

HARMONIZATION AND INTEGRATION OF METADATA

AN URGENT TASK FOR FUTURE
EFFICIENT USE OF THE WEB

Prepared by Dusan Soltes, FM CM
BRATISLAVA, SLOVAKIA

for the IASSIST'2002 at the UCONN
Storrs (USA)

WWW

- HAS ACHIEVED A HIGH LEVEL OF HARMINIZATION ALL OVER THE WORLD:
- HARDWARE
- SOFTWARE
- TELECOMS
- LANGUAGE
- NAVIGATION...HOWEVER...

WWW

- HAS NOT YET ACHIEVED:
- HARMONIZATION OF DATA CONTENT
- DEFINITION OF DATA
- INTERPRETATION OF INFORMATION
- DIRECT AND/OR EASY COMPARABILITY OF DATA /INFORMATION FROM DIFFERENT SOURCES

- NEW ROLE AND CHALLENGE FOR
- METADATA/METAINFORMATION
- AS SPECIFIC DATA/INFORMATION
- ON OTHER (OBJECT) DATA AND/OR INFORMATION

METADATA ITSELF NEEDS HARMONIZATION

- AS THERE ARE:
- - TOO MANY THEORIES
- - MODELS
- - APPROACHES
- - VARIOUS TOOLS AND TECHNIQUES
- - DIFFERENT INTERPRETATIONS
- - OBJECTS
- - ATTRIBUTES
- - ...

ONE OF INTERNATIONAL EFFORTS IS METANET

- A NETWORK OF EXCELLENCE FOR HARMONIZING AND SYNTHESISING THE DEVELOPMENT OF STATISTICAL METADATA UNDER THE EU'S METANET PROJECT

METANET

- FOUR WORKING GROUPS:
- - METADA TOOLS AND MODELS
- - METADA CONTENT-SEMANTICS
- - METADA FUTURE DEVELOPMENT
- - METADATA USERS' ASPECTS

WG 2 – METADATA CONTENTS AND SEMANTICS

- THERE ARE AS MANY APPROACHES AS THERE ARE MANY VARIOUS MODELS OF METADATA
- ALTOGETHER WE HAVE IDENTIFIED OVER 30 OF DIFFERENT METADA MODELS/APPROACHES WITH DIFFERENT CONTENTS, ETC.

METADATA INVENTORY AND ANALYSIS

- METADATA STANDARDS AND TOOLS
ASSESSMENT FORM:
- - NAME OF THE STANDARD OR TOOL
- CONSIDERED
- - REFERENCE
- - REVIEWER
- - VERSION

METADATA STANDARDS AND TOOLS ASSESSMENT FORM

- CHARACTERISTICS & DESCRIPTION:
 - - ORIGIN OF MODEL OR STANDARD
 - - STANDARDISATION STATUS RESP.
- FORMALISATION
 - - MAIN PURPOSE OF STANDARD OR MODEL
- - OPPORTUNITIES FOR SECONDARY USAGE, OTHER FUNCTIONALITIES
- - ACTUAL USER COMMUNITY

CONTINUATION OF CHARACTERISTICS (1)

- - ABSTRACTION LEVEL
- - CONTENTS DESCRIBED BY THE
STANDARD OR MODEL
- - WHICH OBJECTS ARE CONSIDERED?
- - WHICH PROCESSES ARE
CONSIDERED?
- - SEMANTIC VIEW
- - STATISTICAL VIEW
- - INFORMATIC VIEW

CONTINUATION OF CHARACTERISTICS (2)

- - ADMINISTRATIVE VIEW
- - FORMAT (IF RELEVANT)
- - CORRESPONDING UML
REPRESENTATION
- - RELATION TO XML
- - FREE SOFTWARE AND INTERFACES
- - COMMERCIAL SOFTWARE, INTERFACES,
SUPPORT, CONSULTANCY
- - SPECIAL TERMINOLOGY
- - SAMPLE TERMS

PRELIMINARY RESULTS OF ANALYSIS

- MANY VARIOUS SYSTEMS
- TOO MANY ARE EITHER TOO TECHNICAL OR TOO GENERAL OR TOO SPECIFIC FOR ONE GIVEN SYSTEM ONLY
- SOME GOOD EXAMPLES:
 - - DUBLIN CORE
 - - NEUCHATEL GROUP
 - - UN/ECE/METIS

DUBLIN CORE (ISO/IEC 11179-3 BASED)

- 15 ATTRIBUTES (6 ISO COMPLYING):
 - - TITLE
 - - CREATOR
 - - SUBJECT
 - - DESCRIPTION
 - - PUBLISHER
 - - CONTRIBUTOR
 - - DATE
 - - RIGHTS
 - TYPE
 - FORMAT
 - IDENTIFIER
 - SOURCE
 - LANGUAGE
 - RELATION
 - COVERAGE

NEUCHATEL GROUP (SWE, DEN, NOR, SWI)

- CLASSIFICATION OBJECTS AND ATTRIBUTES:
 - - CLASS.NAME
 - - IDENTIFIER
 - - TITLE
 - - DESCRIPTION
 - - OBJECT/UNIT
 - - CONTEXT
 - - SUBJECT AREA
 - - OWNERS
 - - FAMILY
 - - VERSIONS
 - - CURRENT VERS'N
 - - ITEM HISTORY
 - - KEYWORDS
 - - FOOTNOTES

FURTHER PROCEEDING:

- - FINALIZE INVENTORY
- - CREATE META-META DATA BASE
- - ANALYSIS OF MMDB ENTRIES
- - SELECTION OF MOST SUITABLE
- - CREATE STANDARD PROPOSAL
- - ADOPTION OF THE STANDARD
- - WWW IMPLEMENTATION

