METADATA

BUILDING BLOCKS FOR KNOWLEDGE

Nasreen Khan

Digital Resources Management from Data to Knowledge
21-23 July 2009, Durban, South Africa

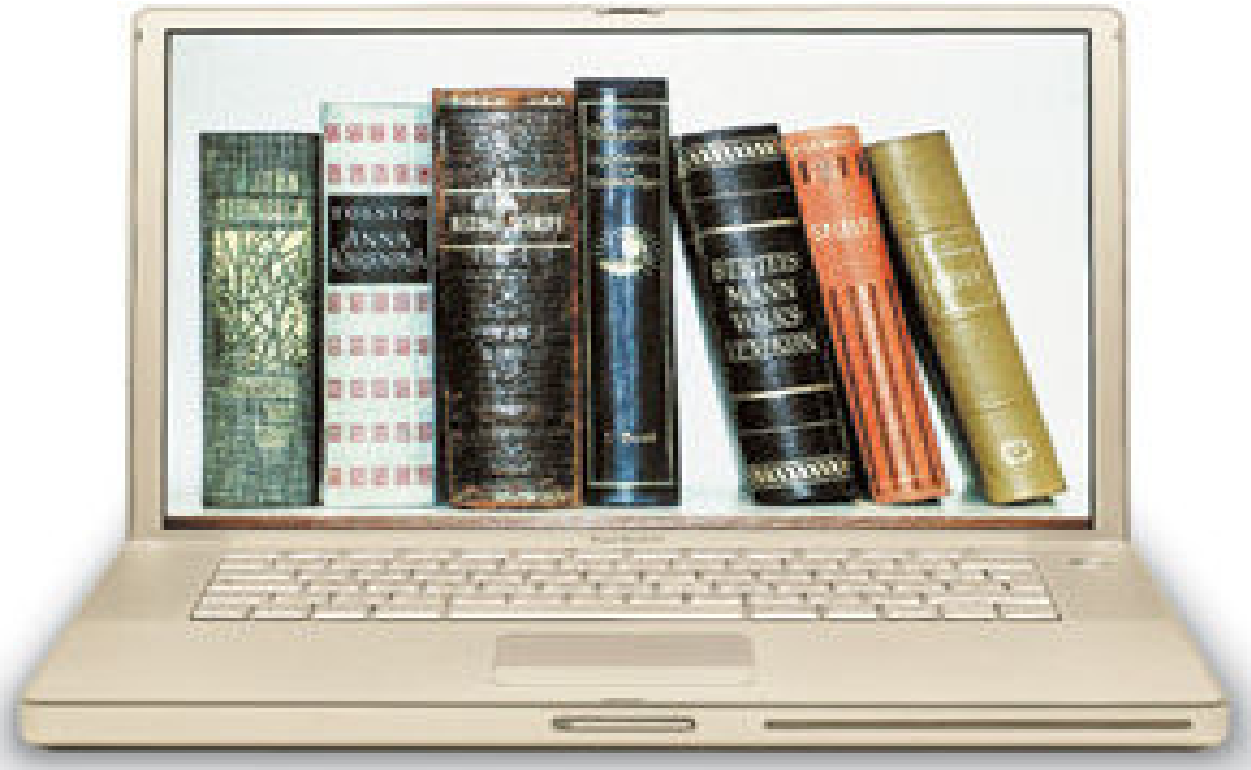
DISA
DIGITAL INNOVATION SOUTH AFRICA

Overview

- To describe and explore the similarities and differences of traditional libraries and digital libraries focusing on metadata
- To understand and identify metadata, types and functions
- To introduce the Dublin Core

DISA
DIGITAL INNOVATION SOUTH AFRICA

Reading on the laptop…



DISA
DIGITAL INNOVATION
SOUTH AFRICA

TRADITIONAL vs DIGITAL LIBRARIES

Traditional libraries

- Emphasis on storage and preservation of physical items such as books and periodicals , videos etc.
- Cataloguing at a high level rather than one of detail, eg. Author and subject indexes as opposed to full text.
- Browsing based on physical location of materials, eg. Books on politics will be close together on the shelves
- Users need to visit the library at a physical location to make use of the material.

