14th Annual Federal CASIC Workshops
Washington, DC, March 16 - 18, 2010

WP-1 Survey Uses of Metadata

Efficiently delivering (micro)data on the web using DDI-XML

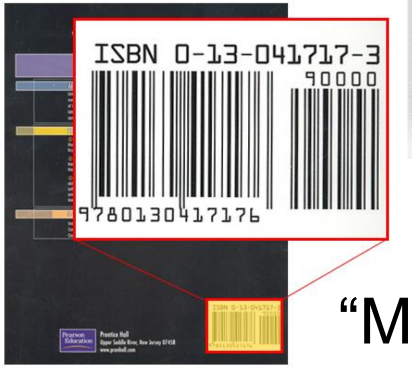
Pascal Heus, Open Data Foundation
http://www.opendatafoundation.org
info@odaf.org

Metadata you use everyday…

“Human Readable” Metadata

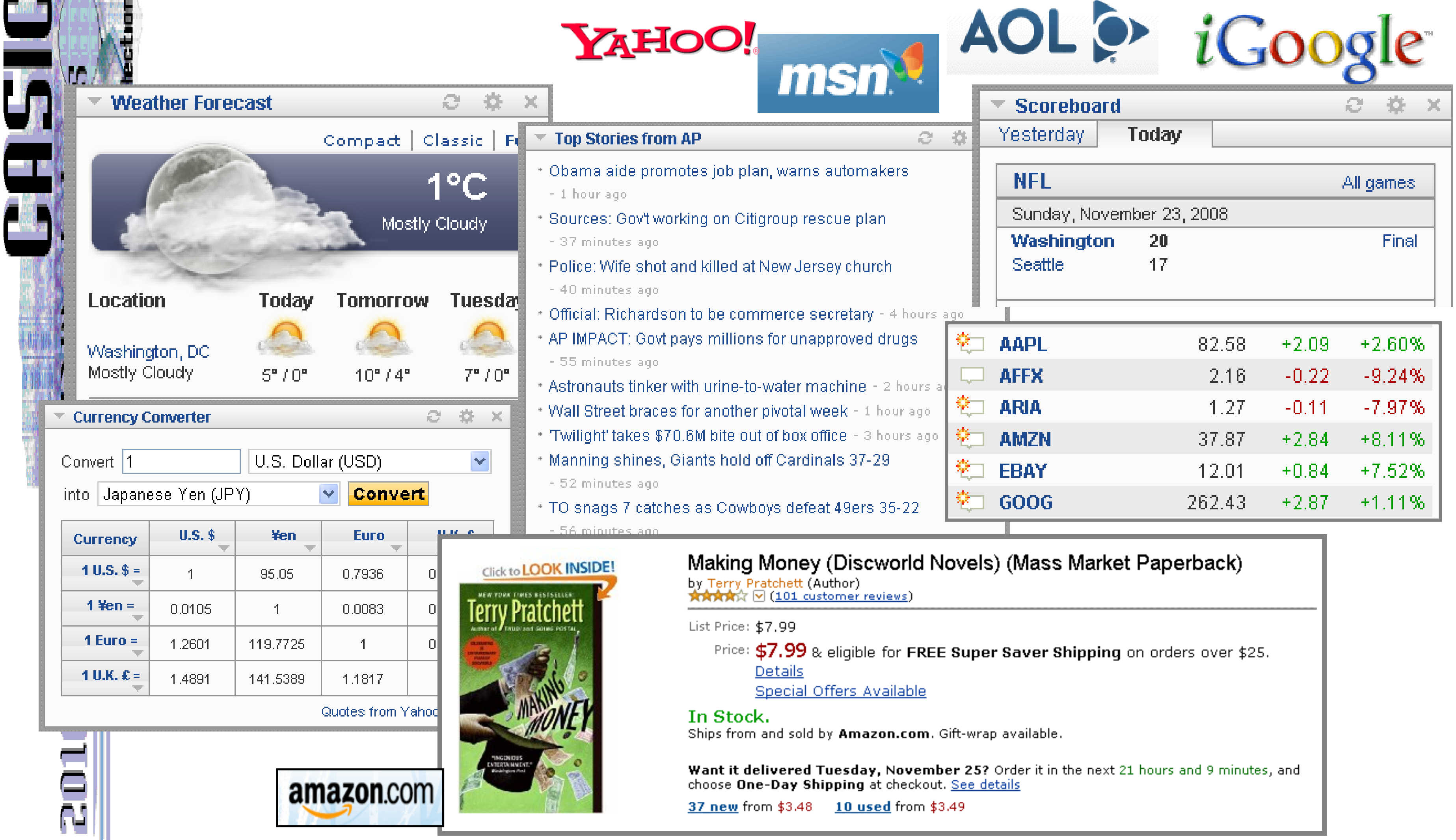
Nutrition Facts
Valeur nutritive
Per 1 bowl (300 g) / Pour 1 bol (300 g)

Amount Teneur	% Daily Value % valeur quotidienne
Calories / Calories 440	
Fat / Lipides 19 g	**29 %**
Saturated / Saturés 4 g + Trans / Trans 0.2 g	**21 %**
Cholesterol / Cholestérol 35 mg	
Sodium / Sodium 860 mg	**36 %**
Carbohydrate / Glucides 53 g	**18 %**
Fibre / Fibres 4 g	**16 %**
Sugars / Sucres 6 g	
Protein / Protéines 15 g	
Vitamin A / Vitamine A	45 %
Vitamin C / Vitamine C	4 %
Calcium / Calcium	20 %
Iron / Fer	20 %



“Machine-actionable” Metadata

Metadata you use everyday…



Weather Forecast

Compact | Classic | Fu

1°C
Mostly Cloudy

Location	Today	Tomorrow	Tuesday
Washington, DC Mostly Cloudy	5° / 0°	10° / 4°	7° / 0°

Top Stories from AP

- Obama aide promotes job plan, warns automakers - 1 hour ago
- Sources: Gov't working on Citigroup rescue plan - 37 minutes ago
- Police: Wife shot and killed at New Jersey church - 40 minutes ago
- Official: Richardson to be commerce secretary - 4 hours ago
- AP IMPACT: Govt pays millions for unapproved drugs - 55 minutes ago
- Astronauts tinker with urine-to-water machine - 2 hours a
- Wall Street braces for another pivotal week - 1 hour ago
- 'Twilight' takes $70.6M bite out of box office - 3 hours ago
- Manning shines, Giants hold off Cardinals 37-29 - 52 minutes ago
- TO snags 7 catches as Cowboys defeat 49ers 35-22 - 56 minutes ago

Scoreboard

Yesterday | Today

NFL — All games

Sunday, November 23, 2008

Washington	**20**	Final
Seattle	17	

AAPL	82.58	+2.09	+2.60%
AFFX	2.16	-0.22	-9.24%
ARIA	1.27	-0.11	-7.97%
AMZN	37.87	+2.84	+8.11%
EBAY	12.01	+0.84	+7.52%
GOOG	262.43	+2.87	+1.11%

Currency Converter

Convert 1 U.S. Dollar (USD)

into Japanese Yen (JPY) Convert

Currency	U.S. $	Yen	Euro	U.K. £
1 U.S. $ =	1	95.05	0.7936	0
1 Yen =	0.0105	1	0.0083	0
1 Euro =	1.2601	119.7725	1	0
1 U.K. £ =	1.4891	141.5389	1.1817	

Quotes from Yahoo

Click to LOOK INSIDE!

