



Treasury Board of Canada  
Secretariat

Secrétariat du Conseil du Trésor  
du Canada

# Government of Canada Metadata Framework

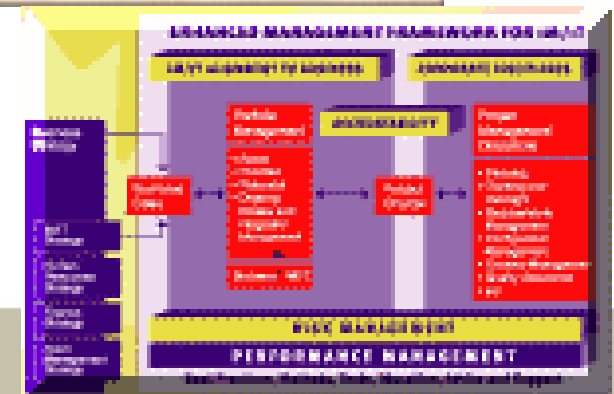
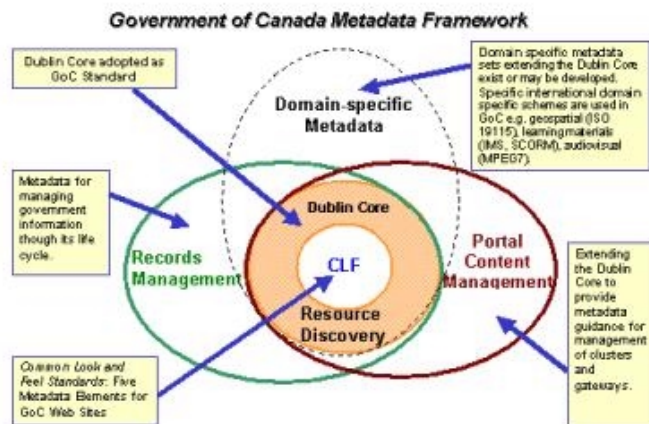
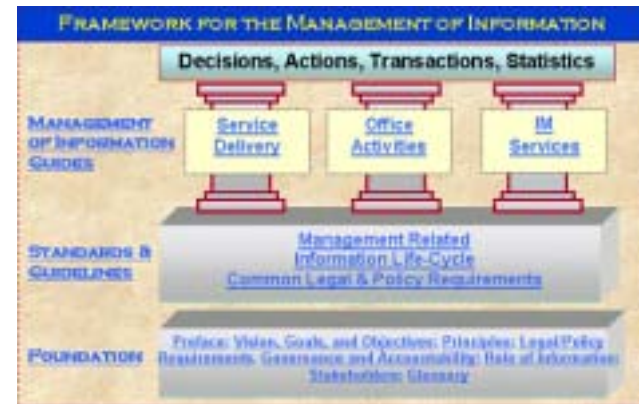
Presentation to  
Canadian Metadata Forum  
September 20, 2003

Nancy Brodie  
Information Management  
Chief Information Officer Branch  
Treasury Board of Canada, Secretariat

Canada 

# Frameworks in the Government of Canada

- ◆ Management Accountability Framework
- ◆ Framework for the Management of Information
- ◆ Enhanced Management Framework





## ***Evolution of GoC Metadata Framework***

1970s

- ◆ Library cataloguing – MARC Format

1980s

- ◆ Rules for Archival Description (RAD)

