# Standard Mechanisms for Data Exchange

by Dave Fulker

as the second in a two-part presentation:

## The Role of Standards for Data & Metadata

BESSIG Launch Meeting, 20-April-2011

# Clarification

* Perhaps this talk should be titled pseudo-standard mechanisms for data exchange
* It's less about formally approved standards than about practices that become "standard" by virtue of being commonplace

# Insert Joke Here

* Three scientific data managers were in a bar, engaged in witty, erudite conversation...

# Technology & Data - Looking Back

* Writing Instruments
* Scrolls/Books/Libraries
* Maps/Charts/Graphs
* Printing Press
* Teletypes & Facsimile Machines
* Computers & Digital Storage
* Internet & World Wide Web
* Numerical Symbols...
* Tables (relations...)
* Representations (qualitative...)
* Dissemination paradigm shift
* Dissemination paradigm shift
* Representations (quantitative...)
* Dissemination paradigm shift

# Technology & Data - Looking Around

* Computers & Digital Storage
  * Data encoding as programming exercises... evolving to
  * Data management/analysis/visualization as library functions or applications (with APIs)
* Internet & World Wide Web
  * File exchanges (like mailing tapes)... evolving to
  * Real-time flows & remote data access, with significant server-side functionality

# Technology & Data - Looking Ahead

* Oops; not now...
* First I'll drill down a bit more re
* Data management/analysis/visualization as library functions or applications (with APIs)
* Real-time flows & remote data access, with significant server-side functionality
    * Mandatory mention (to establish credibility) of my personal experience with punched cards & 7-track tapes

# Data-Oriented Library Functions & Applications

* Evolution of data formatting
  * Metadata as instructions for programmers
  * Library functions for data formatting
    * GBYTES & SBYTES - serialization aids (for programmers)
    * netCDF - hiding serialization; self-describing data files
    * netCDF Java - portable data-structure objects
    * CDF/netCDF - portable geo-science data objects
* Evolution of data-base management systems
  * Please find a more qualified speaker

# (cont.) Data-Oriented Library Functions & Applications

* Data Analysis & Visualization
  * Too long a story to tell here (GIS has its own story)...
  * Common theme: power of a system is directly related to the generality/richness of its underlying data model
* Come to think of it, I have a DBMS-related remark
  * The underlying data model—relations (tables) that can be joined selectively—is unwieldy for representing data structures with dimensionality ≥ 2
  * This is a major factor in the success of netCDF...

# Data Capabilities Enabled by Internet & the Web

* Data flows are increasingly common
  * RSS, Unidata's IDD, etc. move computing closer to real time
    * These strain conventional notions about units of data
    * Units of flow may resemble messages more than files, e.g.
* Remote access is increasingly "standard"
  * E.g., the OPeNDAP protocol (DAP) is designated a "community standard" by NASA and is employed in several servers
    * Unidata's TDS, OPeNDAP's Hyrax, PYDAP...

# Technology & Data - Looking Ahead

* This page is deliberately left blank.