



# IM Standards to Improve Access to Government Information and Services

Electronic Service Delivery Committee

June 4, 2001

Nancy Brodie



## *Approval sought for*

- GoC Metadata Framework for Information Resources
- Adoption of Dublin Core as GoC core metadata standard for resource discovery
- Adoption of a Controlled Vocabulary Standard and GoC Core Subject Thesaurus



## *What is metadata?*

Simply:

- information about information

or

- structured data about data

Current context

- used to aid the identification, description, location or use of information resources. The use of metadata will be needed through the whole lifecycle of an information resource.
- e.g. library catalogue, Web search result



## *The drivers*

- Clients want more precise search access to GoC information on the Web
- Departments need IM standards and guidelines to support implementation of Common Look and Feel
- Portal, Gateways and Cluster managers need IM standards and guidelines to support managing and sharing information across departments



## ***Standard metadata is the solution***

- Metadata is key to:
  - Identifying and organizing information on the Web
  - Adding precision to searching and browsing
  - Life cycle management of information resources
- Metadata according to a standard is key to:
  - Interoperability and sharing of information resources
  - Consistent and predictable search results
  - A coordinated, cost effective and client-focused effort



## *Standardization activities in GoC*

- GILS Subgroup 1995-1999
  - never adopted as a GoC Standard
- CLF Standard 6.3 evolved from GILS work
  - All GoC web sites must adopt the following five metatags as a meta data standard for description of web resources: Title, Originator, Language of Resource, Date and Controlled Subject.
- Record Keeping Metadata Requirements for GoC 1999-2001
  - RDIMS & IM Forum
  - expanded from Dublin Core and GILS



## *Implementations of Metadata in GoC*

- Most departments awaiting
  - demonstrated use of metadata in GoC search engines
  - clearer guidance from TBS
  - adoption of a standard by GoC
- Early implementors
  - Environment Canada - implemented GILS in 1995, converted to Dublin Core in 2000
  - NRCan - geospatial metadata standards
  - Cultural agencies - Dublin Core for digital content
  - SchoolNet - Dublin Core for educational resources



## ***The need to select a metadata standard***

- GOL Metadata Working Group began discussions on a standard in February 2001
- Representatives from 17 departments
- Established scope, benefits, selection criteria and implementation requirements
- Assessed GILS & Dublin Core
  - DC stronger on uptake, government adoption, governance, simplicity, multilingual support
- Selected Dublin Core

***has become urgent***





## *Why Dublin Core*

- Both simple and extensible
- Supports semantic interoperability across different communities
  - e.g. government, cultural, education, health
- International consensus
  - IETF RFC 2413, NISO Z39.85-200x balloting
- Multilingual - User Guide translated by U. Laval
- Growing body of models, implementation guidance and software tools
- Open mechanism for on-going evolution



## *Government Adoption of Dublin Core*

- Australia
  - AGLS 1997 - full scale implementation in support of Government Online
- UK
  - e-Government Metadata Framework May 2001
- Ireland
  - Irish Metadata Consultation Document April 2001
- New Zealand - adopting AGLS
- many US states
- DC Government Working Group and discussion list

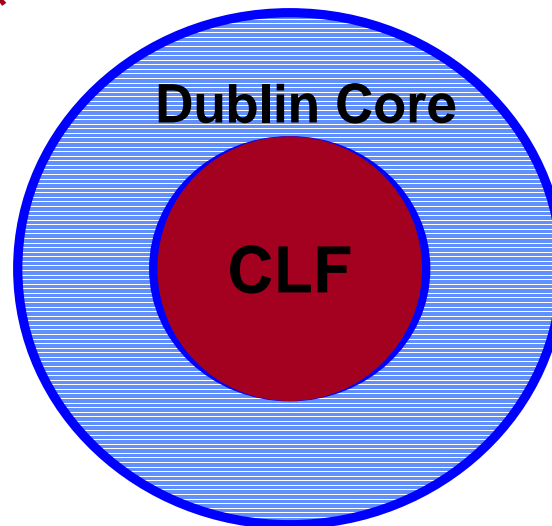


## ***Dublin Core - what is it***

- The Dublin Core is a simple set of 15 metadata elements intended to facilitate *discovery* of electronic resources.
- Dublin Core standardizes semantics i.e. what pieces of information are needed and their labels
- Dublin Core provides guidance on syntax i.e. how the information is coded for computer processing
- Dublin Core provides basic content standards i.e. how the information is formed