1990s

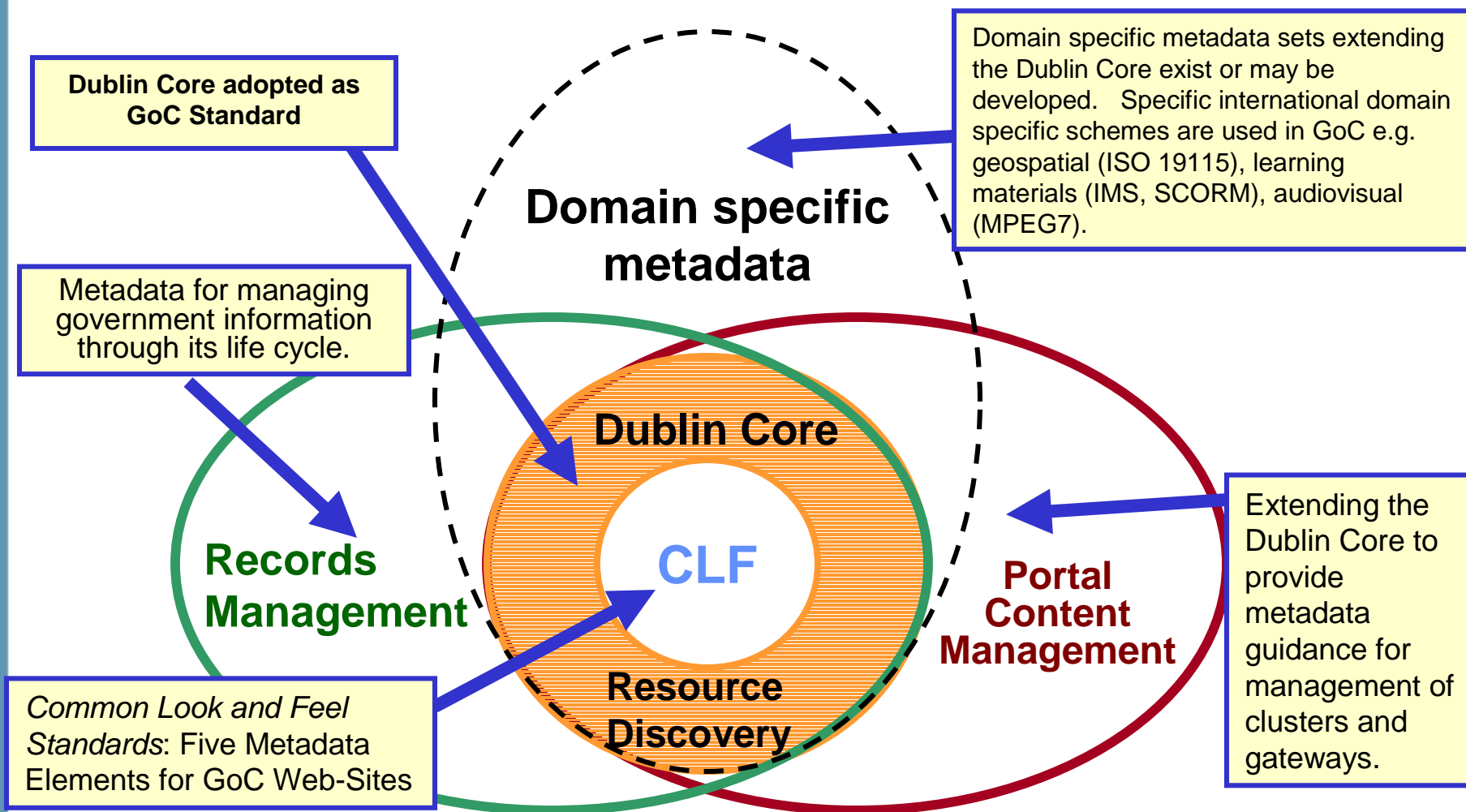
- ◆ FGDC CSDGM (Content Standard for Digital Geospatial Metadata)
- ◆ Government Information Locator Service (GILS)

Diverse needs, many data elements, little commonality

***Complexities***



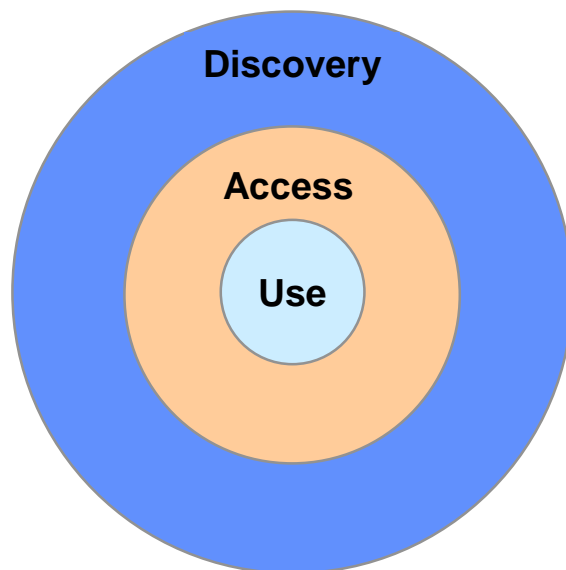
# Government of Canada Metadata Framework



# *Environment Canada's Metadata Framework*

## *Three Levels of Metadata*

***A flexible strategy  
that matches effort to need***



**Using internationally  
recognized standards**

### **Discovery**

Most EC information assets will be discovered at this simplest level. This could be for a collection, database or single object.

