



Introduction to Metadata

Metadata and the Government On-Line
Metadata Standard

Session Goals

- Learn about. . .
 - The nature and purpose of metadata
 - Dublin Core metadata standard
 - Government On-line guidelines and requirements for metadata implementation
 - How to apply metadata for GOL mandatory elements
 - Resources and strategies for problem solving and ongoing learning

Metadata

Commonly Held Views

- More geek speak!
- Here comes something else THEY want us to do!
- Ugh!





Metadata

The Official View

Metadata is . . .

data about data

Metadata

- Data about data!
- Right . . . geek speak!



Metadata

1. Describes the who, what, when, where, why, and how about a data set or resource such as:
 - ◆ a Web page
 - ◆ a computer file
 - ◆ an image
 - ◆ an object
2. Metadata is structured

Scenarios

- Imagine trying to find a book in a library without a catalogue
- Imagine a universe of 300 TV channels without a TV guide
- Imagine visiting an online bookstore and having an invisible presence suggest several titles perfectly matched to your reading tastes

Where is metadata found?

- HTML document
 - ◆ between the `<HEAD>` and `</HEAD>` tags
- Separate file
- Database

Government Web Page

ALLERGY ALERT

UNDECLARED SULPHITES IN AL-DURRA/ALDERRA BRAND WHOLE APRICOTS IN SYRUP PRODUCT

OTTAWA, February 10, 2003 - The Canadian Food Inspection Agency (CFIA) is warning consumers with sensitivities to sulphites not to consume AL-Durra/Alderra brand whole apricots in syrup product. This product may contain sulphites which are not declared on the label. This alert is of concern to those individuals who have sensitivities to sulphites.

The affected product, AL-Durra/Alderra brand whole apricots in syrup product, manufactured by Al-Durra for Food Production Co., Syria, is sold in various sized glass jars. Both brand names, AL-Durra and Alderra appear on the label. All lots and codes of this product are affected by this recall.

The importer is voluntarily recalling the affected product from the marketplace. This product is known to have been distributed in British Columbia, but may have been distributed nationally.

The CFIA is continuing to investigate other importers of the affected product.

Consumption of this product may cause a serious or life-threatening reaction in persons with sensitivities to sulphites. There have been no reported illnesses associated with the consumption of this product.

The CFIA is monitoring the effectiveness of the recall.

For more information, consumers and industry can call the CFIA at one of the following numbers:

In the province of Quebec 1-800-561-3350; or

In other provinces and territories 1-800-442-2342. - 8:00 am to 4:00 pm local time - Monday to Friday.

For information on receiving recalls by electronic mail, or for other food safety facts, visit our web site at www.inspection.gc.ca.

HTML Metadata Example

```
<meta name="dc.title" content=
"Allergy Alert - Undeclared Sulphites in Al-Durra/Alderra Brand Whole Apricots in Syrup Product
">
<meta name="dc.creator" content=
"Government of Canada, Canadian Food Inspection Agency, Operations, Office of Food Safety
and Recall">
<meta name="dc.subject" scheme="gcore" content=
"Food safety; Labelling; Food labelling; Consumers; Consumer protection">
<meta name="dc.date.created" content="2002-01-31">
<meta name="dc.date.modified" content="2003-02-10">
<meta name="dc.language" scheme="ISO639-2" content="eng">
<meta name="dc.description" content=
"The Canadian Food Inspection Agency (CFIA) is warning consumers with sensitivities to
sulphites not to consume AL-Durra/Alderra brand whole apricots in syrup product. This
product may contain sulphites which are not declared on the label.">
<meta name="keywords" content=
"food recall, food recalls, allergy alert, allergy alerts, public warning, food safety, food borne
illness, food poisoning, food allergy, food allergies">
```

Example Tagging Tool

META Tag Generator

Title: [Help](#)

Originator: [Help](#)

[Titles of federal organizations](#)

Controlled Subject: [Help](#)

[GoC Core Subject Thesaurus](#)