# Dublin Core (DC) Metadata Element Set

- DC.Title \*
- DC.Creator \*
- DC.Language \*
- DC.Date \*
- DC.Subject \*
- DC.Description
- DC.Publisher



**\*CLF (5 elements)**

- DC.Type
- DC.Format
- DC.Identifier
- DC.Source
- DC.Relation
- DC.Coverage
- DC.Rights
- DC.Contributor

**15 elements**



## ***Common Look and Feel Standard 6.3***

- Title (dc.title)
  - convey the content, may be presented independently
- Originator (dc.creator)
  - GoC department or agency responsible for content
- Language (dc.language)
  - language code for content
- Date (dc.date.created & dc.date.modified)
- Controlled subject (dc.subject scheme= )
  - terms selected from a registered thesaurus

***the mandatory core***



## *What is a thesaurus?*

- A hierarchical list of controlled terms that is used both by indexers to describe information resources and by users to gain access to them.
- Provides synonyms and related terms to help the user identify the appropriate term within their scope of interest.
- Standardizes the external form and the meaning of index terms, thus ensuring that a particular concept or subject will always be represented in the same way in the index.



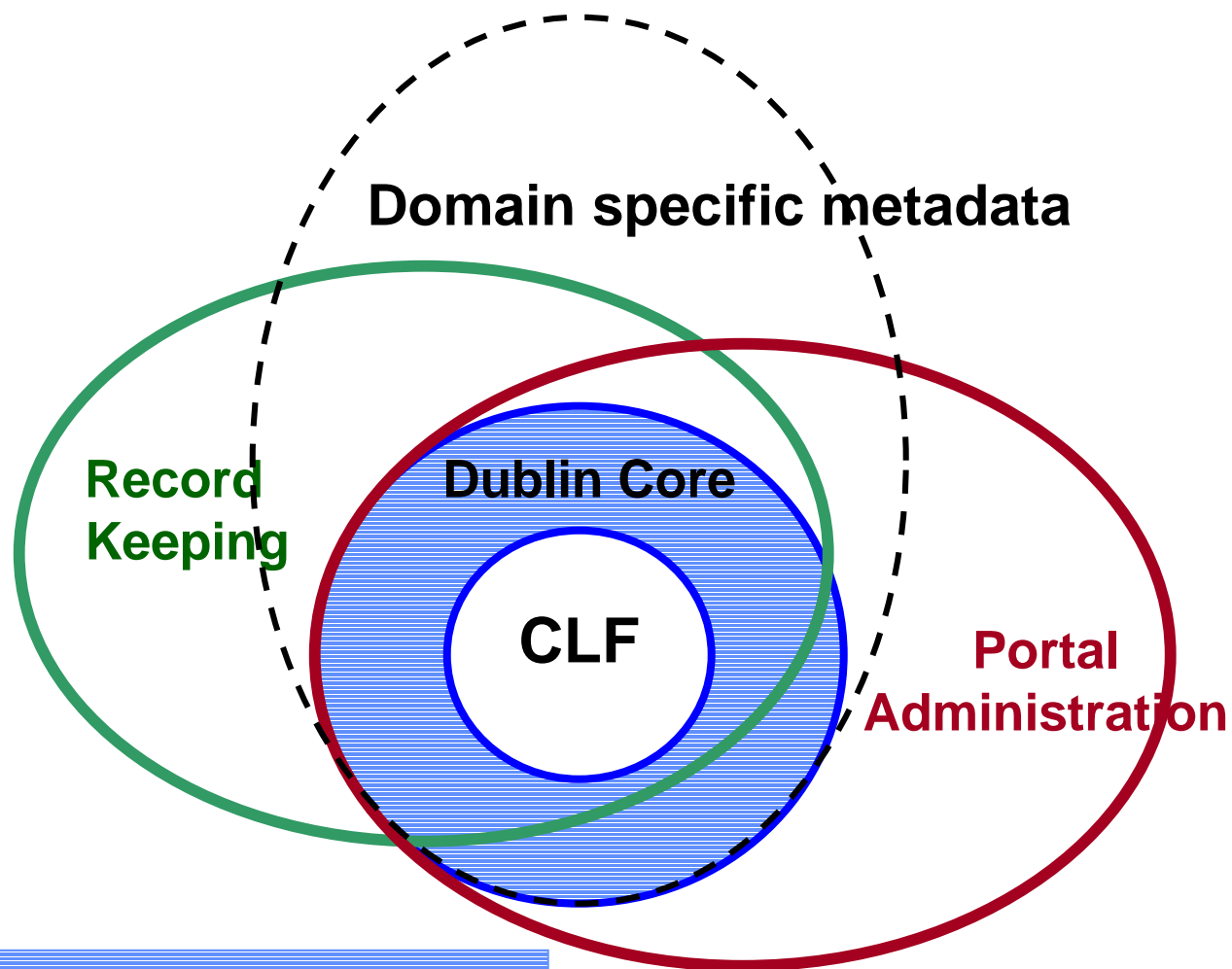
## ***GoC Core Subject Thesaurus***

***<http://dsp-psd.pwgsc.gc.ca/Thesaurus/index-e.html>***

- Recommended as the default thesaurus for use with CLF standard element: **Controlled Subject**
- Meets ISO standards
  - Guidelines for the establishment and development of monolingual thesauri (ISO 2788-1986), and the Guidelines for the establishment and development of multilingual thesauri (ISO 5964-1985).
- Represents all fields of knowledge found in GoC information resources & services
  - Because of this exceptionally wide coverage, the subject descriptors are fairly general in nature.



# GoC Metadata Framework



International Standard  
Recommended by Metadata Working Group





## ***GoC Metadata Framework***

- Only the five elements previously approved as Common Look and Feel Standard 6.3 are mandatory across all GoC Web sites
- The Dublin Core is used when more than 5 elements are needed. No additional elements are mandatory.
- Domain specific metadata sets extending the Dublin Core may be developed and will specify mandatory and optional elements for the domain.



## *Where to apply metadata*

- Welcome pages and Home pages
- Topics/services in high demand
- Information required to understand entitlements
- Pages that provide an E-Service
- Entry points to specific online services and indexes (databases)
- Major formal publications & news releases
- Major collections of links
- manuals used in decision-making affecting the public