### **Access**

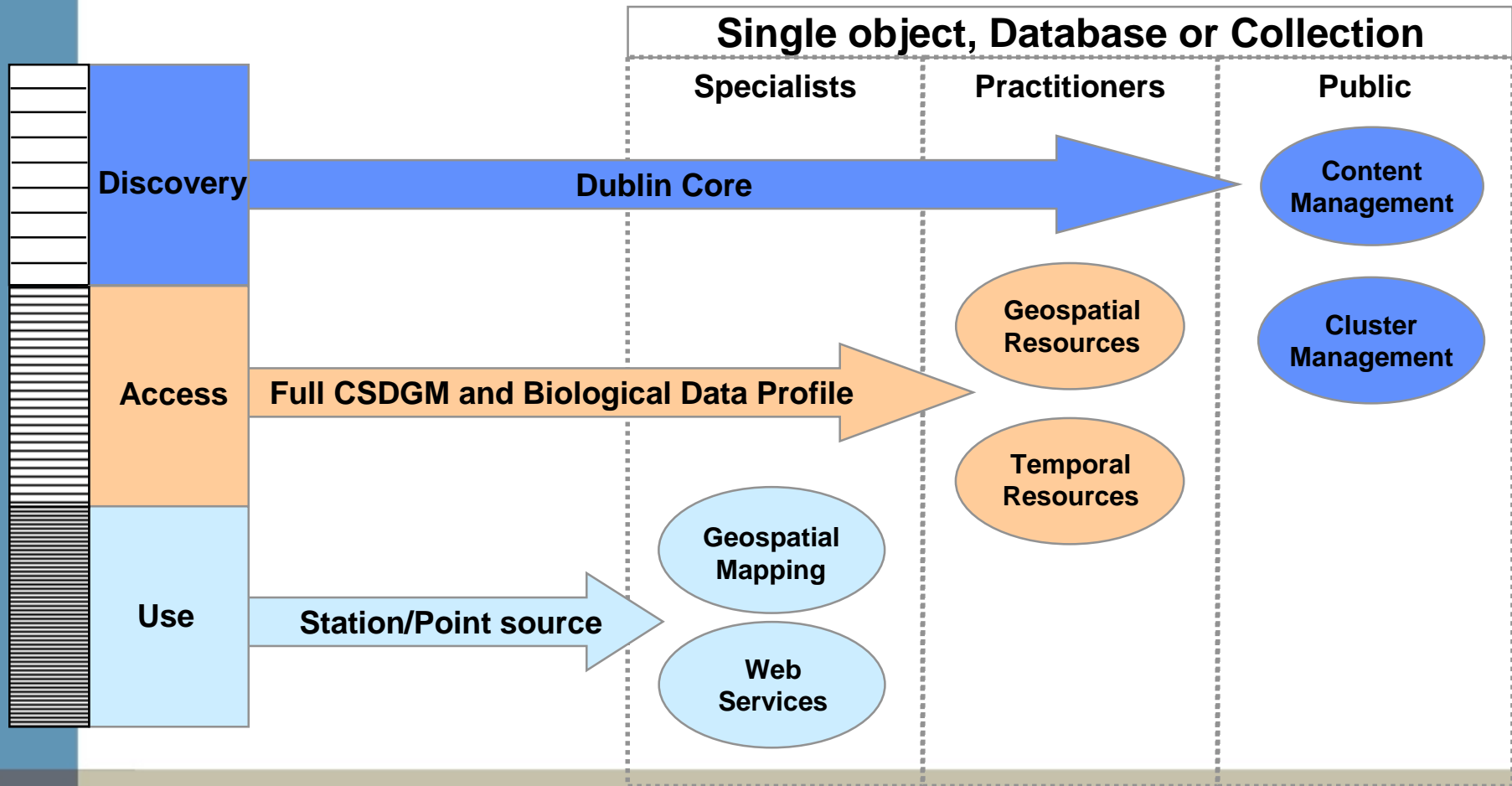
Using the full geospatial and/or biological profile this level will provide for the comprehensive description and disclosure of data.

