

ECM Governance Policies

Metadata and Information Architecture Policy

Document summary

Effective date	13 June 2012	Last updated	17 November 2011
Policy owner	Library Services, ICTS		
Approved by	Council	Reviewed by	Council
Enquiries	Executive Director: University Libraries - Gwenda Thomas, Executive Director: ICT - Sakkie Janse van Rensburg		

Table of Contents

Purpose	2
Definitions	2
Applicable to	3
Exclusions	3
Policy summary	4
Policy details	4
Policy violations	6
Roles and responsibilities	6
Related links	Error! Bookmark not defined.
<i>Appendices</i>	9
Appendix A – Indicative metadata schema and domain standards	9
<i>Approval and version history</i>	10

Purpose

The University must manage metadata and its application to information assets and services to improve the governance, interoperability, retrievability, re-use, storage optimisation, structure and classification of information assets and services.

Definitions

Term	Definition
Metadata	Descriptive or definitive data that is linked to data or to a data collection.
Schema	A set of rules for the structure and content of an information artefact, by which that artefact may be constituted as valid at the time that it is parsed.
Classification	Indicates the permissible access applicable to a record; the level of sensitivity associated with the information contained in a record. (As distinct from “Categorisation”, classification refers specifically to categorisation of its sensitivity, e.g.: Secret, Confidential, Public...)
XML	Extensible Markup Language: a machine-readable, open standard language used to describe (“mark up”) information content and its component elements (including metadata).
Information Architecture	The discipline and frameworks by which information and data are categorised for sourcing, storage and retrieval. For example: Information architecture applies to storage in the domain of database structures, and to retrieval in navigation, filing structures, catalogues etc.
Content artefact	Any collection of physical or electronic information in any uniform medium, that has meaning and context in its own right, as a single unique entity. A content artefact may be a document, an image, a data collection or a web page, and is normally the result of an operational process. Content artefacts may exist at varying degrees of granularity: from a multimedia collection to a single image.
Repository	A storage container for content artefacts, physical or electronic. Physical repositories are also often known as “depositories”.
Taxonomy	An hierarchical structure used to sort information and content into categories and sub-categories.
Ontology	A schematic representation of “what is known”, defining concepts and their relationship to one another.
Controlled vocabulary	A list of permissible descriptors that may be associated with or assigned to a content object as metadata. In an application user interface, controlled vocabularies are often used to populate drop-down lists or combo boxes.

Term	Definition
Attribute	<p>A defining characteristic (such as a title, date, owner, or subject) that is assigned to an object to assist in its identification.</p> <p>In a computer environment, a specific property inherent in a database entity or an object. Attributes usually consist of a name and a value, and they are often considered important metadata elements. (International Records Management Trust).</p> <p>A specification that defines a property of an object, element, or file. It may also refer to or set the specific value for a given instance of such. (Wikipedia)</p>
Open standard	<p>A generally accepted framework that may be applied across many domains, and is accessible in the public domain for common use and re-use.</p>
(Taxonomy) Facet	<p>A domain-specific taxonomy derived from the values and relationships in a given ontology, providing a domain-relevant point of entry to content. For example, a single ontology of a content collection can deliver taxonomies that are relevant to particular business processes, or taxonomies based on time series, or taxonomies based on information classification, all for the same content collection.</p>

Applicable to

Storage and archiving of physical and electronic content artefacts and content objects - including but not restricted to:

- Scholarly resources
- Business records and procedural documentation
- Website content
- Research data
- Publications
- Personal information
- Media collections
- Databases and content collections

Publication and dissemination of physical and electronic content artefacts and content objects – including but not restricted to:

- Physical publications (books, journals, articles, theses, dissertations, newspapers)
- Websites and web-based portals
- Mobile browsing sites

Exclusions

- Transactional data.

