|  |
| --- |
| **Open Government Common Core Metadata Element Set** |
| **Consultation Feedback Summary** |
|  |

Open Government Secretariat

Treasury Board of Canada Secretariat, Government of Canada

July 22, 2015

Contents

[**1.0 Introduction** 3](#_Toc425163767)

[**2.0 Response** 4](#_Toc425163768)

[**3.0 Summary of Feedback** 5](#_Toc425163769)

[3.1 General Comments 5](#_Toc425163770)

[3.2 Comments on new elements 5](#_Toc425163771)

[3.3 Consolidation of Feedback 7](#_Toc425163772)

[3.3.1 Metadata Identifier 7](#_Toc425163773)

[3.3.2 Title 8](#_Toc425163774)

[3.3.3 Publisher 9](#_Toc425163775)

[3.3.4 Creator 12](#_Toc425163776)

[3.3.5 Description 13](#_Toc425163777)

[3.3.6 Keywords 14](#_Toc425163778)

[3.3.7 Subject 15](#_Toc425163779)

[3.3.8 Geographic Region 16](#_Toc425163780)

[3.3.9 Date Published 17](#_Toc425163781)

[3.3.10 Date Modified 18](#_Toc425163782)

[3.3.11 Date Contributed 19](#_Toc425163783)

[3.3.12 Catalogue Type 20](#_Toc425163784)

[3.3.13 Level of Government 21](#_Toc425163785)

[3.3.14 Licence 22](#_Toc425163786)

[3.3.15 Resource Title 23](#_Toc425163787)

[3.3.16 Resource Type 24](#_Toc425163788)

[3.3.17 Resource Format 25](#_Toc425163789)

[3.3.18 Resource – Access URL 26](#_Toc425163790)

[3.3.19 Resource – Language 27](#_Toc425163791)

[3.3.20 Frequency 28](#_Toc425163792)

[3.3.21 Time Period Coverage Start and End Date 29](#_Toc425163793)

[**4.0 Outcome and Next Steps** 30](#_Toc425163794)

[4.1 Final Core and Implementation 30](#_Toc425163795)

[4.2 DCAT and Schema.org 30](#_Toc425163796)

[4.3 Metadata Mapping 30](#_Toc425163797)

[**Appendix A** 30](#_Toc425163798)

[**Appendix B** 36](#_Toc425163799)

[**Appendix C** 4](#_Toc425163800)1

[**Appendix D** 4](#_Toc425163800)3

# 1.0 Introduction

The main objective of the [Open Government Portal](http://open.canada.ca/en) is to provide access to open resources, currently through the Open Data Portal and the Open Information Portal. To ensure effective search and discovery of these resources, the Open Government Secretariat, with the assistance of the Information Management Division within the Treasury Board Secretariat, has been focusing on improving the quality of metadata used to describe the open resources. By developing the ‘Open Government Common Core Metadata Element Set’ , we can establish a set of core elements that can be used to describe all resources found on the open government portal consistently, whether they are datasets or information assets.

By applying common core elements to all datasets and information assets, we will ultimately be able to develop a single search for all open government resources, rather than the separate catalogues currently found on the site. Since the collections will continue to grow over time (i.e open maps) we would like to develop and implement the core set quickly and effectively. This core will align closely with the element set that will be implemented by the Web Renewal Initiative Group to ensure interoperability. In addition to this set of common elements, we will be able to support domain-specific metadata extensions to further describe resources found in a given ‘collection’ including, but not limited to, open maps, open publications, open science.

Developing a common core set of element sets will also assist in federating open data. If provincial and municipal open data providers apply this core, or attempt to map to it, developing an integrated search for all open government resources throughout Canada will be made simpler.

                This document outlines the feedback received from a consultation with the metadata and open data community to gather feedback on the ‘Open Government Common Core Metadata Element Set’. A draft element set was developed and shared on [GCConnex page](http://gcconnex.gc.ca/file/view/12376747/open-government-common-core-metadata-element-set) and [Github](https://github.com/open-data/open-data) (See Appendix A).