### **Use**

This level will allow for the use of biological or geospatial metadata at the station level for visualization and data extraction web services.



# Hierarchical Application of Metadata Standards at Environment Canada



# Dublin Core: Metadata Standard for Resource Discovery

**Title** \* (dc.title)

**Creator** \* (dc.creator)

**Language** \* (dc.language)

**Date** \* (dc.date)

**Subject** \* (dc.subject)

**Description** (dc.description)

**Publisher** (dc.publisher)

**Contributor** (dc.contributor)

**\*CLF (5 elements)**

**Type** (dc.type)

**Format** (dc.format)

**Identifier** (dc.identifier)

**Source** (dc.source)

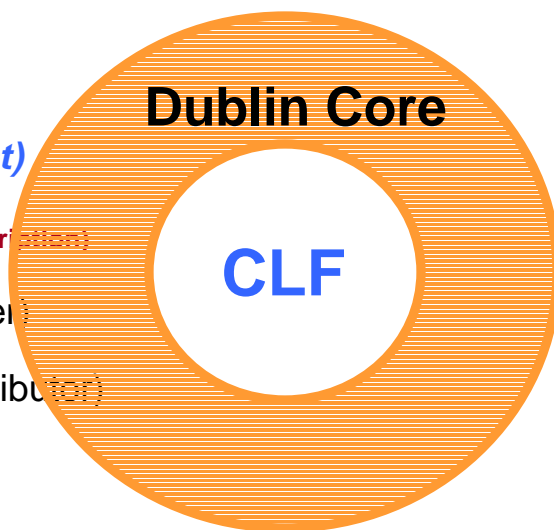
**Relation** (dc.relation)