## *Metadata for Portal Administration*

- A metadata standard for Portal Administration extending the Dublin Core will be developed and will identify mandatory and optional elements for participation in clusters.

### Early assessment

- Common Look and Feel core
- Elements for display (from DC) e.g. abstract (dc.description)
- Elements to support enhanced search (from DC with GoC scheme and content rules) e.g. industrial code
- URL (dc.identifier)
- Elements for administration e.g. contact name



## *Impact on departments*

- No mandatory requirements beyond Common Look and Feel metadata standard
- Enhanced guidance, vocabularies and tools will facilitate CLF implementation
- Use of metadata by search engines and in browsing will provide benefits to justify investment
- Use of standard metadata to submit information to clusters will minimize resource implications of reporting to multiple clusters



## *Implementation requirements*

- Governance and leadership
  - ← TBS as champion and maintenance body
- Content rules
  - ← Established for mandatory elements
    - Will be needed for extensions, portal admin
- Common vocabulary
  - ← GoC Core Subject Thesaurus as default for content of Subject element
    - Will be needed for geographic, audience and other attributes
- Standard mappings and crosswalks
  - to Dublin Core from domain specific metadata sets



## *Implementation requirements*

- Where to apply metadata (level of granularity)
  - ← Guidance developed based on Australian experience
- Tools for metadata creation
  - ← Commercial and public domain tools exist
    - Upgrade GoC tools to Dublin Core - Canada Site tool and departmental tools
    - Portal Content Admin Pilot outcomes



# *Implementation requirements*

- Training
  - ← Enhanced Common Look and Feel guidance
  - ← GovNet01 Metadata Workshop presentations available on-line (in translation)
  - ← On-line guidance on how to use GoC Core Subject Thesaurus
    - Develop train the trainers strategy (National Library and Council of Federal Libraries)
- Search engine use of metadata
  - Canada Site search engine must be “metadata enabled”
  - Support browsing via metadata and thesaurus



## ***Consultations undertaken to date***

- GOL Metadata Working Group - 20+ departments
- IM Forum
- Council of Federal Libraries
- IM Advisory Group
- GovNet01 Workshop
- Departments - e.g. Transport, Health
- Domains - cluster managers, digital content, instructional materials





## *Approval sought for*

- GoC Metadata Framework for Information Resources
- Adoption of Dublin Core as GoC core metadata standard for resource discovery
- Adoption of a Controlled Vocabulary Standard and GoC Core Subject Thesaurus

Comments to Nancy Brodie by June 18, 2001

[brodie.nancy@tbs-sct.gc.ca](mailto:brodie.nancy@tbs-sct.gc.ca)