Making Money (Discworld Novels) (Mass Market Paperback)
by Terry Pratchett (Author)
(101 customer reviews)

List Price: $7.99

Price: $7.99 & eligible for **FREE Super Saver Shipping** on orders over $25.
Details
Special Offers Available

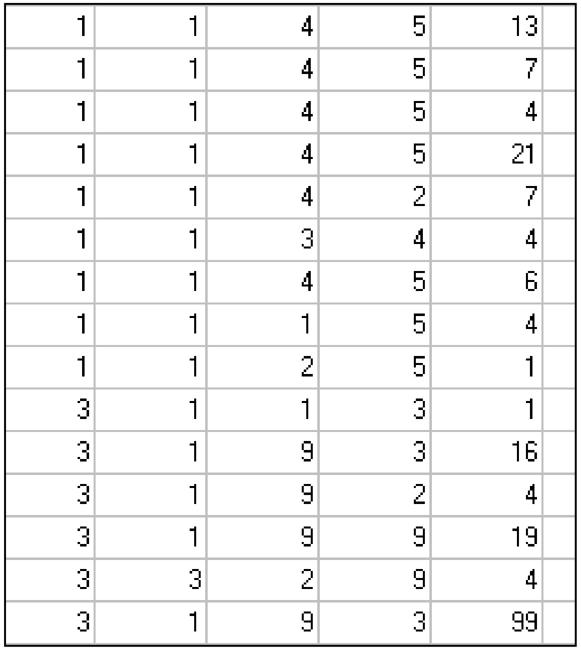
In Stock.
Ships from and sold by **Amazon.com**. Gift-wrap available.

Want it delivered Tuesday, November 25? Order it in the next 21 hours and 9 minutes, and choose **One-Day Shipping** at checkout. See details

37 new from $3.48 10 used from $3.49

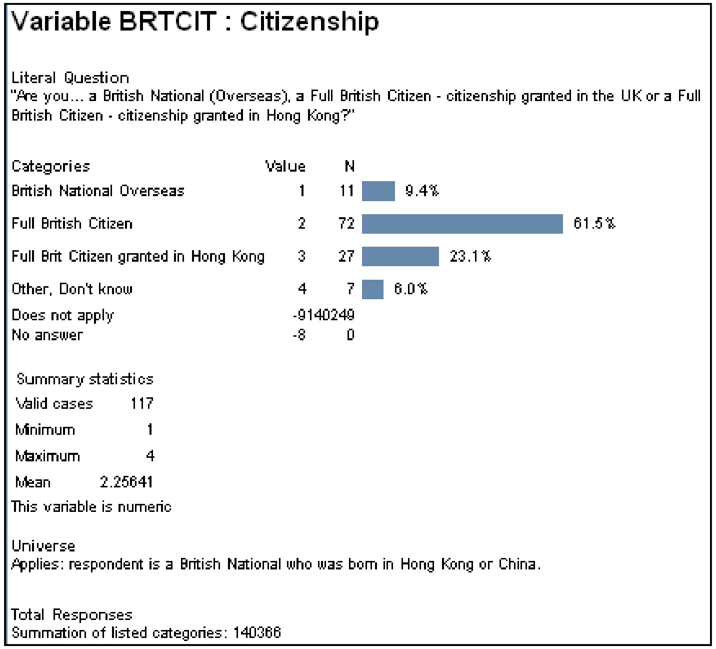
What are metadata?

Common definition: Data about Data



1	1	4	5	13
1	1	4	5	7
1	1	4	5	4
1	1	4	5	21
1	1	4	2	7
1	1	3	4	4
1	1	4	5	6
1	1	1	5	4
1	1	2	5	1
3	1	1	3	1
3	1	9	3	16
3	1	9	2	4
3	1	9	9	19
3	3	2	9	4
3	1	9	3	99

Unlabeled stuff



Variable BRTCIT : Citizenship

Literal Question
"Are you... a British National (Overseas), a Full British Citizen - citizenship granted in the UK or a Full British Citizen - citizenship granted in Hong Kong?"

Categories	Value	N	
British National Overseas	1	11	9.4%
Full British Citizen	2	72	61.5%
Full Brit Citizen granted in Hong Kong	3	27	23.1%
Other, Don't know	4	7	6.0%
Does not apply	-9	140249	
No answer	-8	0	

Summary statistics	
Valid cases	117
Minimum	1
Maximum	4
Mean	2.25641

This variable is numeric

Universe
Applies: respondent is a British National who was born in Hong Kong or China.