Policy summary

1. Metadata and Information Architecture are jointly managed on behalf of UCT by Library Services and ICTS.
2. All content objects generated, managed and published by the University of Cape Town and its direct affiliates must be tagged and stored with sufficient metadata.
3. Where applicable, metadata should support re-use and interoperability of content between content management systems and content publication media.
4. Reduction of duplication: In the interests of efficiency and alignment, metadata standards, schemas and attribute values should be re-used wherever possible.
5. Metadata should be registered with a central metadata governance body.
6. Standards and accessibility of standards: Open standards for metadata and metadata schemas must be applied wherever possible and practical.
7. Metadata delivery: By default and wherever practical, content metadata should be transferred between information systems in Extensible Markup Language (XML) format.
8. Publication of metadata: Selection of which metadata values are to be published together with a content object is at the discretion of the content channel owner.
9. Metadata quality: Content and collection owners are responsible for ensuring that metadata are compliant with standards, valid, and current.

Policy details

- 1. Metadata and Information Architecture are jointly managed on behalf of UCT by Library Services and ICTS.**
 - 1.1. Metadata and Information Architecture for all of UCT's scholarly resources are managed on behalf of UCT by Library Services.
 - 1.2. All other metadata (such as business, governance and operational process metadata) are managed by ICTS.
 - 1.3. Ownership of and responsibility for metadata as they are applied to information resources, while under the custodianship of Library Services and ICTS, nevertheless resides with the owner of the data or information collection.
 - 1.4. For the purposes of metadata governance and information architecture definition and maintenance, Library Services and ICTS share a metadata registry, which facilitates sharing of resources, metadata, schemas, frameworks, methods, standards and structures wherever applicable.
 - 1.5. Close alignment and collaboration between ICTS and Library Services metadata management and maintenance staff is essential to ensure effective governance.
 - 1.6. The selection of appropriate definitive universal metadata standards and frameworks is at the discretion of Library Services and ICTS for scholarly resources and business information resources respectively, in consultation with each other and with the relevant stakeholders.
- 2. All content objects generated, managed and published by the University of Cape Town and its direct affiliates must be tagged and stored with sufficient metadata.**
 - 2.1. Metadata must provide for unique identification, adequate classification, transparent governance and simple retrieval of content objects.

- 2.2. Metadata capture mechanisms must make every effort to make metadata capture as easy and simple as possible, and metadata fields should be pre-populated with contextually sourced metadata if this exists.
- 2.3. Mandatory requirements for metadata on a content object type must be kept to an absolute minimum, without compromising reasonable requirements for content dissemination.
3. **Where applicable, metadata should support re-use and interoperability of content between content management systems and content publication media.**
 - 3.1. Within content storage and management systems, metadata structures and values should adhere to standards and principles that allow the transfer of content objects to other content management systems.
4. **Reduction of duplication: In the interests of efficiency and alignment, metadata standards, schemas and attribute values should be re-used wherever possible.**
5. **Metadata should be registered with a central metadata governance body.**
 - 5.1. A centralised UCT metadata management body must be established that maintains a record of the metadata standards and architectures in effect in official UCT content storage mechanisms, so that these are available for application to multiple collections across organisational units and information systems.
 - 5.2. It is in the best interests of UCT that owners of new and existing Content Management systems register their metadata standards, metadata architectures controlled vocabularies, and data classification models with the metadata registry.
6. **Standards and accessibility of standards: Open standards for metadata and metadata schemas must be applied wherever possible and practical.**
 - 6.1. While the standard chosen is at the discretion of the content collection owners, the name and scope of the standard employed, and any deviations from the standard, must be registered with the central metadata registry. Where the technology used does not support an open standard, the metadata schemas employed must be registered in the metadata registry.
7. **Metadata delivery: By default and wherever practical, content metadata should be transferred between information systems in Extensible Markup Language (XML) format.**
 - 7.1. If a metadata standard or content system prescribes or requires a format other than XML, the feasibility of transforming data into XML format should be thoroughly investigated.
8. **Publication of metadata: Selection of which metadata values are to be published together with a content object is at the discretion of the content channel owner.**
 - 8.1. Publication of metadata must not expose the University to undue risk or constitute any unwarranted disclosure or information.
 - 8.2. Metadata publication by a channel should be by prior arrangement and agreement with the content owner.
9. **Metadata quality: Content and collection owners are responsible for ensuring that metadata are compliant with standards, valid, and current.**
 - 9.1. Metadata quality applies to the disciplined application of selected metadata standards used as well as to the content of metadata values.
 - 9.2. Audits of metadata quality are the responsibility of the content owner.

- 9.3. Changes, revisions and updates to standards must be kept track of, and implemented where necessary and appropriate.

