Proceedings of the International Conference on Dublin Core and Metadata Applications 2013

Designing an Application Profile using qualified Dublin Core: A case study with fracture mechanics datasets

João Aguiar Castro joaoaguiarcastro@gmail.com
Cristina Ribeiro mcr@fe.up.pt
João Rocha da Siva joaorosilva@gmail.com

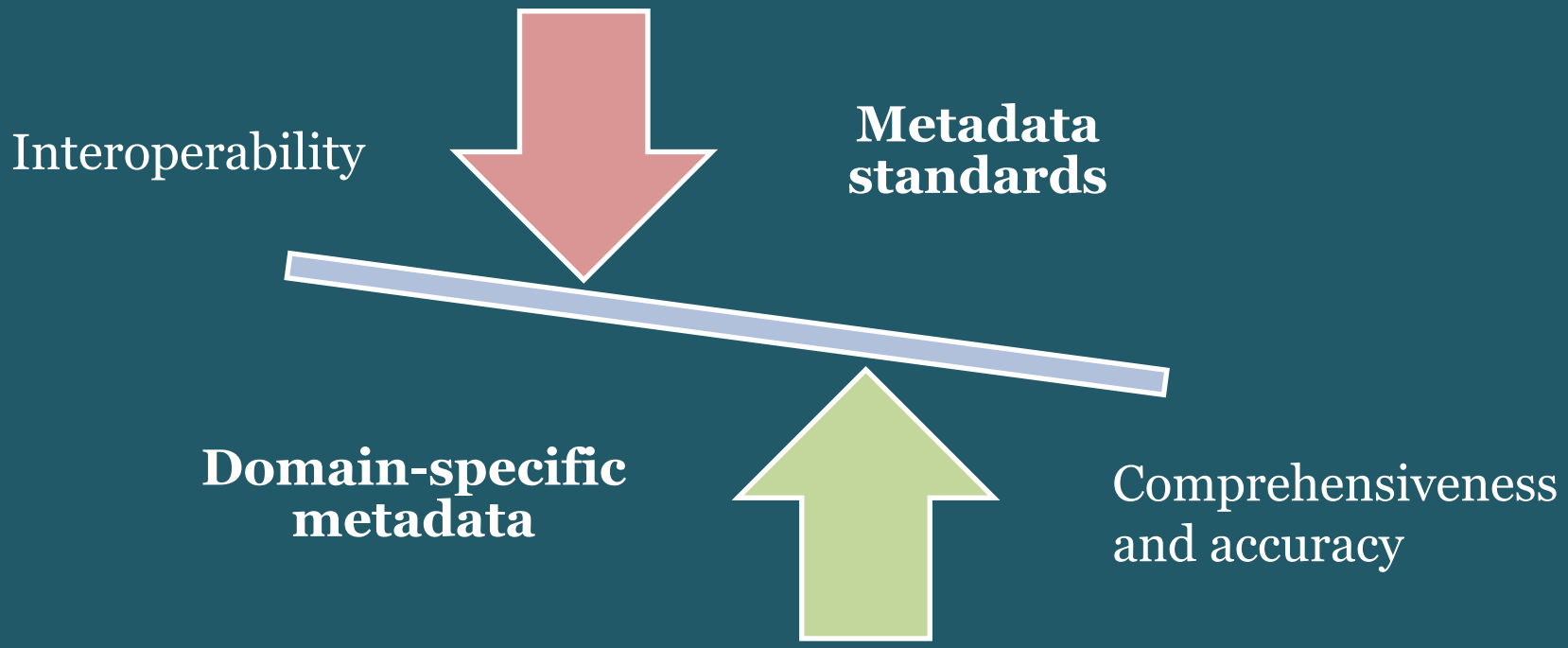
Lisbon, September 2013

CONTENT

- Introduction / Related Work
- Research data management workflow at the University of Porto
- Data management in the research group – Current pratices and requirements
- Application Profile for Fracture Experiments
- Conclusion / Future Work

INTRODUCTION

Metadata for research data management



"mix and match" approach as the best alternative

Application Profile:
- *"set of data elements (...) from one or more schemas combined together"*
(Heery and Pattel, 2000)

Heery, Rachel and Manjula Patel (2000). Application Profiles: Mixing and Matching Metadata Schemas. Ariadne 25, September 2000.