Total Responses
Summation of listed categories: 140366

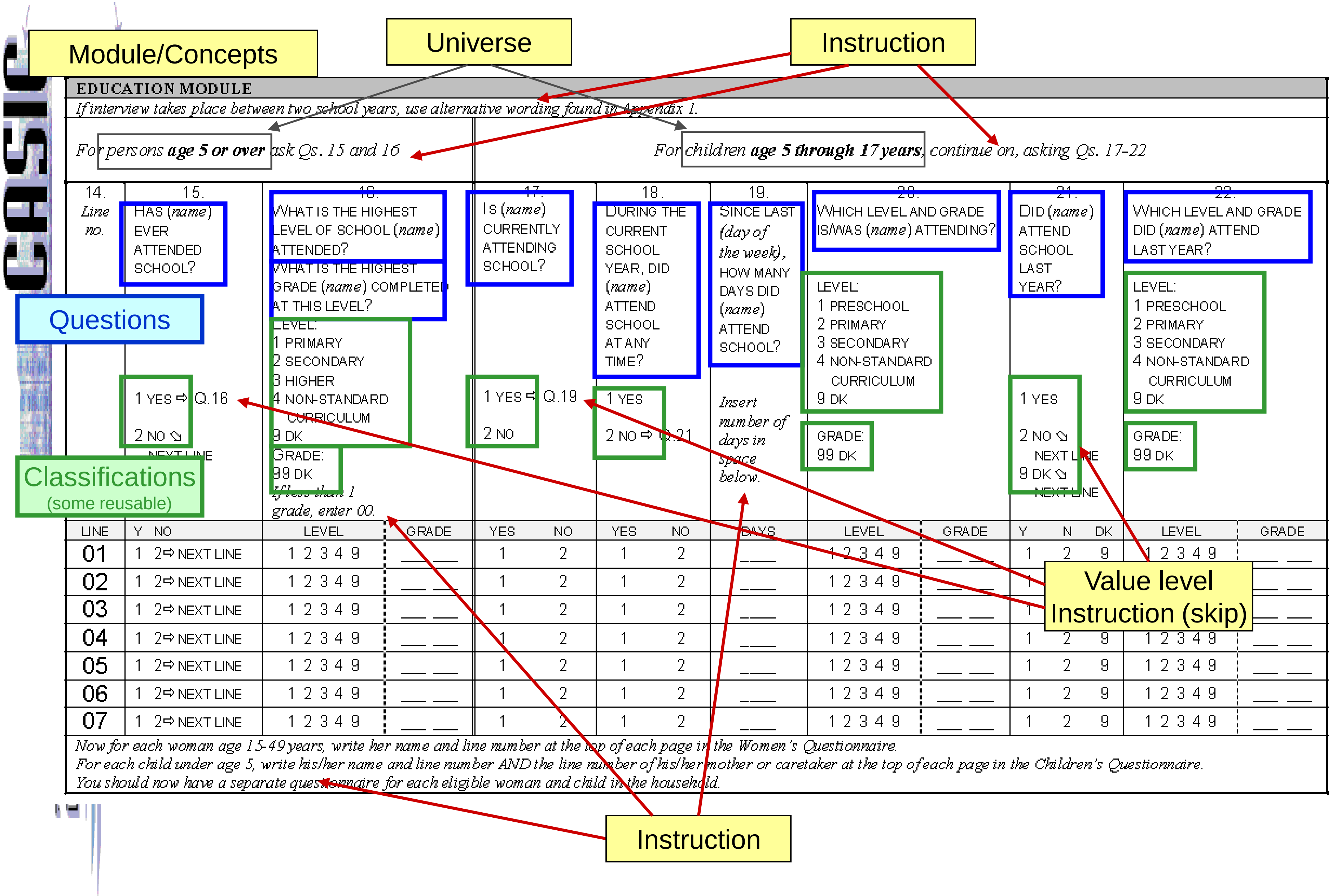
Labeled stuff

The bean example is taken from: A Manager's Introduction to Adobe eXtensible Metadata Platform, http://www.adobe.com/products/xmp/pdfs/whitepaper.pdf

Metadata for microdata

- Need more that data dictionary and a couple of documents….
- Survey level
  - Data dictionary (variable labels, names, formats,…)
  - Questionnaires: questions, instructions, flow, universe
  - Dataset structure: files, structure/relationships,
  - Survey and processes: concepts, description, sampling, stakeholders, access conditions, time and spatial coverage, data collection & processing,…
  - Documentation: reports, manuals, guides, methodologies, administration, multimedia, maps, …
- Across surveys
  - Groups: series, longitudinal, panel,…
  - Comparability: by design, after the fact
  - Harmonization
  - Common metadata: concepts, classifications, universes, geography, universe

Questionnaire Example



Module/Concepts

Universe

Instruction

Questions

Classifications
(some reusable)

Value level Instruction (skip)

Instruction

EDUCATION MODULE

If interview takes place between two school years, use alternative wording found in Appendix 1.

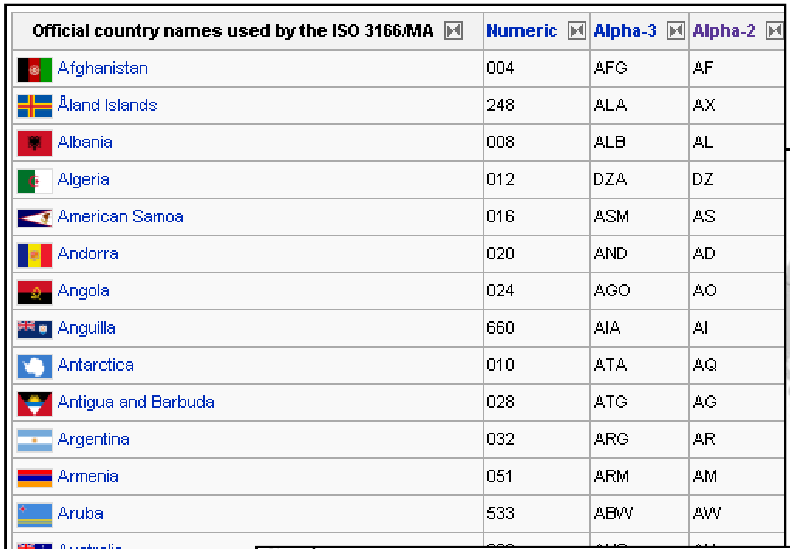
*For persons **age 5 or over** ask Qs. 15 and 16*

*For children **age 5 through 17 years**, continue on, asking Qs. 17-22*

14.	15.	16.		17.		18.		19.	20.		21.			22.	
Line no.	HAS (*name*) EVER ATTENDED SCHOOL?	WHAT IS THE HIGHEST LEVEL OF SCHOOL (*name*) ATTENDED? WHAT IS THE HIGHEST GRADE (*name*) COMPLETED AT THIS LEVEL?		IS (*name*) CURRENTLY ATTENDING SCHOOL?		DURING THE CURRENT SCHOOL YEAR, DID (*name*) ATTEND SCHOOL AT ANY TIME?		SINCE LAST *(day of the week)*, HOW MANY DAYS DID (*name*) ATTEND SCHOOL?	WHICH LEVEL AND GRADE IS/WAS (*name*) ATTENDING?		DID (*name*) ATTEND SCHOOL LAST YEAR?			WHICH LEVEL AND GRADE DID (*name*) ATTEND LAST YEAR?	
	1 YES ⇨ Q.16 2 NO ⇘ NEXT LINE	LEVEL: 1 PRIMARY 2 SECONDARY 3 HIGHER 4 NON-STANDARD CURRICULUM 9 DK GRADE: 99 DK *If less than 1 grade, enter 00.*		1 YES ⇨ Q.19 2 NO		1 YES 2 NO ⇨ Q.21		*Insert number of days in space below.*	LEVEL: 1 PRESCHOOL 2 PRIMARY 3 SECONDARY 4 NON-STANDARD CURRICULUM 9 DK GRADE: 99 DK		1 YES 2 NO ⇘ NEXT LINE 9 DK ⇘ NEXT LINE			LEVEL: 1 PRESCHOOL 2 PRIMARY 3 SECONDARY 4 NON-STANDARD CURRICULUM 9 DK GRADE: 99 DK	
LINE	Y NO	LEVEL	GRADE	YES	NO	YES	NO	DAYS	LEVEL	GRADE	Y	N	DK	LEVEL	GRADE
01	1 2⇨ NEXT LINE	1 2 3 4 9	__ __	1	2	1	2	___	1 2 3 4 9	__ __	1	2	9	1 2 3 4 9	__ __
02	1 2⇨ NEXT LINE	1 2 3 4 9	__ __	1	2	1	2	___	1 2 3 4 9	__ __	1	2	9	1 2 3 4 9	__ __
03	1 2⇨ NEXT LINE	1 2 3 4 9	__ __	1	2	1	2	___	1 2 3 4 9	__ __	1	2	9	1 2 3 4 9	__ __
04	1 2⇨ NEXT LINE	1 2 3 4 9	__ __	1	2	1	2	___	1 2 3 4 9	__ __	1	2	9	1 2 3 4 9	__ __
05	1 2⇨ NEXT LINE	1 2 3 4 9	__ __	1	2	1	2	___	1 2 3 4 9	__ __	1	2	9	1 2 3 4 9	__ __
06	1 2⇨ NEXT LINE	1 2 3 4 9	__ __	1	2	1	2	___	1 2 3 4 9	__ __	1	2	9	1 2 3 4 9	__ __
07	1 2⇨ NEXT LINE	1 2 3 4 9	__ __	1	2	1	2	___	1 2 3 4 9	__ __	1	2	9	1 2 3 4 9	__ __