Policy violations

Non-registration of metadata schemas with the UCT central metadata management body has implications for the security of the University's Intellectual Capital and the re-usability of content, as well as implications for compliance with related policies such as the Records Management Policy and the Web Content Management Policy. Inadequate, invalid and inaccurate metadata may result in inappropriate access to content, poor return on investment, and inefficiency. Violations of this policy will be dealt with by the Executive Director: University Libraries and Executive Director: ICT as applicable.

Roles and responsibilities

Roles

<u>Department/designation</u>	<u>Role</u>
Executive Director: University Libraries	<ul style="list-style-type: none"> Joint Metadata and Information Architecture Policy owner
Executive Director: ICT	<ul style="list-style-type: none"> Joint Metadata and Information Architecture Policy owner
Internal Audit	<ul style="list-style-type: none"> Monitoring of compliance with policy
Information and Communication Technology Services (ICTS)	<ul style="list-style-type: none"> Provision and support of information technology to support the centralised management of metadata (metadata registry) Support for security management of metadata registry Maintenance of content application landscape Maintenance of metadata registry content (schemas, thesauri, controlled vocabularies) for non-scholarly information Oversight of metadata standards for non-scholarly information Advice and consulting on metadata standards and schemas to be used Collaboration and alignment with Library Services metadata management team
Library Services	<ul style="list-style-type: none"> Maintenance of metadata registry content (schemas, thesauri, controlled vocabularies) Oversight of metadata standards Advice and consulting on metadata standards and schemas to be used Collaboration and alignment with ICTS metadata management team
Content owners (of information systems)	<ul style="list-style-type: none"> Ownership of metadata in applicable information systems, and maintenance of metadata standards in these systems

Department/designation	Role
Channel owners (of information channels)	<ul style="list-style-type: none"> • Ownership of metadata requirements for channel publishing
Business analysts / Content analysts	<ul style="list-style-type: none"> • Definition of business requirements and content object structures in the context of the metadata frameworks

Responsibilities

Department/designation	Responsibility
Executive Director: University Libraries	<ul style="list-style-type: none"> • Joint sign-off, executive sponsorship and overall advocacy for Metadata and Information Architecture Policy • Joint maintenance of policy review procedures
Executive Director: ICT	<ul style="list-style-type: none"> • Joint sign-off, executive sponsorship and overall advocacy for Metadata and Information Architecture Policy • Joint maintenance of policy review procedures • Temporary de facto adoption of ED University Libraries responsibilities in respect of metadata policy
Internal Audit	<ul style="list-style-type: none"> • Monitor policy compliance as necessary
Information and Communication Technology Services (ICTS)	<ul style="list-style-type: none"> • Provide and support of information technology to support the centralised management of metadata (metadata registry) • Support security management of metadata registry • Maintain content application landscape • Maintain metadata registry content (non-scholarly resources) • Oversee metadata standards in information applications • Advise and consult on metadata standards and schemas to be used
Metadata management team	<ul style="list-style-type: none"> • Maintain metadata registry content (scholarly resources) • Oversee metadata standards in information applications • Advise and consult on metadata standards and schemas to be used
Content owners (of information systems)	<ul style="list-style-type: none"> • Ensure that metadata standards are applied and monitored in information systems • Approve inter-application metadata transfers • Ensure that metadata quality is maintained in information systems

Department/designation	Responsibility
Channel owners (of information channels)	<ul style="list-style-type: none">• Approve business requirements definitions of metadata sourcing and publishing• Liaise with content owners to negotiate information sourcing requirements and publication criteria• Ensure that metadata standards are applied and monitored in publishing channels
Business analysts / Content analysts	<ul style="list-style-type: none">• Define business requirements and content object structures in the context of the metadata frameworks (including content integration requirements and schema definitions)

Appendices

Appendix A – Indicative metadata schema and domain standards

Schema/ Language	Domain/ Purpose
IEEE Learning Object Metadata	Education
JPEG-2000	Images
Machine-Readable Cataloguing (MARC)	Libraries
Moving Picture Expert Group MPEG-21 and MPEG-7	Multimedia
ONIX	Publishing
Extensible Rights Metadata Language (XrML)	Rights Management
Dublin Core Metadata Element Set	Web