RELATED WORK

ANDS	ISO RIF-CS as data interchange format
DANS	Dublin Core + D. D. I. + other domain specific standards
DataOne, EDINA	Data management plans recommend data documentation from early stages; also recommends standards
Dryad Application Profile	In conformance with DC guidelines
University of Edinburgh*	Institutional data repositories – DC metadata

**Rice, Robin (2008). Applying DC to Institutional Data Repositories. Proceedings of the International Conference on Dublin Core and Metadata Applications, North America, 2008*

RESEARCH DATA MANAGEMENT AT THE UNIVERSITY OF PORTO

Data Repository (Dspace)

Dataset audit at University of Porto (iPRES 2012):

- ✓ Researcher needs and requirements
- ✓ Dataset samples from several domains
- ✓ Prototype tested with real datasets

UPBox ("Dropbox")

- ✓ Easy storage & collaboration
- ✓ Full institutional control over data

DataNotes (Wiki)

- ✓ Collaborative data annotation
- ✓ Offers *Application Profiles* for easier description

Greater researcher autonomy in data management and description

DATA MANAGEMENT IN THE RESEARCH GROUP - CURRENT PRACTICES AND REQUIREMENTS

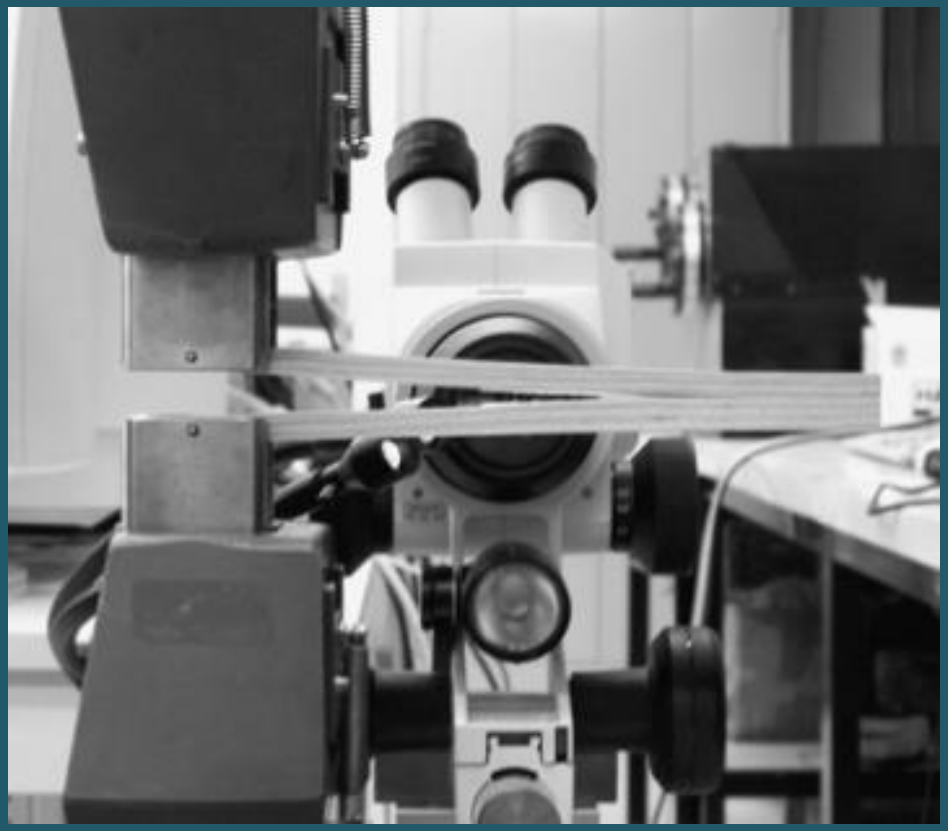
Approach

1. Introduce the two management applications
2. Provide insight about data curation and management
3. Make researchers familiar with “metadata” and “descriptors” concepts