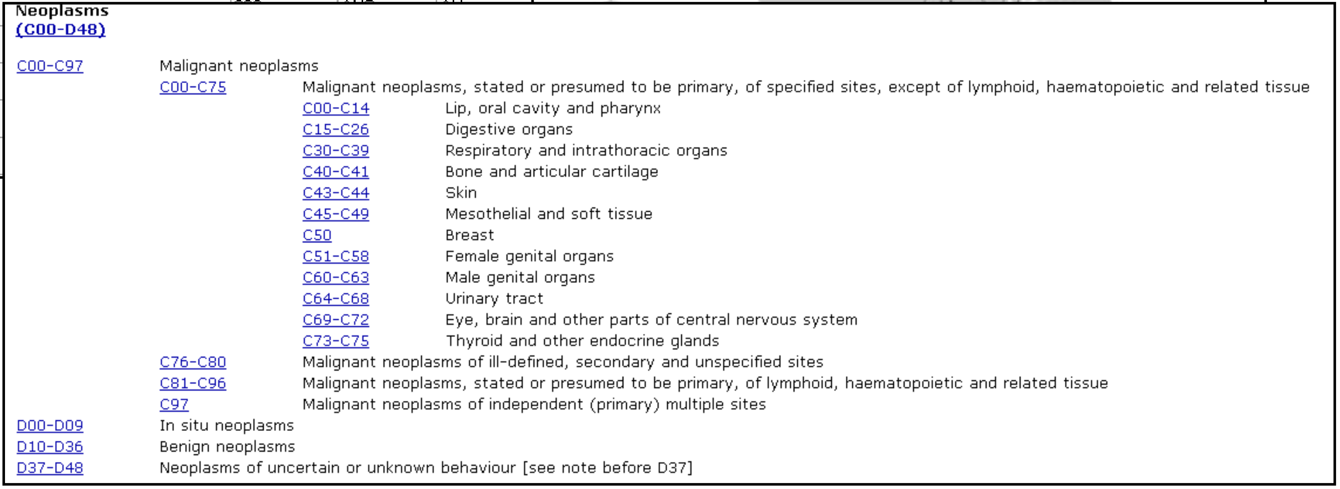
Now for each woman age 15-49 years, write her name and line number at the top of each page in the Women's Questionnaire.
For each child under age 5, write his/her name and line number AND the line number of his/her mother or caretaker at the top of each page in the Children's Questionnaire.
You should now have a separate questionnaire for each eligible woman and child in the household.

Common metadata example



Official country names used by the ISO 3166/MA	Numeric	Alpha-3	Alpha-2
Afghanistan	004	AFG	AF
Åland Islands	248	ALA	AX
Albania	008	ALB	AL
Algeria	012	DZA	DZ
American Samoa	016	ASM	AS
Andorra	020	AND	AD
Angola	024	AGO	AO
Anguilla	660	AIA	AI
Antarctica	010	ATA	AQ
Antigua and Barbuda	028	ATG	AG
Argentina	032	ARG	AR
Armenia	051	ARM	AM
Aruba	533	ABW	AW
Australia			
Austria			
Azerbaijan			
Bahamas			
Bahrain			

WA
OR
ID
MT
ND
MN
WI
MI
NY
VT
NH
ME
MA
CT
RI
CA
NV
UT
WY
SD
NE
IA
IL
IN
OH
PA
NJ
MD
DE
CO
KS
MO
KY
WV
VA
AZ
NM
OK
AR
TN
NC
SC
MS
AL
GA
TX



Neoplasms
(C00-D48)

- C00-C97 Malignant neoplasms
  - C00-C75 Malignant neoplasms, stated or presumed to be primary, of specified sites, except of lymphoid, haematopoietic and related tissue
    - C00-C14 Lip, oral cavity and pharynx
    - C15-C26 Digestive organs
    - C30-C39 Respiratory and intrathoracic organs
    - C40-C41 Bone and articular cartilage
    - C43-C44 Skin
    - C45-C49 Mesothelial and soft tissue
    - C50 Breast
    - C51-C58 Female genital organs
    - C60-C63 Male genital organs
    - C64-C68 Urinary tract
    - C69-C72 Eye, brain and other parts of central nervous system
    - C73-C75 Thyroid and other endocrine glands
  - C76-C80 Malignant neoplasms of ill-defined, secondary and unspecified sites
  - C81-C96 Malignant neoplasms, stated or presumed to be primary, of lymphoid, haematopoietic and related tissue
  - C97 Malignant neoplasms of independent (primary) multiple sites
- D00-D09 In situ neoplasms
- D10-D36 Benign neoplasms
- D37-D48 Neoplasms of uncertain or unknown behaviour [see note before D37]

Data with no or limited metadata

- Has little usefulness (low quality)
- Cannot be discovered
- Cannot be used effectively and responsibly
- Cannot be processes
- Is difficult to publish
- Cannot be cited
- Cannot be linked to other data or documents
- Increases burden on data provider
- …
- Incomplete metadata can even be more risky
  - users will “guess” which leads to disparate / contentious “valid” results
- This applies to the institutional, national and global levels
  - Metadata is not only for public use…