# 2.0 Response

During the consultation period, of nearly three weeks throughout the months of June and July 2015, feedback was received from the following stakeholders:

* Open North
* Government of Canada
  + Open Government Technical Working Group
  + Open Government Information Management Listserv
  + Open Government IM Awareness Listserv
  + Open Government Federal Geospatial Platform
    - Agriculture and Agri-Food Canada
    - Canada Revenue Agency
    - Canada School of Public Service
    - Department of Finance
    - Environment Canada
    - National Research Council Canada,
    - Natural Resources Canada (Federal Geospatial platform)
    - Standards Council Canada
    - Treasury Board of Canada Secretariat
    - Public Works and Government Services Canada
    - Statistics Canada
* Open Canada Working Group
  + Province of Ontario
  + Province of Newfoundland
  + City of Montreal
  + City of Vancouver,
* Open Ontario Working Group
* University of Ottawa Library, and members of the Canadian Library Association

# 3.0 Summary of Feedback

## 3.1 General Comments

It was commonly suggested to ensure that definitions were clear, and applicable to both open data and open information. Many revisions were made with this in mind, and definitions were leveraged from Dublin Core and DCAT for standardization and consistency. More in depth definitions and explanations of each element will be included in the Open Government Core Metadata Application Profile. This will be developed in August 2015, and will be shared on the [GCConnex page](http://gcconnex.gc.ca/file/view/12376747/open-government-common-core-metadata-element-set) and [Github](https://github.com/open-data/open-data). It was also recommended to encourage the development of more robust and detailed explanations for specific collections added to the Open Government Portal. For example, the metadata element set developed for the ‘Open Maps’ collection could include refined definitions, which is fully supported.

Another common recommendation was to apply DCAT (and RDF vocabulary) as well as Schema.org (an HTML language) in order to improve search of the metadata records found on the Open.Canada.ca. This was considered an important step to improve discoverability and interoperability, and as a result we have developed a project plan to implement both DCAT and Schema.org to the metadata found on the portal. The proposed project plan can be found in Appendix C, and will be consulted on in late July 2015. It will be posted on the [GCConnex page](http://gcconnex.gc.ca/file/view/12376747/open-government-common-core-metadata-element-set) and [Github](https://github.com/open-data/open-data).

## 3.2 Comments on new elements

The following is a breakdown of new elements that were suggested to be added to the common core :

| **Suggested Element** | **Response/Action** |
| --- | --- |
| Resource – Identifier | This element has been added |
| Resource – Date Modified- is needed to capture when the data on which the resource was updated as some datasets will need to be updated daily, weekly, etc | This has not been added. Although it would be a beneficial field, it would be too difficult to maintain, as some resource update hourly. We will encourage departments to populate the ‘Frequency’ field accurately in order to inform users of the frequency of updates. |
| Resource – Date Published | This element has been added |
| Resource – Format Version | This element has not been added, as we will encourage departments to provide data in various formats, and ensure that they offer continuous access, which may involve updated to new data formats. |
| Resource – Character Encoding | This has not been added because all resources released on the Open Governmental Portal need to comply with the Government of Canada’s Standard on Web Interoperability. This standard states that the mark-up language should be UTF8. |
| Resource- Description | This has not been added, as we did not think the element would be populated consistently, and therefore the added value was minimal. We will encourage them to provide supporting documentation to distinguish and describe resources, if necessary. |
| Resource Size | This element was added. |
| Related Resources – to enable departments to link to a metadata record in a different ‘catalogue’ (i.e. data, information) | This element was added. |
| Association Type (controlled list ‘is part of’, ‘supplementary’, ‘superseded’ | This element was added. |
| Metadata Contact – could be defaulted to [open-ouvert@tbs-sct.gc.ca](mailto:open-ouvert@tbs-sct.gc.ca), or other generic email boxes for various catalogues.  If a “Dataset - Contact Point” field is added, it can be formatted as structured data using the vCard spec [4], like in DCAT | This element was added. |
| Additional Publisher / Secondary Publisher | The original ‘Publisher’ element will be made repeatable so that more than one publisher can be applied to a record (the number of potential contributors could be many). |
| HomepageURL: dcat: LandingPage should be used in the machine readable data. | This element was added. |

## 3.3 Consolidation of Feedback

Below is a consolidation of feedback received, with responses, for each element provided in the draft. Note, changes made to the element are indicated in strikethrough formatting and red font.