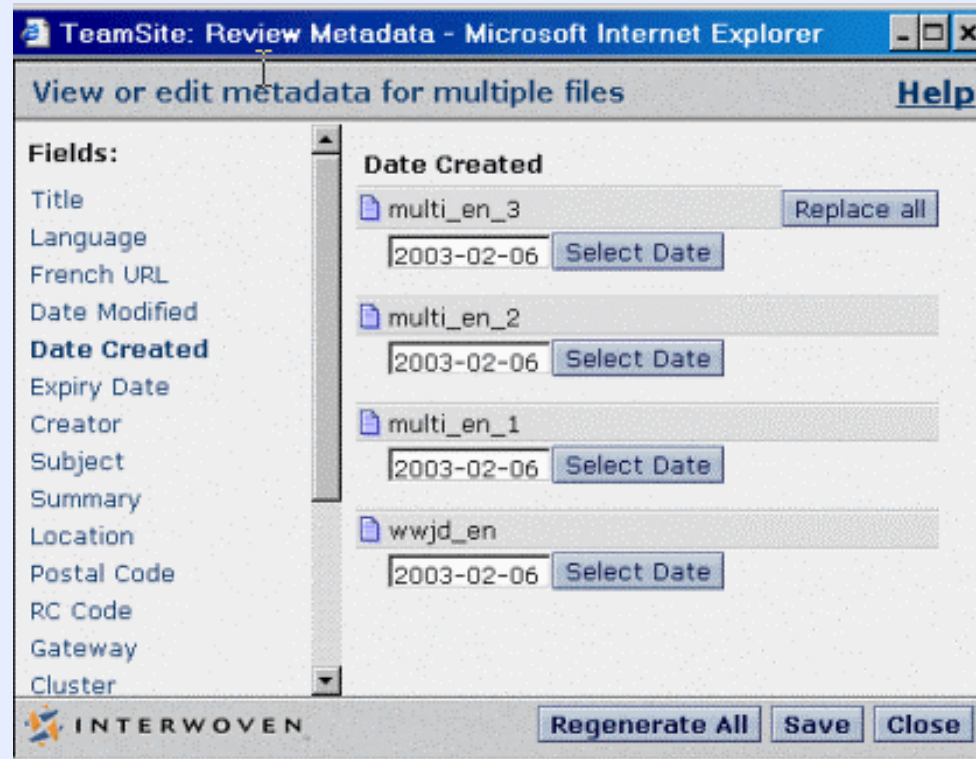
Date created: Date modified:

Review date: [Help](#)

Language of resource: [Help](#)

[The ISO639-2 guide](#)

Example Tagging Tool



Why Use Metadata

- Improved Searching
- Improved Content Management
- Improved Content Sharing

Improved Searching

- Better search results through . . .
 - fielded searching
 - relevance ranking
 - results lists grouped by subject
- Metadata will provide better access to your work

Metadata Search

HotMeta - Metadata Search Engine



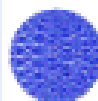
Queensland Government Search Engine

Find:

Match words exactly (no stemming)

Ret.

[Help](#) [Advanced Search](#)

 DISTRIBUTED SYSTEMS TECHNOLOGY CENTRE

HotMeta v1.2
Copyright 1998 DSTC Pty Ltd.
All Rights Reserved.

- All elements
- Title
- Creator
- Subject
- Description
- Publisher
- Contributor
- Date
- Type
- Format
- Identifier

Improved Searching

- Search results after metadata:
 - ◆ 20% better on average
 - ◆ 40% better on average for “novice” searchers
 - ◆ Searches completed more quickly

→ Source: U.K. Office of the e-Envoy; *Government Metadata Standard Testing Project.*

Content Sharing through Metadata Harvesting

- Support content sharing
- Collects metadata from multiple providers
- Metadata is “harvested” and indexed centrally
- Single point of query for user

GOL Metadata

- Canada Site
- HRDC -- Automatic Directory
- Canada News Site -- customized news alerts
- Cluster/gateway application --
 - ◆ Canada Business Centre Gateway

Metadata Standards

- Establish uniform “ground rules”
- Ensure consistency
- Support interoperability and sharing
- Pave the way for new technologies