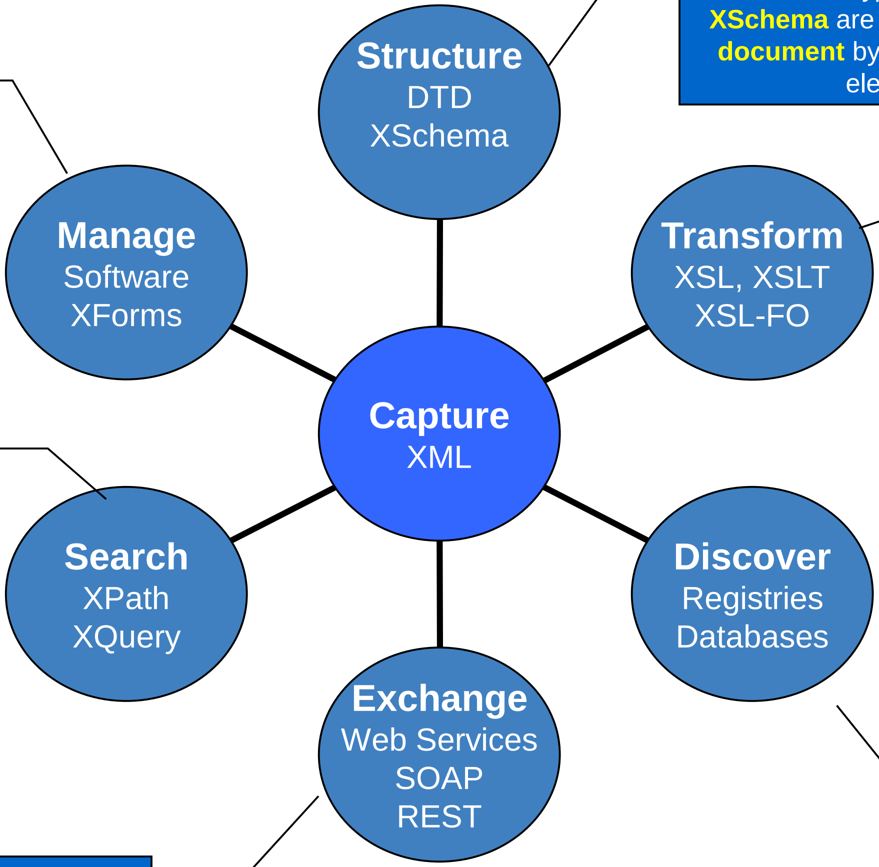
The e**X**tensible **M**arkup **L**anguage
The **D**ata **D**ocumentation **I**nitiative

Leveraging on industry standard technology
to support microdata management and research processes

What is XML?

- Today's Universal language on the web
- Purpose is to facilitate sharing of structured information across information systems in a generic fashion
- XML stands for e**X**tensible **M**arkup **L**anguage
  - eXtensibe → can be customized
  - Markup → tags, marks, attach attributes to things
  - Language → syntax (grammatical rules)
- HTML (**H**yper**T**ext **M**arkup **L**anguage) is a markup language but not extensible! It is also concerned about presentation, not content.
- XML is a text format (not a binary black box)
- XML is a also a collection of technologies (built on the XML language)
- It is platform independent and is understood by modern programming languages (C++, Java, .NET, pHp, perl, etc.)
- It is both machine and human readable

XML: an information management technology suite



Document **T**ype **D**efinition (DTD) and **XSchema** are use to **validate an XML document** by defining namespaces, elements, rules

Specialized software and database systems can be used to **create** and **edit** XML documents. In the future the XForm standard will be used

XML separates the metadata storage from its presentation. XML documents can be **transformed** into something else, like HTML, PDF, XML, other) through the use of the eXtensible Stylesheet Language, XSL Transformations (XSLT) and XSL Formatting Objects (XSL-FO)

Very much like a database system, XML documents can be **searched** and **queried** through the use of XPath oe XQuery. There is no need to create tables, indexes or define relationships

XML metadata or data can be **published** in “smart” catalogs often referred to as **registries** than can be used for discovery of information.

XML Documents can be sent like regular files but are typically **exchanged** between applications through Web Services using the SOAP and other protocols

The need for “standards”

- When sharing/exchanging/publishing information, we need to agree on a common set of similar elements and attributes to describe objects or concepts
  - Book, car, press releases, stock market, weather, etc.
  - Surveys, variables, questions, time series, classification, etc.
- In XML, this is a “specification” (DTD or Schema) that describes the information model
  - In some case this may be an official “standard” (i.e. ISO)
- Many different specifications exists for the different domains
- Typically maintained by consortium of organizations

Metadata specifications for SBE

- A single specification is not enough
  - We need a set of complementary metadata structures
    - That can map to each other to (maintain linkages)
    - Will be around for a long time (global adoption, strong community support)
    - Based on technology standards (XML)
- Suggested set
  - Data Documentation Initiative (DDI) – survey / administrative microdata
  - Statistical Data and Metadata Exchange standard (SDMX) – aggregated data / time series
  - ISO/IEC 11179 – concept management and semantic modeling
  - ISO 19115 – Geographical metadata
  - METS – packaging/archiving of digital objects
  - PREMIS – Archival lifecycle metadata
  - XBRL – business reporting
  - Dublin Core – citation metadata
  - Etc.

The Data Documentation Initiative

- The Data Documentation Initiative is an XML specification to capture structured metadata about “microdata” (broad sense)
- First generation DDI 1.0…2.1 (2000-2008)
  - Focus on single archived instance
- Second generation DDI 3.0 (2008)
  - Focus on life cycle
  - Go beyond the single survey concept
  - Multi-purpose
- Governance: DDI Alliance
  - Membership based organizations (35 members)
  - Data archives, producers, research data centers, academic
  - http://www.ddialliance.org

DDI Timeline / Status

- Pre-DDI 1.0
  - 70’s / 80’s OSIRIS Codebook
  - 1993: IASSIST Codebook Action Group
  - 1996 SGML DTD
  - 1997 DDI XML
  - 1999 Draft DDI DTD
- 2000 – DDI 1.0
  - Simple survey
  - Archival data formats
  - Microdata only
- 2003 – DDI 2.0
  - Aggregate data (based on matrix structure)
  - Added geographic material to aid geographic search systems and GIS users
- 2003 – Establishment of DDI Alliance
- 2004 – Acceptance of a new DDI paradigm
  - Lifecycle model
  - Shift from the codebook centric / variable centric model to capturing the lifecycle of data
  - Agreement on expanded areas of coverage

2005

Presentation of schema structure

Focus on points of metadata creation and reuse

2006

Presentation of first complete 3.0 model

Internal and public review

2007

Vote to move to Candidate Version (CR)

