



U.S. GOVERNMENT PRINTING OFFICE | KEEPING AMERICA INFORMED

# Getting from Theory to Practice: A metadata model for GPO's digital future

George Barnum

Office of Innovation & New Technology

U.S. Government Printing Office

# Metadata?

“Nobody will know what you’re talking about.”

-Mike DiMario

# The view from 38,000 feet

- Metadata helps **locate, interpret, and manage** data

# It's not really complicated (theoretically)

- how data is formatted or arranged;
- how,
- when, and
- by whom it was compiled;
- where it resides;
- how is it represented. named, characterized

# Early assumptions about metadata in FDsys

- Every process will either create or require metadata (or both)
- Metadata will be captured from processes, acquired from outside sources, or created
- Metadata will be built up in layers
- Metadata will probably be centrally stored
- Multiple schema will be required, and must interact

# Categorized by broad function

- Descriptive
- Technical/Structural
- Administrative
- Preservation

# GPO's Flexible Digital Content System

- Fundamentally a system for managing content for the long term
- Based on the OAIS model
- Aspects of
  - composition
  - preservation
  - content management
  - access

# What has FDsys planning told us?

- Our high level view is valid
- Metadata is increasingly part of the fabric of what we're managing, rather than a discrete entity
- We will store metadata in a common form, and will employ different schema to meet needs throughout the organization



# From the abstract to the concrete...

- There is no cookbook
- There are some recipes, we're concocting a menu
  - METS
    - For expressing the structure of packages; extensible by reference
  - MODS
    - For expressing descriptive (e.g. bibliographic) info; an extension
  - MARC21
    - The gold standard, but not XML; a vast, invaluable asset
  - PREMIS
    - Preservation information; another extension

# Side dishes on the menu...

- Dublin Core
- ONIX
- Other yet to be determined