**Coverage** (dc.coverage)

**Rights** (dc.rights)

**Audience** (dc.audience)

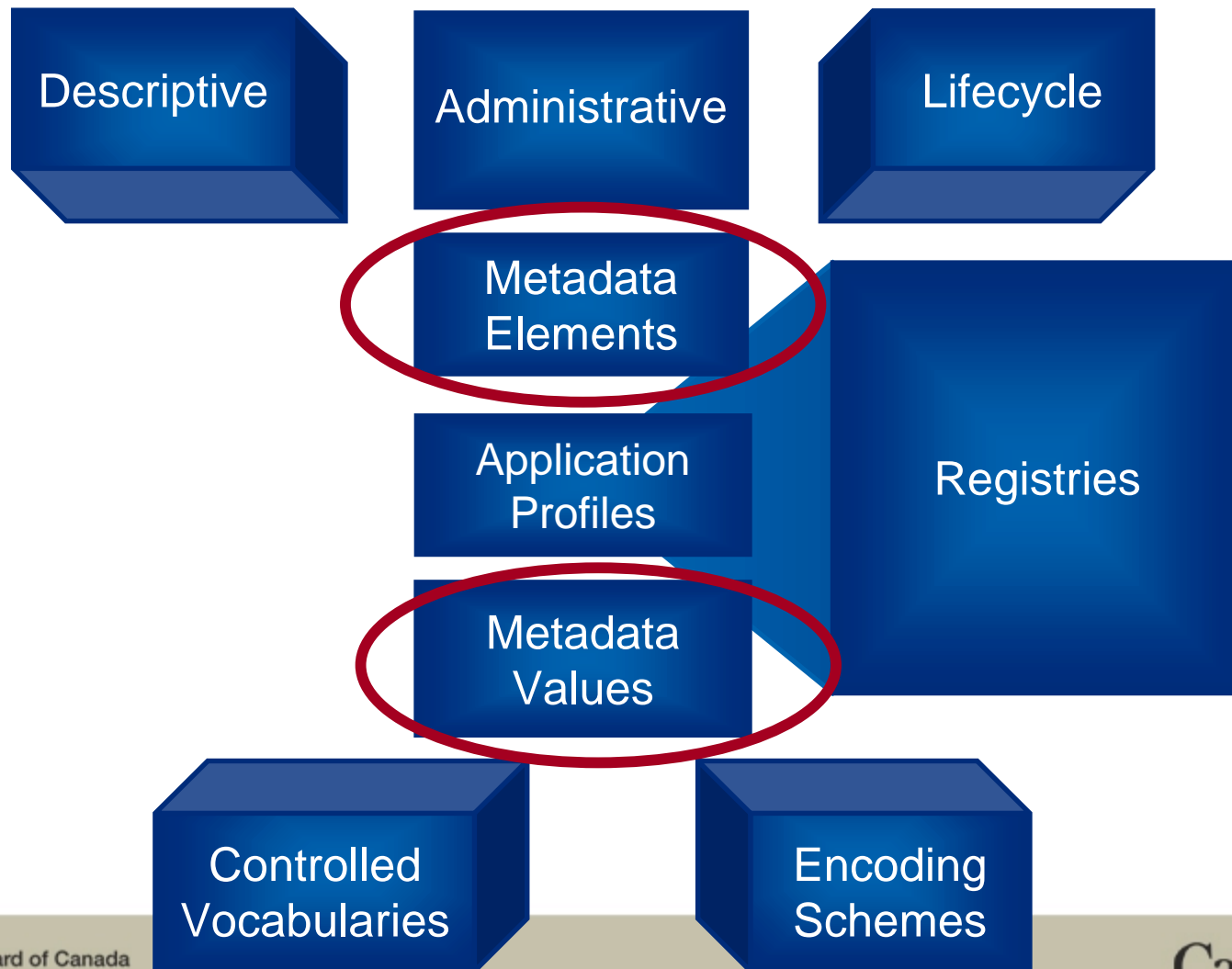
**16 elements**



- Only the five elements previously approved as Common Look and Feel (CLF) Standard 6.3 are mandatory across all GoC Web sites
- The Dublin Core is used when more than 5 elements are needed.
- **Guidance and schemes have been developed.**

# *Information Assets Management Framework from Health Canada - extract*

## Metadata Standards





# *Semantic model*

## ◆ **Metadata Elements**

- Dublin Core defines the semantics (meaning) of the 16 elements

## ◆ **Metadata Values**

- ◆ A controlled element : the content or value is
  - formatted in a standard way (encoding scheme), or
  - a term from a specified list of acceptable values or controlled vocabulary
- ◆ Controlled vocabularies should be named, registered and identified in metadata records as schemes
- ◆ An element can have multiple schemes

*Common semantics are key to interoperability*





## ***Dublin Core schemes for interoperability***

- ◆ Dublin Core Metadata Element Set
- ◆ Dublin Core Qualifiers
- ◆ DC Element Encoding Schemes
  - Rules and controlled vocabularies
  - DCMI Type Vocabulary
  - ISO 639-2: Codes for the representation of names of languages
- ISO 3166: Codes for the representation of names of countries
- Getty Thesaurus of Geographic Names
- Internet media type (MIME)
- ◆ Subject Schemes
  - LCSH, MeSH, DDC, LCC, UDC

<http://dublincore.org/usage/terms/dc/current-schemes/>

***Build on what is widely accepted and used***





## ***Schemes for interoperability within GoC***

### ***Accommodate diversity***

- ◆ Define usage of DC schemes
- ◆ Broad, high-level GoC-wide schemes
- ◆ More specific domain or departmental schemes

### ***Issues***

- ◆ Defining broad and high-level
- ◆ Managing scope of schemes
- ◆ Mandatory use of GoC scheme

***Wide-spread use of common schemes is key to interoperability***





## ***Principles for registering GOC Schemes***

- ◆ Controlled vocabularies used by the Government of Canada should be registered and publicly available (TBITS 39.2)
- ◆ The National Library of Canada (NLC) is the registrar of standardized vocabularies used in GoC
- ◆ Vocabularies must be **bilingual**
- ◆ Vocabularies must be created and maintained by **trusted authorities** which have a mandate within a department to develop and maintain the vocabulary
- ◆ Departmental thesauri must be capable of being mapped to the GoC ***Core Subject Thesaurus***





## ***Evolving principles for developing GoC schemes***

- **Applicable:** terms represent content found on a significant number of Government of Canada Web sites, and/or are of substantial significance to Government of Canada programs/services
- **Recognizable:** terms are understandable by implementers/indexers
- **Unique:** no terms will be a synonym of an existing term [within one scheme]
- **Client-centric:** concepts and terminology used in presentation layer should be tested with the public