Establishment of a set of use cases to test application and implementation

October 3.0 CR2

2008

February 3.0 CR3

March 3.0 CR3 update

April 3.0 CR3 final

April 28th 3.0 Approved by DDI Alliance

May 21st DDI 3.0 Officially announced

Initial presentations at IASSIST 2008

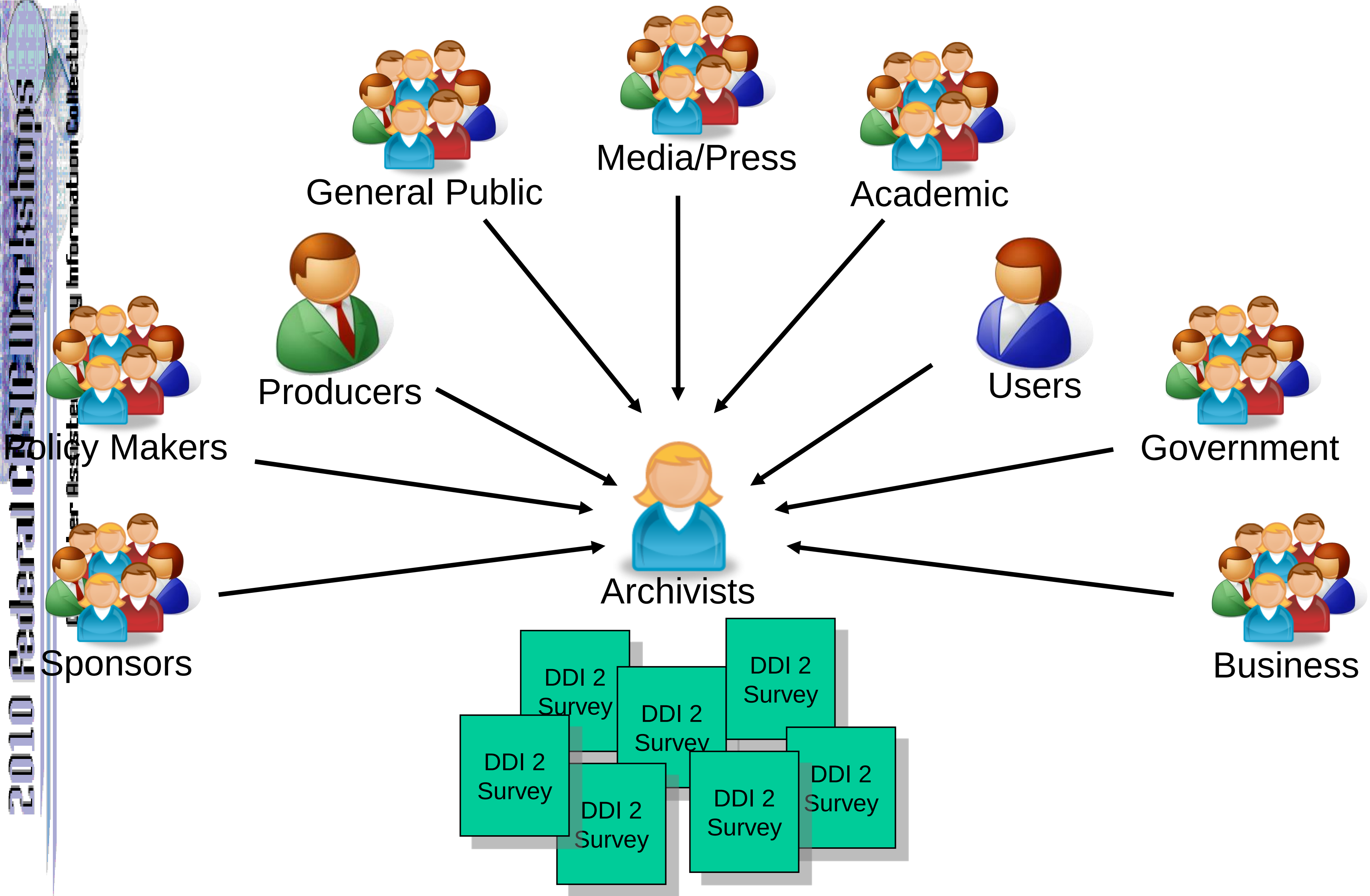
2009

DDI 3.1 and beyond

DDI 1.0 – 2.1 – Archival Metadata

- Focus on preservation of a survey
- Often see survey as collection of data files accompanied by documentation
  - Code book-centric
  - Report, questionnaire, methodologies, scripts, etc.
- Covers elements such as study, files, variables, questions, data cubes, geography, other materials
- Result in a static event: the archived survey
- Maintained by a single agency
- Is typically documentation after the facts
- Success story and widely adopted around the globe
- Tools available today
- Powerful but has limitations / constraints

