



InterPARES² Project
International Research on Permanent Authentic Records in Electronic Systems

Description Cross Domain

Presented by Professor Sue McKemmish
Co-Chair Description Cross Domain
Chair of Archival Systems, Monash University
30 September 2004

Overview

- Objectives
- Research Activities
- Research Outputs
- Progress To Date
- Challenges
- Publications and Presentations

This slide provides an overview of my presentation. I will cover the objectives of the DCD, our research activities and planned outputs, progress to date, the challenges we face, and our dissemination activities.

Objectives

Evaluate whether existing and emerging metadata schemas, descriptive standards and metadata tools meet recordkeeping metadata and archival description requirements relating to the appraisal, creation, management, preservation and use of reliable and authentic records in e-government, artistic activities and scientific activities

Recommend the development of new metadata schemas, descriptive standards and metadata tools, or the extension of existing/emerging ones

Develop metadata specification models and specifications for metadata tools to support this

These are our main objectives as set down in our research plan. In order to achieve these objectives we need to work closely with all the other research teams in InterPARES 2.

Research Activities

- Analysis of Case Studies
- Literary Warrant Database
- Metadata Schema Registry
- Metadata Specification Models
- Evaluation and Specification of Metadata Tools
- Interface with other relevant R&D initiatives:
 - ISO 23081 Records Metadata Standard
 - Monash Clever Recordkeeping Metadata Project
 - SDSC work on metadata tools development
- Dissemination of research

This is a list of our research activities which I'll describe in a little more detail.

A couple of things to note about this list:

- The Metadata Schema Registry which is both a major research instrument and one of our deliverables in its own right will be demonstrated to you later today.
- We have been working closely with the ISO Working Group responsible for the Metadata Standard and, as Luciana said, we are in the process of establishing a formal relationship between the ISO and InterPARES to facilitate that. The first part of the ISO standard provides a benchmark against which we are evaluating metadata schemas, including archival description standards. At the same time outputs from the Metadata Registry will constitute Part III of the ISO, and our research findings are informing the development of Part II which relates to implementation
- The Monash Clever Recordkeeping Metadata Project which focuses on implementation issues relating to the creation, use, re-use and re-purposing of metadata for business, current recordkeeping and archival description purposes is funded by the ARC and is also closely linked to InterPARES 2. My Co-Chair Professor Anne G-S from UCLA is a Partner Investigator in that project and other members of the Australian Team are also involved in it.

Research Activities

Analysis of Case Studies

Identify use of metadata, schemas and standards, and related issues (**Focus Groups**)

Literary Warrant Database

Identify warrant, or authoritative sources relating to metadata-related requirements for the accuracy, reliability, and authenticity of records and the ways in which metadata and archival description can support appraisal, creation, management, preservation and use (**All Teams**)

These next few slides give you a bit more detail on our research activities. I have highlighted in red the links with other InterPARES work and other R&D initiatives.

Metadata Schema Registry

- Database which describes, analyses and evaluates schemas with reference to ISO, InterPARES 1 Benchmark and Baseline Requirements and Australian RKMS ([link to Monash CRKM](#))
- Centralized repository of information to assist organizations and archival institutions in identification of relevant schemas/combinations of schemas to meet their specific needs
- Standardized framework for assessing functionality of schemas, e.g. by standards-setting bodies ([link to ISO 23081 Part III](#))

As I mentioned we will demonstrate this for you later. It has multiple purposes as ...

Metadata Specification Models

- Develop and test metadata specifications relating to InterPARES models using case study data (**Focus Groups**)
- Validate the metadata model through walk-throughs using case study data and scenarios
- Involves identifying requirements re what metadata needs to be created at different points in the models, how, and by whom (**Modeling Team**)
- Includes specifying type, source and application of metadata, relevant schemas and standards, and requirements for new or extended ones (**using Literary Warrant Database, case study analyses and Metadata Schema Registry**)

We are also working closely with the Modeling team to develop metadata specifications relating to the models they are developing.

Metatool Evaluation and Specification

- Evaluate usefulness of existing tools
- Develop specifications for metadata management tools for activities such as automatic metadata specification and extraction

(Links to SDSC work on metatools and Monash CRKM Project)

And we are looking at the functionality of metatools.

Research Outputs

Terms and proposed definitions relating to metadata and archival description submitted for inclusion in InterPARES 2 Glossary and Thesaurus

Recommendations and model specifications for extension of existing generic and sector specific schemas and standards or development of new ones to meet recordkeeping metadata and archival description requirements, particularly in e-government, artistic and scientific communities

Analytical data on the adequacy of existing metadata schemas and descriptive standards as input to the draft ISO 23081, Part III

The next couple of slides list our planned research outputs. Work is underway in relation to all of these.