Definitions

- ◆ **Standard** -- A set of rules or a protocol that define how something will be done
- ◆ **Element** -- A discrete unit of data or metadata
- ◆ **Application Profile** -- set of metadata elements, policies, and guidelines defined for a particular application or domain.
- ◆ **Best Practices** -- Guidelines and documentation to describe and standardize the use of metadata elements

Metadata Standards and Applications -- Examples

- GILS -- Global Information Locator Service
- AGLS -- Australian Government Locator Service
- Cancore -- Canadian e-learning application profile
- Dublin Core -- International Standard adopted by GoC

Metadata Standards

- Dublin Core (<http://dublincore.org>)
- International metadata standard for resource discovery
- Features
 - ◆ Simplicity
 - ◆ 16 Elements
 - ◆ Extensibility
 - Element refinements
 - Encoding schemes

Simple Dublin Core

Title

Description

Date

Identifier

Source

Audience

Creator

Publisher

Type

Language

Coverage

Subject

Contributor

Format

Relation

Rights

Metatags

- **2 parts for each element (html format)**
 - ◆ meta name=" "
 - ◆ content=" "
- **Dublin Core metatag examples**
 - ◆ `<meta name="dc.title" content="Sustaining the Environment and Resources for Canadians - What is it?" >`
 - ◆ `<meta name="dc.creator" content="Government of Canada, Environment Canada, Deputy Minister's Office" >`

DC Extensibility

- Element refinements

- <meta name="dc.date.created" content="2002-03-27" >

- <meta name="dc.date.modified" content="2002-10-28" >

- Encoding schemes

- <meta name="dc.subject" scheme="gcore" content="Environment; Environmental management">

- <meta name="dc.language" scheme="ISO639-2" content="eng">

TBITS 39.1

- TBITS 39.1 = GOL Metadata Standard
- Goals
 - ◆ Coordination across domains, clusters and business functions
 - ◆ consistency
 - ◆ interoperability
 - ◆ cost-effectiveness

TBITS 39.1

- Adopts the Dublin Core Standard as core standard for resource discovery
- Specifies how the Dublin Core element set is to be implemented for Government On-Line

TBITS 39.1

- Five mandatory DC elements (as per CL&F 6.3)
 - ◆ dc.creator
 - ◆ dc.title
 - ◆ dc.language
 - ◆ dc.date
 - ◆ dc.subject

TBITS 39.1

- Applies to Web resources

- ◆ Welcome pages and Home pages
- ◆ Topics and services in high demand by the community
- ◆ Information required by the public
- ◆ Pages that provide an actual on-line service to the public
- ◆ Pages required to meet a prescribed legal or service obligation
- ◆ Entry points to specific online services and indexes

TBITS 39.1

- Applies to Web resources, cont.
 - ◆ Major formal publications
 - ◆ Media releases
 - ◆ Major entry points or indexes and menus
 - ◆ Information about agency powers affecting the public
 - ◆ Manuals and other documents used in decision-making affecting the public
 - ◆ Substantial descriptive or marketing information about institutions, their services, activities and collections

General Guidelines

- Elements are repeatable where appropriate
- Resources must be described in the language of the resource
 - ◆ Bilingual resources are described in both official languages
- DC syntax is not translated
- Use lower case for DC tags
 - dc.title
 - Not: DC.TITLE or Dc.title or Dc.Title



General Guidelines

- English example:

```
<meta name="dc.title" content="What is Micro-Credit?">
```

```
<meta name="dc.creator" content="Government of Canada, Industry Canada">
```

```
<meta name="dc.language" scheme="ISO639-2" content="eng">
```

```
<meta name="dc.date.created" content="2002-03-20">
```

```
<meta name="dc.date.modified" content="2002-06-11">
```

```
<meta name="dc.subject" scheme="gccore" content="Finance; Credit">
```

- French example:

```
<meta name="dc.title" content="Qu'est-ce que le micro-crédit?">
```

```
<meta name="dc.creator" content="Gouvernement du Canada, Industrie Canada">
```

```
<meta name="dc.language" scheme="ISO639-2" content="fre">
```

```
<meta name="dc.date.created" content="2002-03-20">
```

```
<meta name="dc.date.modified" content="2002-06-11">
```

```
<meta name="dc.subject" scheme="gccore" content="Finances; Crédit">
```