DDI 2.0 perspective



General Public

Media/Press

Academic

Producers

Users

Policy Makers

Government

Archivists

Sponsors

Business

DDI 2 Survey

DDI 2 Survey

DDI 2 Survey

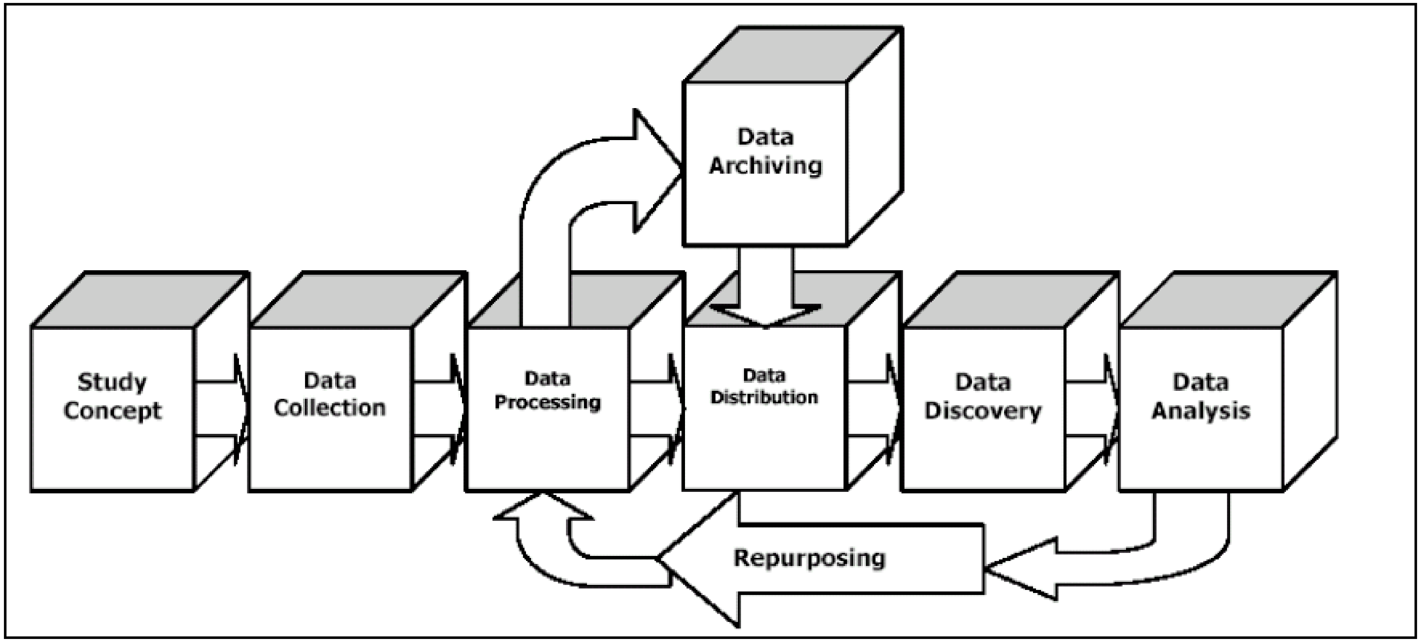
DDI 2 Survey

DDI 2 Survey

DDI 2 Survey

DDI 2 Survey

DDI 3.0 and the Survey Life Cycle

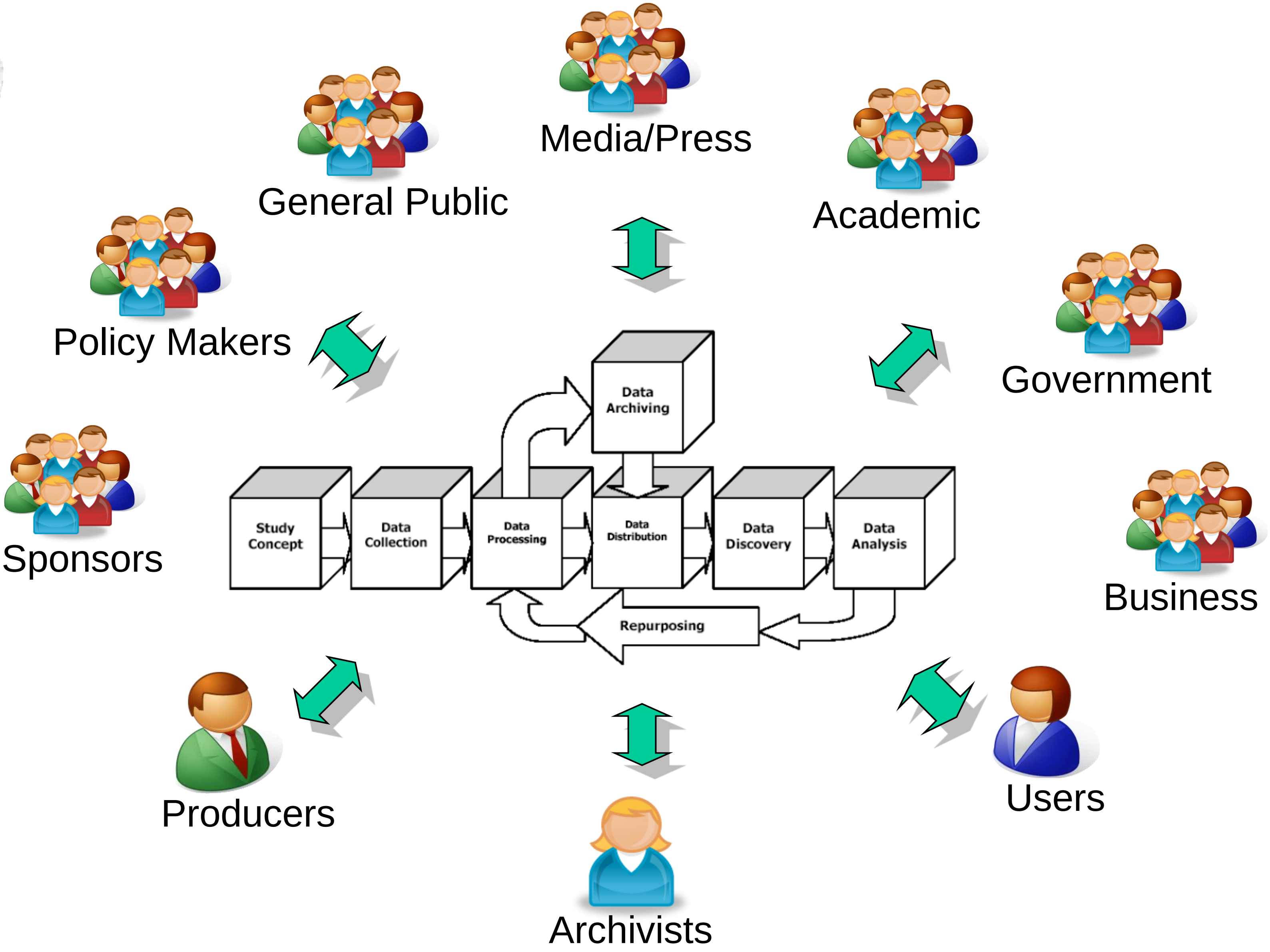


- A survey is not a static process: It dynamically evolved across time and involves many agencies/individuals
- DDI 2.x is about archiving, DDI 3.0 across the entire “life cycle”
- 3.0 focus on metadata reuse (minimizes redundancies/discrepancies, support comparison)
- Also supports multilingual, grouping, geography, and others
- 3.0 is extensible

DDI 3.0 Use Cases

- DDI 3 is composed of several schemas/modules
  - You only use what you need!
  - DDI 3.0 provides the common metadata language to maintain links and consistency across the entire life cycle
- Some examples
  - Study design/survey instrumentation
  - Questionnaire generation/data collection and processing
  - Data recoding, aggregation and other processing
  - Data dissemination/discovery
  - Archival ingestion/metadata value-add
  - Question /concept /variable banks
  - DDI for use within a research project
  - Capture of metadata regarding data use
  - Metadata mining for comparison, etc.
  - Generating instruction packages/presentations
  - Data sourced from registers
- The same specification is used across the lifecycle by different actors

DDI 3.0 perspective



General Public

Media/Press

Academic

Policy Makers

Government

Sponsors

Business

Study Concept

Data Collection

Data Processing

Data Archiving

Data Distribution

Data Discovery

Data Analysis

Repurposing

Producers

Users

Archivists

DDI 3 Relationship to Other Standards

- SDMX (from microdata to indicators / time series)
  - Completely mapping to and from DDI NCubes
- Dublin Core (surveys and documents gets cited)
  - Mapping of citation elements
  - Option for DC namespace basic entry
- ISO 19115 – Geography (microdata gets mapped)
  - Search requirements
  - Support for GIS users
- METS
  - Designed to support profile development
- OAIS (alignment of archiving standards)
  - Reference model for the archival lifecycle
- ISO/IEC 11179 (metadata mining through concepts)
  - Variable linking representation to concept and universe
  - Optional data element construct in ConceptualComponent that allows for complete ISO/IEC 11179 structure as a maintained item

What can DDI-XML do for you?