### 3.3.1 Metadata Identifier

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **1** | **Metadata** Record **Identifier** | ~~A unique phrase or string which uniquely identifies the~~ **~~metadata record~~** ~~within the catalogue~~  A unique phrase or string which uniquely identifies the metadata record within the Open Government Portal | M | Single | System generated (or this element can be mapped to an exisiting i.d. | <http://purl.org/dc/elements/1.1/identifier> | <http://purl.org/dc/terms/identifier> |

| **Comment** | **Response/Action** |
| --- | --- |
| This field should be renamed: - the term "Metadata Identifier" is vague and does not give context to what exactly is this "identifier". It is understood that this is the unique identifier for the metadata record of a dataset (and a dataset metadata record may include multiple resources (data items)), suggestions: "Dataset Record Identifier"? | Changed to ‘Metadata Record Identifier’ |
| Rename this field to “Dataset identifier” rather than/in addition to “Metadata identifier”.   * Is the object to uniquely identify the dataset or the metadata record – | The objective of this field is to uniquely identify the metadata record, not the resource itself. This is because there could be many associated resources available on one metadata record.  However, we will add the element ‘Resource Identifier to capture a unique i.d. for a resource. |

### 3.3.2 Title

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **Title English** | The English name by which thedataset is formally known  The English name by which the dataset or information asset is known | M | Single | Free text | <http://purl.org/dc/elements/1.1/title> | <http://purl.org/dc/terms/title> |
| **Title French** | ~~The French name by which the~~~~dataset is formally known~~  The French name given to the dataset or information asset is formally known | M | Single | Free text | <http://purl.org/dc/elements/1.1/title> | <http://purl.org/dc/terms/title> |

| **Comment** | **Response/Action** |
| --- | --- |
| There are "Resource - Title English" and "Resource - Title French" (suggestions: use "Resource Title - English" and "Resource Title - French") would you consider using: Dataset Title – English / Dataset Title – French or to be consistent with the labeling of the proposed "Dataset Record Identifier": Dataset Record Title – English / Dataset Record Title – French | The definition needs to apply to all assets found on the portal (dataset or information), therefore we won’t add ‘Dataset’ to the field name. |
| Are these must be equivalent titles? Would this require translation, if a translation does not exist/is not present in documentation accompanying the data? | Yes, we will require the creation on a single bilingual metadata record. All free text fields will need to be translated. So, yes this field would need to have an English and French equivalent. |
| Title French” should use dcterms:title like “Title English”. Same with "Resource - Title English” and "Resource - Title French”. In RDF, language tags are used to distinguish the languages. | Noted, this change has been made. |

### 

### 3.3.3 Publisher

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Obligation** | **Occurrence** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **4** | **Publisher** | ~~An entity (GC Department or Agency) primarily responsible for publishing the~~ **~~resource~~**  An entity (GC Department or Agency) primarily responsible for publishing the dataset / information asset | M | ~~Single~~ Repeatable. | [Controlled Vocabulary  Government of Canada Registry of Applied Titles](http://www.tbs-sct.gc.ca/fip-pcim/reg-eng.asp) | <http://purl.org/dc/elements/1.1/publisher> | <http://purl.org/dc/terms/publisher> |