# ***Established GoC Schemes***

## **Mandatory elements**

### **dc.creator**

- ◆ Titles of Federal Organizations, GEDS

### **dc.subject**

- ◆ GoC Core Subject Thesaurus

## **Optional elements**

### **dc.coverage**

- ◆ Canadian Geographic Names Data Base
- ◆ Regions of Canada

### **dc.type**

- ◆ Government of Canada Type Scheme

### **dc.audience**

- ◆ Government of Canada Audience Scheme

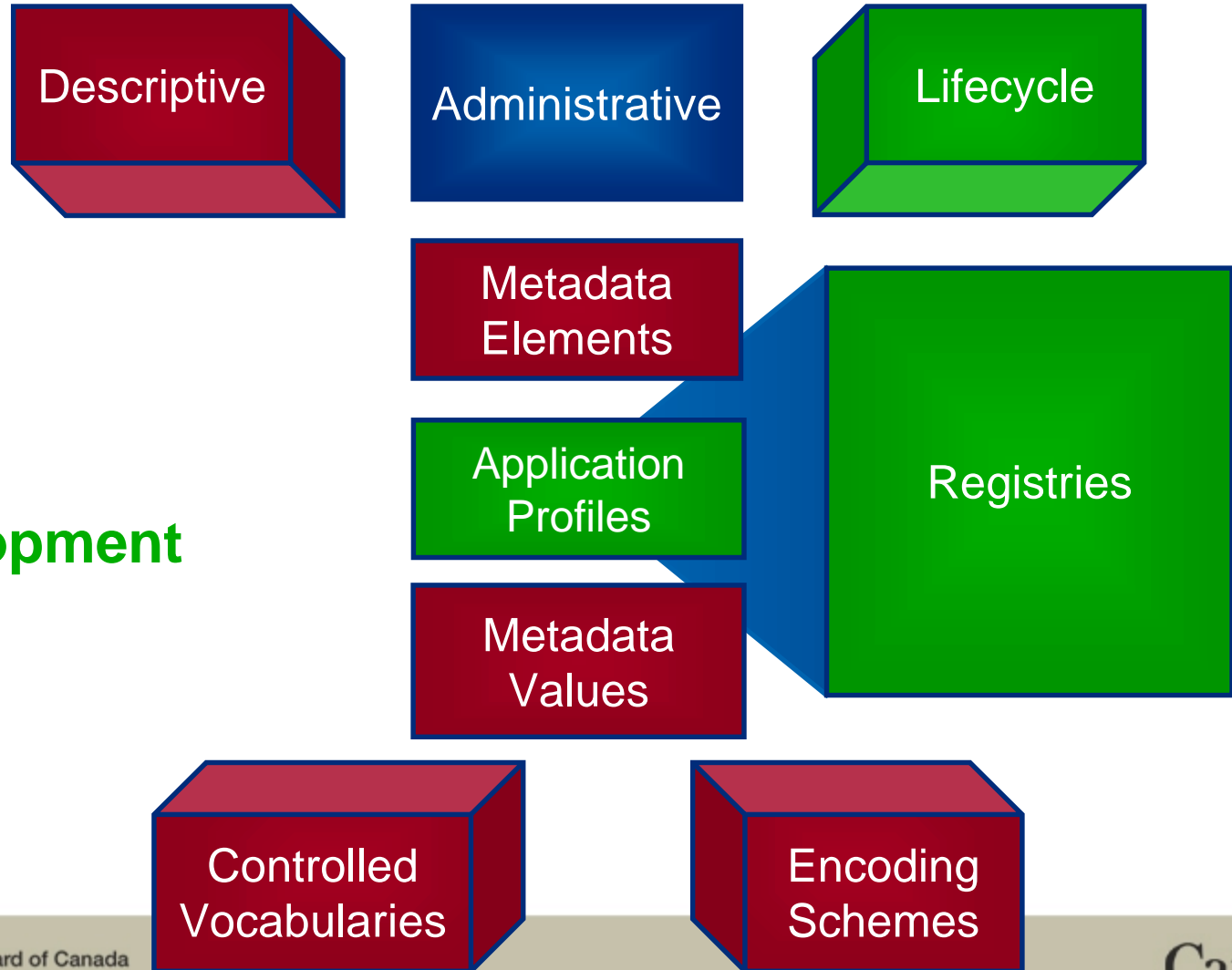
[http://www.cio-dpi.gc.ca/im-gi/meta/meta-cdn\\_e.asp](http://www.cio-dpi.gc.ca/im-gi/meta/meta-cdn_e.asp)



