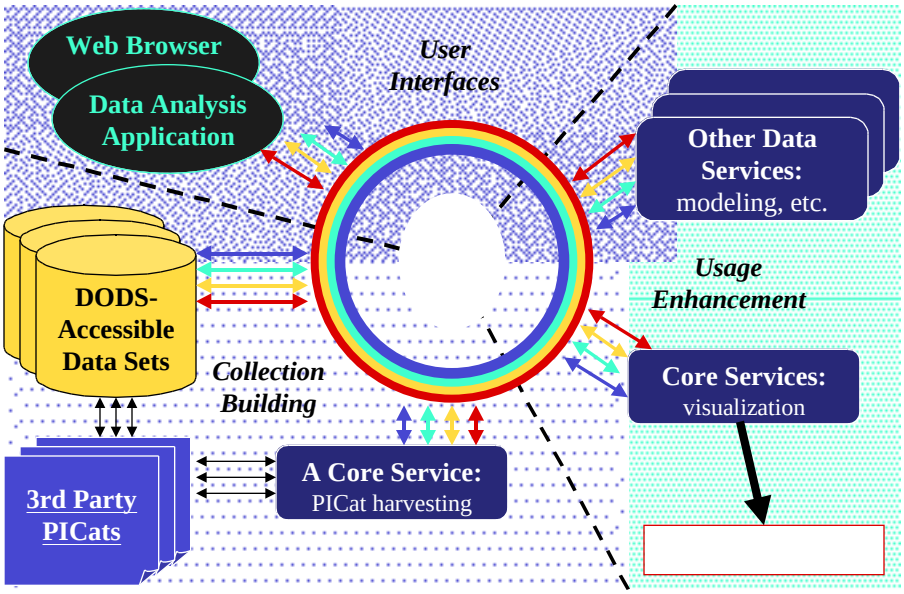
- Data providers can't meet all needs
  - ***Providers perceive little benefit from the work of adding metadata***
  - ***Meaning is always problem-specific, so complete semantic metadata is impossible***
  - ***Granularity requirements also are problem-specific***
- Data users typically lack expertise/time to
  - ***Form suitable units of information***
    - Aggregations of multiple files (time series, e.g.)
    - Subsets (sub-samplings of specified variables, e.g.)
  - ***Acquire metadata needed for tools & understanding***

Idea:

3rd Party Metadata Publication

- Publishable Inventories & Catalogs (PICats) could be built anywhere using XML
  - ***To create "virtual aggregations" of data sets***
  - ***To characterize data sets in meaningful ways (including metadata needed by applications)***
- Data providers would see reduced demands for filling user-specific requests
- Data users would gain new views of data
  - ***Relating them to specific (educational) contexts***
  - ***Facilitating their use in familiar software tools***

PICat Consistency with NSDL Architecture



Example of a (Prototype) PICat

Note: XML => ease of inclusion in digital libraries...