Assess data management practices – Script adapted from the Data Curation Tool Kit

Pardue University & University of Illionais at Urbana-Champaign

DATA MANAGEMENT IN THE RESEARCH GROUP - CURRENT PRACTICES AND REQUIREMENTS



The fracture mechanics experiments

- Cantilever fracture tests
- Force is applied to a sample up to the point of fracture
- Specialized equipment measures the evolution of force and the cantilever displacement

DATA MANAGEMENT IN THE RESEARCH GROUP
CURRENT PRACTICES AND REQUIREMENTS

Data lifecycle

1– Data is captured by proprietary software, producing an Excel spreadsheet

	A	B	C	D	E	F
1	Provete 1		Provete 2		Provete 3	
2	Forca	Desloc	Forca	Desloc	Forca	Desloc
3	(N)	(mm)	(N)	(mm)	(N)	(mm)
4	1.2	0.003	0.4	0.003	0.92	0.003
5	1.56	0.003	1.04	0.003	1.4	0.003
6	1.2	0.003	0.88	0.003	1.2	0.003

2 – Data is processed and converted into energy values

Also – Data sharing among the research group (inter-university) without established procedures

DATA MANAGEMENT IN THE RESEARCH GROUP
CURRENT PRACTICES AND REQUIREMENTS (CONT'D)

Issues

- ✘ No data management plan
- ✘ Lack of awareness to the need of long-term preservation
- ✘ No standardized metadata
- ✘ Finding a particular document consumes too much time

APPLICATION PROFILE FOR FRACTURE EXPERIMENTS

Requirements for the documentation of research data (Willis et al. 2012)

- Comprehensive and sufficient
- Simple
- Data interchange
- Enhance data documentation
- Retrieval

Dublin Core
+
EML (Ecologic Metadata Language)
+
Domain-specific elements

Dublin Core terms used in the Fracture Mechanics application profile (fm.)

DC		Qualified DC
dc.title (required)	dc.date (required)	dc.References (if available)
dc.subject (required)	dc.description (required)	dc.isReferencedBy (if available)
dc.Identifier (if available)		dc.medium (recomended)

Willis, Craig, Jane Greenberg, Holly White (2012). Analysis and Synthesis of Metadata Goals. Journal of the American Society for Information Science and Technology 63, 8: 1505 -152.

APPLICATION PROFILE FOR FRACTURE EXPERIMENTS

EML and domain-specific metadata elements

Namespace	Description	Example
eml.methods	Actual procedures that are use in the creation or the subsquent processing of a dataset	Free text
eml.instrumentation	The Instrumentation field allows the description of any instruments used in the data collection ore quality control and quality assurance	INSTROM-1125
fm.specimen	A word or phrase describing the type of specimen collected	Wood (pinus pinaster)
fm.specimenLenght	Specimen geometric lenght	400 mm
fm.specimenWidth	Specimen geometric width	20 mm
fm.specimenHeight	Specimen geometric height	20 mm
fm.specimenInitialCrackLenght	The initial crack of the DCB specimen prior to the fracture test	150 mm
fm.specimenProperties	Specific proporties related to the specimen, such as elastic, stress properties	Free text
fm.temperature	The ambient temperature of the location where the experiment was performed, in degrees Celsius	18º
fm.moisture	The moisture percentage at the location where the experiment was performed	55%
fm.testVelocity	Velocity at which the experiment was performed (millimeters per minute)	3 mm/m

CONCLUSION AND FUTURE WORK

✓ **Promoting the deposit of research data**

Helping researchers take an active part in the organization and description of project data; better awareness of the value of well-described datasets facilitates the transition to a public data repository

✓ **Application Profile validated**

Preliminary usage demonstrate improvements in the data management workflow; paving the way for its use by groups working on the same domain

✓ **Work with groups from different domains**

To provide further insight on how to improve the management tools

As more researchers describe and share datasets, the repository can grow

These datasets can be cited, making it worthwhile to curate them

João Aguiar Castro joaoaguiarcastro@gmail.com
Cristina Ribeiro mcr@fe.up.pt
João Rocha da Siva joaorosilva@gmail.com

End