Why use XML?

- Industry standard
- Set of open technologies
  - Free, cross platform, embedded in IT tools, etc.
- Capture information in a non-proprietary format
- Can convert to traditional format
  - HTML, PDF, XLS, RTF, DOC, Text, etc.
  - *Not true the other way around!*
- Reduced development / implementation cost
- Allows for the reuse of tools
  - Collaborative efforts, no need to work in isolation
- Hybrid database systems understand XML
- Global adoption of standard
- Fit in public and private information networks
- …

Why use DDI?

- Builds on XML
- Internationally recognized specification
- Mature specification supported by a large community
- With DDI3:
  - Provides common framework / language across the entire life cycle
  - Allows for multiple contributors
  - Maximizes reuse!
  - Unique and persistent identifiers
  - Support for many use cases
- Works hand in hand with other XML specification / standards (from respondent to policy maker)

Leveraging on DDI-XML

- Unlock the data
- With human readable metadata: document your data!
  - But this is only part of the story
- With machine actionable metadata: automate processes:
  - Production, Archive / Preservation, Discovery / Dissemination, Use / Analysis / Repurposing
- Facilitate harmonization / comparability
- Manage “Banks” (question, variables, concepts, classifications)
- Provide public information on protected datasets
- Maintain institutional or national standards
- Bridge legacy / proprietary systems through standard based publication /exchange (crucial in federated environment)
- Explore new possibilities
  - Understand data usage, manage disclosure processes
- Plug into industry standard web services architecture
- Bridge to rich web applications, social networks and the semantic web
  - Foster user provided metadata

Where to start?

Components of a metadata driven framework

- Metadata surround data with:
  - human readable information (knowledge)
  - machine actionable information (processing / automation)
- XML
  - Provides the common language
  - Combines with a set of powerful industry standard open technologies to process / manage the metadata
- Standards
  - Common agreed upon structures that allows for publication, exchange, processing, reuse of tools, etc.
  - → Which one to use: DDI, SDMX, ISO 11179, Dublin Core, etc.
- Other necessary components for success
  - Institutional, national and international practices / endorsement
  - Guidelines, best practices, training
  - Tools + integration into existing environments / adoption by vendors
  - Public metadata registries / web services
  - Change management

Suggested Readings

- **"*Metadata*"**, Arofan Gregory (ODaF), Pascal Heus (ODaF), German Council for Social and Economic Data Working Paper no. 57/2009, March 2009, http://www.ratswd.de/download/workingpapers2009/57_09.pdf
- ***"DDI and SDMX: Complementary, Not Competing, Standards"***, A. Gregory, P. Heus, Open Data Foundation, July 2007
- ***"Combining Metadata Standards: Approaches and benefits"***, Arofan Gregory, Open Data Foundation, Work Session on Statistical Metadata (METIS) (Geneva, Switzerland, 10-12 March 2010), http://www.unece.org/stats/documents/ece/ces/ge.40/2010/wp.3.e.pdf
- ***"The Common Metadata Framework",*** http://www1.unece.org/stat/platform/display/metis/The+Common+Metadata+Framework
  - **Part A - Statistical Metadata in a Corporate Context**
  - Part B - Metadata Concepts, Standards, Models and Registries
  - Part C - Metadata and the Statistical Business Process
  - Part D - Implementation
- ***"Data Documentation Initiative: Toward a Standard for the Social Sciences"***, Mary Vardigan (ICPSR), Pascal Heus (ODaF), Wendy Thomas (MPC), International Journal of Digital Curation, Vol 3, No 1, Aug 2008, http://www.ijdc.net/index.php/ijdc/article/view/66
- ***"Data Access in a Cyber World: Making Use of Cyberinfrastructure"***, Julia Lane (NSF), Tim Mulcahy (NORC), Pascal Heus (ODaF), Transactions on Data Privacy (TDP), Volume 1, Issue 1, 2008, http://www.tdp.cat/issues/abs.a002a08.php
- See also
  - http://odaf.org/?lvl1=resources&lvl2=papers
  - http://www.ddialliance.org/resources/publications
  - http://www.sdmx.org

Internet Resources

- DDI Alliance - http://www.ddialliance.org
- SDMX - http://www.sdmx.org
- METIS - http://www1.unece.org/stat/platform/display/metis/METIS-wiki
  - UNECE METIS Work Session on Statistical Metadata (Geneva, 10-12 March 2010) - http://www.unece.org/stats/documents/2010.03.metis.htm
- Metadata Technology - http://www.metadatatechnology.com
- Open Data Foundation - http://www.opendatafoundation.org
- IASSIST - http://www.iassistdata.org/
  - 2010 Conference - http://ciser.cornell.edu/IASSIST/

Data.gov

Use case: About data.gov

- The purpose of Data.gov is to **increase public access** to high value, **machine readable datasets** generated by the Executive Branch of the Federal Government.
- As a **priority** Open Government Initiative for President Obama's administration, Data.gov increases the ability of the public to **easily find**, **download**, and **use datasets** that are generated and held by the Federal Government. Data.gov **provides descriptions** of the Federal datasets (**metadata**), information about **how to access** the datasets, and tools that leverage government datasets. The data catalogs will continue to **grow as datasets are added**. Federal, Executive Branch data are included in the first version of Data.gov.

Use case: About data.gov

- **Public participation and collaboration** will be one of the keys to the success of Data.gov. Data.gov enables the public to participate in government by providing downloadable Federal datasets to build applications, conduct analyses, and perform research. Data.gov will continue to improve based on **feedback, comments, and recommendations** from the public and therefore we encourage individuals to **suggest datasets** they'd like to see, rate and comment on current datasets, and suggest ways to improve the site.

Use case: About data.gov

- A **primary goal** of Data.gov is to **improve access** to Federal data and expand creative use of those data beyond the walls of government by **encouraging innovative ideas** (e.g., web applications). Data.gov strives to make government **more transparent** and is committed to creating an **unprecedented level of openness** in Government. The openness derived from Data.gov will strengthen our Nation's democracy and promote efficiency and effectiveness in Government.
- → DDI, SDMX and related standards have been designed to answer such mandate.
  - particularly relevant in the highly federated US statistical system