Digital libraries

- Emphasis on access of digitized materials wherever they may be located
- Cataloguing down to individual words.
- Browsing based on hyperlinks, keywords, or any other defined relation; materials on the same subject need not be in the same location in any physical sense
- Users need to make use of the world wide web to access the digital library.

DISA
DIGITAL INNOVATION SOUTH AFRICA

Finding physical resources and information

- In everyday life
  - clothing tags
  - television guides to programmes
  - information on packaging
- catalogues are useful for finding resources
  - Concise information representing the resource
- Indexes, directories are useful for finding resources
  - Lists guiding you to the resource

DISA
DIGITAL INNOVATION SOUTH AFRICA

Digital resources

- Using computers to create, manage and store information in digital format

and

- Using computers to access these materials in digital format
  - Collections of digital files or datastreams
  - Digital files managed in a different way
  - End-users are accessing digital resources without going to a physical location

7

DISA
DIGITAL INNOVATION SOUTH AFRICA

World Wide Web

- looking for information on the Web
  - Large returns of irrelevant information
  - Browsing is time consuming
  - Clicking through numerous links and pages
  - Abandoned searches
- you can't always find what you are looking for …

DISA
DIGITAL INNOVATION SOUTH AFRICA

What is missing?

- pathways to the resources
- lack of information about the resources on the web
  - concise descriptions
  - consistent descriptions
  - structured descriptions
- “data about data” - metadata

DISA
DIGITAL INNOVATION SOUTH AFRICA

What is Metadata?

- Data about Data
- Information about Data
- Of any sort and in any format
- Readable by computers
- Understandable by humans

DISA
DIGITAL INNOVATION
SOUTH AFRICA

Definition of Metadata

- Metadata describes other data and provides information about a certain item’s content.

- For example, an image may include metadata that describes how large the picture is, the colour depth, the image resolution, when the image was created, and other data. A text document's metadata may contain information about how long the document is, who the author is, when the document was written, and a short summary of the document.

DISA
DIGITAL INNOVATION SOUTH AFRICA

Why use metadata?

- Allows for **standardised**, precise description of resources
- Allows for **sharing** of information
- Enables resource **discovery**
- Enables **interoperability** between institutions
- **Authentication** and **reliability** of data

12

DISA
DIGITAL INNOVATION SOUTH AFRICA

IFLA's functional requirements for bibliographic/metadata records (FRBR)

- Find
- Identify
- Select
- Obtain

DISA
DIGITAL INNOVATION SOUTH AFRICA

Types of Metadata

- Descriptive
- Administrative (including Rights)
- Preservation
- Technical
- Structural

DISA
DIGITAL INNOVATION SOUTH AFRICA

Descriptive Metadata

- Title,
- Author
- Description of resource
- Subjects and Keywords
- Publisher
- Date

- Information describing the intellectual content of the object
- Descriptive metadata = cataloguing

DISA
DIGITAL INNOVATION SOUTH AFRICA

Structural Metadata

- Information that ties each object to make up logical units.
- Information that relates individual images of pages from a book to the others that make up the book.

DISA
DIGITAL INNOVATION SOUTH AFRICA

Administrative and Rights Metadata

- Information used to manage the object or control access to it.

- Included here is information on how it was scanned, its storage format, copyright, and the necessary information required for the long-term preservation of digital objects.

DISA
DIGITAL INNOVATION SOUTH AFRICA

International Standards and Schemes

- Metadata should be created according to appropriate international standards or schemes.
- This will depend on the type and format of the physical object which is being described
- Types include: Dublin Core (DC); Text encoding Initiative (TEI) and Encoded Archival Description (EAD)

DISA
DIGITAL INNOVATION SOUTH AFRICA

Dublin Core

- “Dublin” in the name Dublin Core refers to Dublin, Ohio, U.S., where the work originated from an original workshop hosted by OCLC IN 1995
- “Core” refers to the fact that the metadata element set is a basic yet expandable core list of elements
- All fields on Dublin Core are optional and all are repeatable
- 1 Metadata record = 1 Item