Mandatory Elements

- `<meta name="dc.creator" content=" " ">`
- `<meta name="dc.title" content=" " ">`
- `<meta name="dc.date.created" content=" " ">`
- `<meta name="dc.date.modified" content=" " ">`
- `<meta name="dc.language" scheme="ISO639-2" content=" " ">`
- `<meta name="dc.subject" scheme="gccore" content=" " ">`

Mandatory Elements -- dc.creator

- dc.creator
 - ◆ Indicates who is primarily responsible for creating the content of the resource
- dc.creator example
 - <meta name="dc.creator" content="Government of Canada, Department of Canadian Heritage, Parks Canada Agency">

dc.creator Guidelines

- ◆ dc.creator element should reflect organizational hierarchy
- ◆ Avoid abbreviations and acronyms
- ◆ Repeat the creator element to acknowledge important participants
- ◆ Use the title of the department or agency listed as the *applied* title in **Titles of Federal Organizations**
- ◆ For specific divisions, use the **Government Electronic Directory Services (GEDS)**

Titles of Federal Organizations

- Example usage
 - ◆ Canadian Museum of Civilization
[not Museum of Civilization]
 - ◆ Canadian Security Intelligence Service
[not CSIS]
- ***Titles of Federal Organizations***

GEDS

- Provides standard forms for names of government agencies
- Organized by departmental hierarchies
- GEDS example
 - ◆ Government of Canada, Industry Canada, Atlantic Region, Energy and Marine Branch.
- Government Electronic Directory Services (GEDS)

Mandatory Elements

dc.title

- dc.title
 - ◆ Indicates name given to the resource
- dc.title example
 - ◆ `<meta name="dc.title" content="Canadian Rural Information Service - Government of Canada Programs and Services Pathfinder">`

Mandatory Elements

dc.title - Guidelines

- ◆ Every Web page should have a unique and meaningful title:
 - *Canadian Consumer Information Gateway - Canada's Most Comprehensive Consumer Information Site*
 - *Public Safety*
- ◆ Assign a title that succinctly describes the contents of a page (70 to 115 characters)
- ◆ Ideally, the content for <dc.title> should be identical to the HTML <title> </title>

HTML Title Example



Mandatory Elements

dc.language

- dc.language
 - ◆ Indicates the language of the intellectual content of the resource
- dc.language examples
 - <meta name="dc.language" scheme="ISO639-2" content="eng">
 - <meta name="dc.language" scheme="ISO639-2" content="fre">

Mandatory Elements

dc.language - Guidelines

- ◆ French language content is tagged in French (fre)
- ◆ English language content is tagged in English (eng)
- ◆ Bilingual documents are tagged in both languages
- ◆ Scheme="ISO639-2" is a reference to the language designation scheme

Mandatory Elements

dc.language - Example

How would you tag the following page?



The screenshot shows the header of the Canada International website. It includes the Canadian flag, the text "Government of Canada" and "Gouvernement du Canada", and the "Canada" logo. Below the header is a banner for "CANADÁ INTERNACIONAL" with a background of a Canadian flag. The main content area contains a paragraph in Spanish explaining that the official languages are English and French, but some content is translated into Spanish. Below the text are three red buttons labeled "English", "Français", and "Español". At the bottom, there is a date "Última modificación: 04/04/2002" and a link "Anuncios importantes".