# Where we are today...

Existing

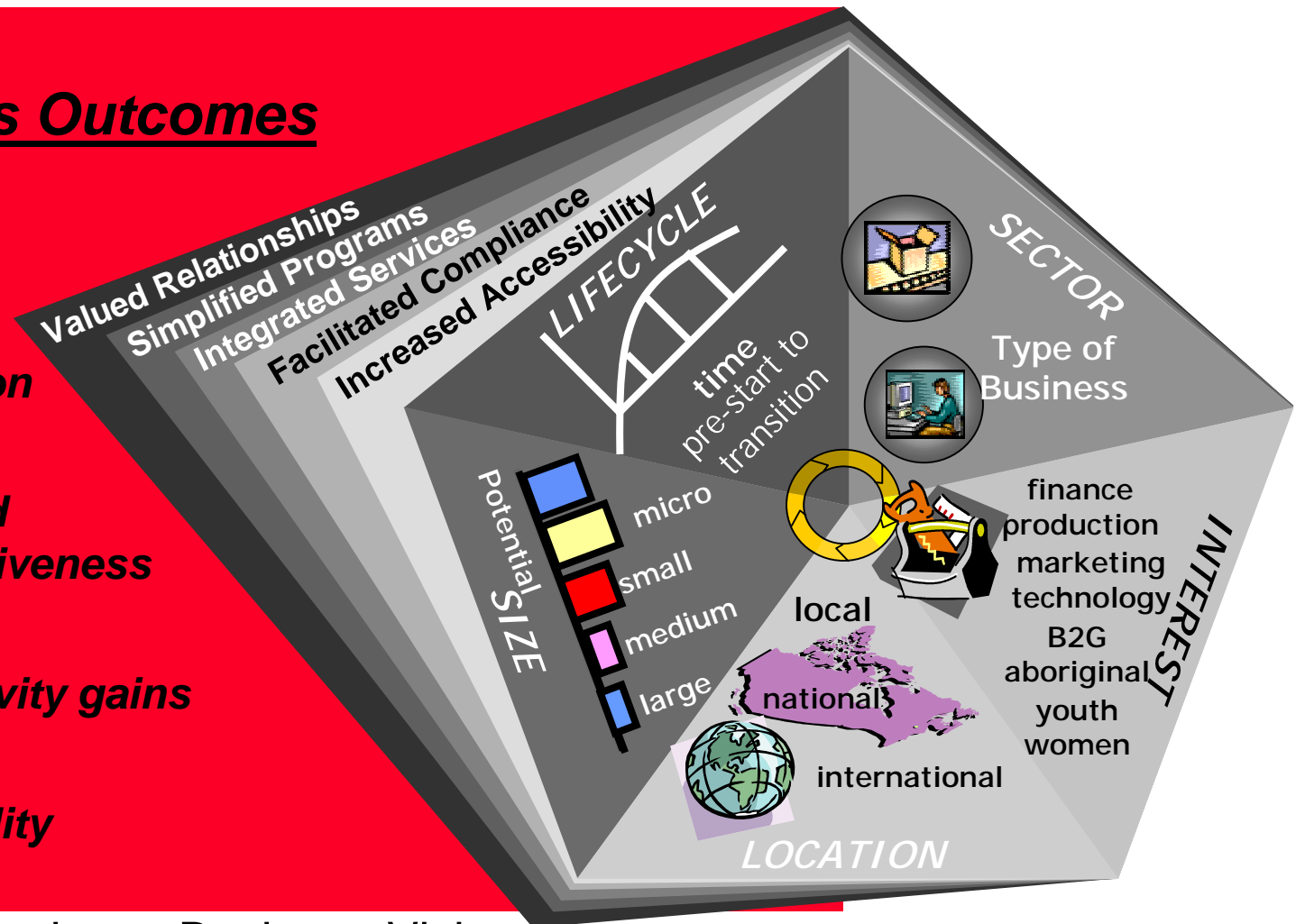
In development



# Facets to support access

## Business Outcomes

- **Growth**
- **Innovation**
- **Improved competitiveness**
- **Productivity gains**
- **Profitability**



Source: Service to Business Vision





# The Metadata Environment