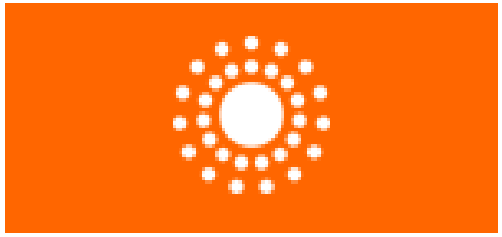
DISA
DIGITAL INNOVATION SOUTH AFRICA

Dublin Core Metadata Elements

Title	Creator
Subject	Publisher
Description	Contributor
Language	Rights
Source	Date
Relation	Format
Coverage	Identifier
	Type

DISA
DIGITAL INNOVATION SOUTH AFRICA

Dublin Core



- international standard for metadata creation
- simple element set – 15 elements
- describing simple resources (eg photos)
- “simple” and “qualified”
- expressed in XML – “Interoperability Standard”

DISA
DIGITAL INNOVATION SOUTH AFRICA

What is XML?

- eXtensible Mark-up Language
- XML is for structuring data
- XML looks a bit like HTML
- XML is text but isn't meant to be read
- XML is a family of technologies – X-link, XSL
- XML is license-free and platform-independent
- *De facto* standard for metadata on the Web

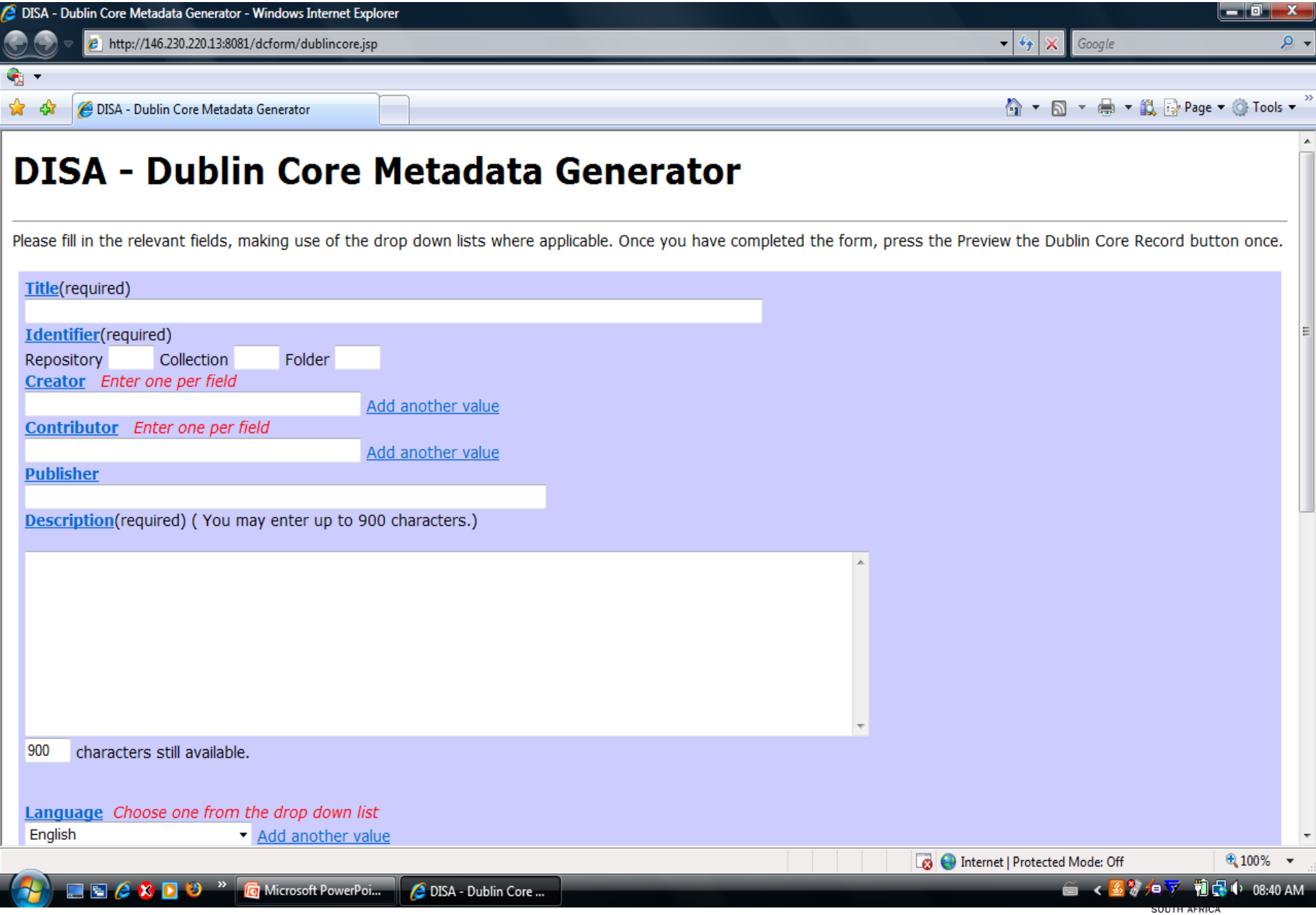
22

DISA
DIGITAL INNOVATION SOUTH AFRICA

Use of keywords/thesaurus/subject headings

- Use of controlled vocabularies
- Consistency
- Efficient retrieval of resources

DISA
DIGITAL INNOVATION SOUTH AFRICA



DISA - Dublin Core Metadata Generator

Please fill in the relevant fields, making use of the drop down lists where applicable. Once you have completed the form, press the Preview the Dublin Core Record button once.

Title(required)

Identifier(required)

Repository Collection Folder

Creator *Enter one per field*

Add another value

Contributor *Enter one per field*

Add another value

Publisher

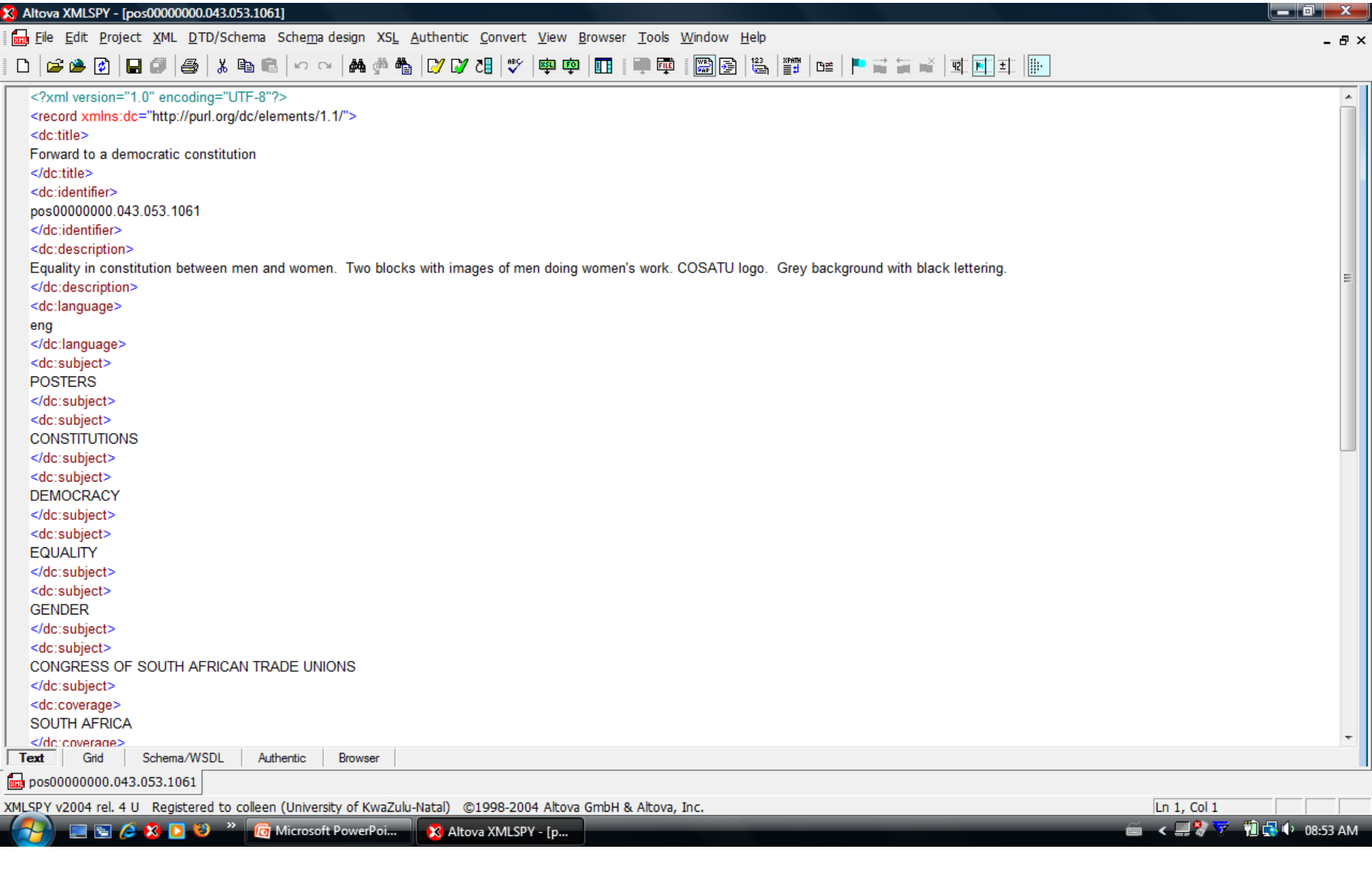
Description(required) (You may enter up to 900 characters.)

