Metadata Share Project
Project Scope

Background and Goals
Background

• Our institutions support data collections that are similar/complimentary

• Current data documentation is inefficient and results in much unnecessary duplication of effort

• Our idea - divide up metadata creation and share records across institutions
Desired Outcomes and "Deliverables"

- Process/tools in place for creation of metadata instances
- Collection of (valid!) DDI instances for our various data resources
- Sharing of instances across institutions for each of us to borrow and adapt for particular needs
Intended Goals:

- Enhanced Searching/Discovery of Data Resources
- Sharing of DDI Instances
- Facilitation of Participation in Other Digital Library Initiatives
Goal #1 - Enhanced Search and Discovery

- Descriptions of data resources in library catalogs are often thin
- DDI allows for more "rich" descriptions with greater amount of useful detail
- Provide users with options for more powerful and controlled searches
Goal #2 - Sharing Records

• Take advantage of common resources across collections

• Strengthen catalogs of holdings and resources

• Reduce amount of time that would be devoted to metadata creation
Goal #3 - Participate in Digital Initiatives

• Trend across libraries to consolidate searching of holdings and databases into one interface
Implementation - Creating and Sharing Instances

• Identification of DDI Tags

• Standardization of Values for Elements and Attributes

• Metadata Entry

• Comparability and Sharing Across Institutions
Identification of DDI Tags

- Examples and Templates - e.g. ICPSR abstracts
- Standard: DDI 2.1
- `<docDscr/>` - Instance Itself
- `<stdyDscr/>` - Geography, Time, Contents
- `<fileDscr/>` - File Formats
- Utilizing Repeatable Elements and `<notes/>`
Standardization of Values for Elements and Attributes:

- Variation across resources in documentation, terminology
- Absence of controlled vocabularies
- E.G. `<nation/>` and Country Names
Metadata Entry

- Emory - webform
  (http://cti.library.emory.edu/EDC/datafreewayform/)
Metadata Entry

- Duke - NoteTab template
Examples

IRIS Dataset - current Duke interface
http://library.duke.edu/cgi-bin/data/dao.cgi?keywords=iris

IRIS Dataset - working draft of DDI version (next slide)
Iris Dataset
IDNO: 350198419

Principal Investigator(s): Knack, Steve IRIS Center at the University of Maryland
Keefer, Philip IRIS Center at the University of Maryland

Abstract: The IRIS Dataset was originally constructed in 1993 by Steve Knack and Philip Keefer for the IRIS Center at the University of Maryland, based on data obtained from the International Country Risk Guide (ICRG). The dataset includes computed scores for six ICRG political risk variables: corruption in government, rule of law, bureaucratic quality, ethnic tensions, repudiation of contracts by government, and risk of expropriation. Knack produced subsequent issues of the data for an ongoing series of working papers from the IRIS Center.

Related Terms: corruption, risk, political risk, expropriation, icrg

Coverage

Dates of Coverage: 1982 to 1997
Geographic Coverage: global
Smallest Geographic Unit: country

Methods

Analytic Unit: county-year

Access: http://library.duke.edu/data/duke-only/iris/iris.zip

Restrictions:
Data are restricted to current Duke faculty, staff, and students only. Redistribution is prohibited.
DAO Version Date: February 21, 2008
This DDI instance was developed as part of a joint initiative between Emory's Electronic Data Center and Duke's Data and GIS Services.
Comparability and Sharing

• Exchanging Records
• Checking for Duplication
• Evaluating Access Restrictions
Challenges

• Inter-Coder Reliability

• Institutional Priorities
Challenge #1 - Inter-Coder Reliability

- Lack of controlled vocabularies
- Differences in emphasis on and intent of use of resources
- Entry by students not familiar with metadata or data
Challenge #2 - Institutional Priorities

• Need for programming support

• 'Data's' Role in the 'Digital' Library
Conclusions
Conclusions

• Metadata entry is the greatest hurdle to adopting DDI.
Conclusions

• Metadata entry is the greatest hurdle to adopting DDI.

• Identifying a controlled vocabulary for DDI instances presents localization challenges when sharing records.
Conclusions

• Metadata entry is the greatest hurdle to adopting DDI.

• Identifying a controlled vocabulary for DDI instances presents localization challenges when sharing records.

• Sharing DDI records enhances the holdings of each institution and leaves more time for working with researchers.
Metadata Share Project

Joel Herndon
Duke University
Perkins Library
joel.herndon@duke.edu
http://library.duke.edu/data

Rob O'Reilly
Emory University
Woodruff Library
roreill@emory.edu
http://einstein.library.emory.edu
I have no idea what you're talking about...

...so here's a bunny with a pancake on its head.