Research Outputs

Other policy and standards recommendations
(Policy Cross Domain; ISO; Standards Australia)

Metadata Schema Registry

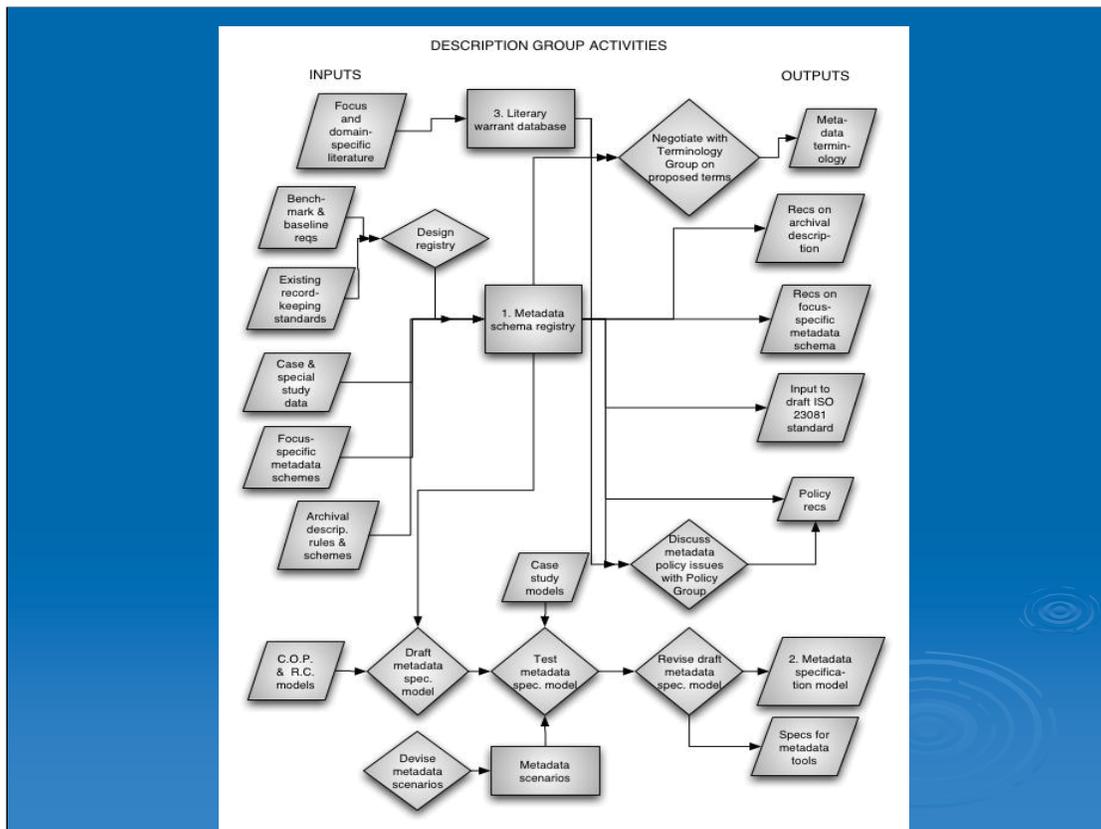
Literary Warrant Database

Specifications for Metadata Tools

Contribution of new knowledge, understandings,
models and tools to wider metadata community

Collaborations/sharing research outputs with other
R&D initiatives, including ISO, Monash CRKM
project, SDSC projects

Publications and presentations



This diagram is still being developed. It represents some of our main activities, the related inputs and outputs, indicating also the inter-relationships with the work of other IP researchers and other R&D initiatives, for example in relation to the Metadata Schema Registry, it references inputs including the IP case studies, and external standards and benchmarks, as well as outputs including recommendations re improving sector specific schemas in the arts, scientific and e-government domain, inputs to Part III of the ISO, and recommendations re policy development, for instance in archival description which could be fed into the ICA's frameworks.

Progress To Date

- Case Study Analysis – framework developed
- Literary Warrant Database – designed, developed and being populated
- Metadata Registry – designed, developed, and being populated
- Metadata Specification Model and Metatool Specification – about to commence
- Evaluation of Metadata Tools – in progress
- Definition of Input to ISO and Establishment of Formal Arrangements – in progress
- Publications and Presentations

This is a quick overview of our progress to date.

Challenges

- Geography
- Collaborative Research
- Communication
- Coordinating interaction with other teams
- Optimising links to related research initiatives
- Cost-effective use of resources

There are many challenges in engaging in international collaborative Research involving academics, students, and practitioners across domains, disciplines and hemispheres. Some of them are listed on this slide – and Luciana has addressed this question more comprehensively and for the Project as a whole in her presentation. DCD has met these challenges through:

- Careful project Management, including research planning; specification of activities and timelines; allocation of responsibilities and roles
- Communication strategies which include regular and irregular face-to-face meetings of the whole or parts of the team, and virtual working meetings via web conferencing/net meeting
- Working with other InterPARES teams via cross membership of teams, liaison people, joint meetings
- Establishing formal links with related projects – ISO, Monash CRKM, SDSC – and coordinating research planning to maximise sharing of inputs and outputs
- Leveraging the resource mix, which includes funding from SSHRC, NHPRC, UCLA, Monash, National Archives Australia, Netherlands National Archives, and SDSC.

Publications and Presentations

- ICA Vienna, August 2004: Session on Smart Metadata and Archives of the Future – 4 inter-related papers
- InterPARES West Coast Symposium, UCLA, May 2004 – showcasing InterPARES Description Cross Domain work
- Archival Science x 3 (forthcoming)
- Shanghai DC 2004 Conference (October 2004)
- 8 journal and conference papers, 2002-2004
- Future publication strategy

Finally a summary of publications and presentations. We have also developed a future publications strategy which targets different domains and stakeholders, e.g. IEEE Journal re technical aspects of the Registry and JASIST re the metadata schema we developed for the Registry itself, and the journals and conferences of the focus group domains.