900 characters still available.

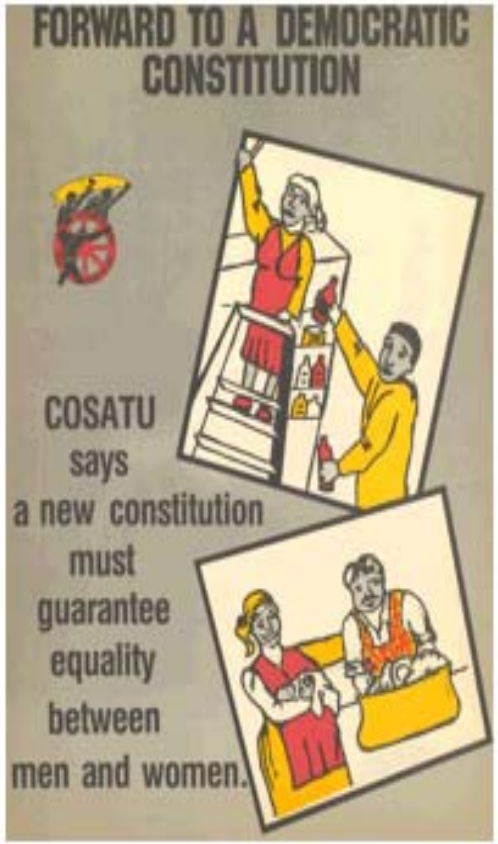
Language *Choose one from the drop down list*

English Add another value

XML metadata record



Web display of metadata



Title	**Forward to a democratic constitution**
Creator	
Contributor	
Publisher	
Date	0000-00-00
Resource type	Posters
Language	English
Keywords	POSTERS CONSTITUTIONS DEMOCRACY EQUALITY GENDER CONGRESS OF SOUTH AFRICAN TRADE UNIONS
Coverage	SOUTH AFRICA
Source	**SAHA**
Extent	1 poster
Description	Equality in constitution between men and women. Two blocks with images of men doing women's work. COSATU logo. Grey background with black lettering.
Relation	
Digital rights	Digital Innovation South Africa

In conclusion…

When creating descriptive metadata ask yourself …

- Is this record useful for resource discovery?
  - Think like a researcher
- Is the content correct?
  - Use reliable reference resources
  - Spelling errors?
  - Controlled vocabularies?
  - Consistency
- Have I adequately described the physical object?

DISA
DIGITAL INNOVATION SOUTH AFRICA

References

- *Digital vs. Traditional Libraries* http://www.wtec.org/loyola/digilibs/02_03.htm: 08/07/2009
- *DISA Guidelines for best practice: indexing and metadata* http://www.disa.ukzn.ac.za/index.php?option=com_docman&task=cat_view&gid=62&Itemid=88 08/07/2009
- *Metadata* http://www.techterms.com/definition/metadata 15/07/2009
- Rowley, Jennifer and Hartley, Richard J. 2008. *Organising knowledge*
- *Playing with building blocks* http://www.edupics.com/en-coloring-pictures-pages-photo-playing-with-building-blocks-i7336.html 08/07/2009

DISA
DIGITAL INNOVATION SOUTH AFRICA