Government of Canada Gouvernement du Canada Canada

CANADÁ
INTERNACIONAL

Los idiomas oficiales de Canadá son el inglés y el francés, sin embargo, algo del contenido de este sitio ha sido traducido al español como parte de un ensayo piloto. Para acceder el contenido disponible en español, haga clic en el botón CanadaInternational-Spanish que aparece a continuación:

English Français Español

Última modificación: 04/04/2002 [Anuncios importantes](#)

Mandatory Elements

dc.date

- dc.date
 - ◆ Indicates a date associated with a resource
- dc.date examples
 - <meta name="dc.date.created" content="2002-04-16">
 - <meta name="dc.date.modified" content="2003-02-20">

Mandatory Elements

dc.date - Guidelines

- ◆ Format is YYYY-MM-DD (e.g. 2003-03-15)
- ◆ Use dc.date.created to indicate date that a resource is created or posted
- ◆ Use dc.date.modified to designate a date when significant changes were made
- ◆ Fill in unknown month and day with 01 (e.g. 1995-01-01)

Mandatory Elements

dc.subject

- dc.subject
 - ◆ Indicates the topic of a resource
- dc.subject example
 - ◆ `<meta name="dc.subject" scheme="gcore" content="Student loans; Scholarships; Debt; Financial assistance; Grants; Financial management; Budgets; Personal finance; Loans">`

Mandatory Elements

dc.subject - Guidelines

- ◆ Terms used must be selected from a controlled vocabulary
 - Example: *GoC Core Subject Thesaurus* (gccore)
- ◆ Indicate the subject vocabulary used in the metadata entry
 - Example: scheme=gccore

Mandatory Elements

dc.subject - Guidelines (cont.)

- ◆ Use the *GoC Core Subject Thesaurus* to locate one or more terms for the <dc.subject> metadata element
- ◆ Use the HTML <keywords> tag to include terms that are not linked to a controlled vocabulary

dc.subject and keywords

-- Examples

- Page title: Travel at Home and Abroad
- `<meta name="dc.subject" scheme="gccore" content="Federal Departments; Foreign Affairs; Travel; Information; Government Services">`
- `<meta name="keywords" content="travel in Canada, travel abroad, living abroad">`

Controlled Vocabularies

-- Examples

- ◆ *Canadian Immigration and Citizenship Indexing Terms*
- ◆ *CRTC Thesaurus*
- ◆ *Health Canada Core Controlled Vocabulary*
- ◆ *GoC Core Subject Thesaurus*

GoC Core Subject Thesaurus -- Example

- **Aircraft accidents**

FRENCH:

Accident d'avion

Subject Category:

HE Health and Safety

TR Transport

Used For:

Air disasters

Aviation accidents

Plane crashes

Broader Term:

Accidents

Related Term:

Air safety

Aircraft

GoC Core Subject Thesaurus -- Guidelines

- ◆ Search or browse for suitable terms
- ◆ “Drill-down” into any entry that is found
- ◆ Choose the most specific descriptor available.
- ◆ Choose as many descriptors as needed
- ◆ Use the HTML <keywords> metatag for descriptors not linked to a specific vocabulary

GoC Core Subject Thesaurus -- Tips

- ◆ The *GoC Core Subject Thesaurus* uses broad, high-level terms
- ◆ Select and search for a term that describes what your document is about.
- ◆ It is necessary to “drill-down” into a term to find out if it is actually used.
- ◆ Select the term that most specifically matches your concept.
- ◆ If you cannot locate the term you are looking for, use the browsing feature of the thesaurus.

Using the *GoC Core Subject Thesaurus*

- Your task is to locate correct terminology for the document you want to describe.
- Access the *GoC Core Subject Thesaurus* at:
<http://www.thesaurus.gc.ca/>

Review of TBITS 39.1 Application

- Indicates priority areas for metadata application.
- Identifies the kinds of resources that represent the most important aspects of Government On-Line information and services

Non-mandatory Elements

- HTML elements
 - ◆ description
 - ◆ keywords
- DC elements
 - ◆ dc.description
 - ◆ dc.type (e.g. guide, educational resource, policy, etc.)
 - ◆ dc.coverage.spatial (e.g. Manitoba, Akwesasne Indian Reserve 15)

Review

- Why use metadata?
- What is Dublin Core?
- What is TBITS 39.1?
- What are the 5 CL&F mandatory elements?
- What are the two parts to a metadata element?
- What is a controlled vocabulary?



Questions?

Where to find out more!

- ◆ Information Management Resource Centre (IMRC) http://www.cio-dpi.gc.ca/im-gi/meta/meta_e.asp
- ◆ Ask your departmental [metadata specialist!](#)