| **Comment** | **Response/Action** |
| --- | --- |
| Does the term "publishing" in the description in column D refer to  a. when the dataset is submitted to the open data portal and made available to the public? (Submitted by?)  b. when the dataset was previously published in another Web location? (e.g. Web sites of depts./agencies)  c. when the dataset was previously published by other commercial publisher? (research paper?) | a. No  b. Yes, when the item was first published  c. No |
| This "Publisher" element is in the "dataset" level, and within one dataset, there can be multiple resources, question:  a. Are all the resources necessarily be published by the same "publisher"?  b. Would it help to replace the term "publisher" with "Submitted by" or "Contributing Department/Agency" or "Contributor"?  c. Add "Publisher" to individual resource level? | a. Yes  b. We will add this as a comment, in the Application Profile, but keep the field name ‘Publisher’  c. We will keep this element at the record level. |
| Sometimes there will be more than one publisher.  Will that information be lost? | Duel ownership has been explored as a possibility; therefore we will make this element repeatable. We will apply the GC ‘Registry of Applied Title’s’, but consider modifying or updating this controlled vocabulary if necessary. |
| What if the publisher is outside of the GC but there is a sponsoring department who paid to have the resource created for distribution? | Since all resources found on the Open Government Portal must be owned by the Government of Canada\* and released under the Open Government Licence, this controlled vocabulary will not expand to support outside organizations.  \*This is subject to change, and may need to be updated when we federate resources from outside portals and organizations. |
| Department names may change over time – how will this be accommodated? Will there be reference to old department names so users can easily track the history of datasets from a department? | Currently, we are unable to capture the changes to department names. The Web Renewal Initiative will be investigating how this can be resolved, as it will need to be integrated into the ‘Digital Asset Management’ tool. We will leverage the methodology that is developed. |
| We were worried about this one since it doesn’t connect up with the Standard on Metadata where we have always used the Creator for the departments instead. As well, when we add in the publications, it will then make it harder to distinguish the actual publisher as we’ll be using Publisher for the department’s name. If you want to capture the department that is primarily responsible for creating the resource, then use Creator and use the Registry of Applied Titles. If you want the actual publisher, then I suggest Publisher and use a different controlled vocabulary which includes journal publishers, etc. | The Information Management Division within TBS indicates that the Standard on Metadata is currently being reviewed, and this discrepancy will be corrected. The future guidance will align with this application of ‘Creator’ and ‘publisher’. |
| What if data were published only in part, e.g. as part of a study for which they were prepared - **or** - released, all or in part, through some other means (e.g. website) without being formally published?  DC seems to intend formal publication – what about OG Portal? In case of data previously or separately published **in part**, e.g. as part of a study, how can the current release on OG Portal be linked to that publication/resource?  Need a field for this? | If a resource is created in part, we will enable users to add each part to an individual record, or they may create a new metadata record, and link these records by using the ‘Related Items’ element. |
| Any current list of applied titles will not be adequate for capturing certain obsolete GC entities such as Communication Canada or the Department of the Interior. | We will used the registry of Applied titles for now, and explore the need to expand or replace this controlled vocabulary in the future. |

### 3.3.4 Creator

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **5** | **Creator** | ~~An [individual] entity primarily responsible for making the~~ **~~dataset~~**  An [individual] entity primarily responsible for making the dataset / information asset | O | R |  | <http://purl.org/dc/elements/1.1/creator> | <http://purl.org/dc/terms/creator> |

| **Comment** | **Response/Action** |
| --- | --- |
| The name ‘Author’ would be ideal. | We have selected the element ‘Creator’ to align with Dublin Core. |
| This element should be required as it is not always going to be the same value as the Publisher. | Changed to ‘Optional’. |
| As the nature of general Web resources is different from the dataset in an open data portal, I think we need to add clarity to the term "Creator" here:  a. the entity which makes (creates) the dataset in the open data portal (making the dataset available in the open data portal) (based on the description, I gather this is what it meant)  b. Does it refer to: - the entity who/which is responsible for the dataset intellectual content?  c. Why is it "optional”? Is it possible/ difficult to have the "Creator" information auto-captured and populate by the system? | This definition comes from Dublin Core; however, additional notes to clarify this element will be added.  a. Yes (‘makes it available’ in terms of developing, which may or may not also be the individual responsible for providing access to it)  b. Yes  c. This element should capture an individual responsible for making the dataset / information resource, not the metadata record. However, ‘submitted by’ is stored in CKAN when a user logs in to add or update a metadata record assigned to their account or organization |

### 3.3.5 Description

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **6** | **Description English** | ~~A brief narrative summary of the~~ **~~dataset~~**~~’s contents, in English~~  An account of the dataset / information asset, in English. A description may include but is not limited to: an abstract, a table of contents, a graphical representation, or a free-text account of the resource. | M | Single | Free text (CharacterString) | <http://purl.org/dc/elements/1.1/description> | <http://purl.org/dc/terms/description> |

| **Comment** | **Response/Action** |
| --- | --- |
| This element needs to be consistent with the data format regarding the expression of the language qualifier. For example, is it "Term - English" (Description - English) or "Term English" (Description English)? | Noted, we will be consistent. |
| Is there guidance on the length and scope of descriptions that should be supplied; for example, should descriptions be like abstracts or should they provide a more general overview of the contents? | Clear guidance will be outlined in the Application Profile and Implementation Guide. |
| Does this have to be mandatory? Would the keywords/subject suffice where no abstract/summary is available? (Not all records contain a description that would map here) | Our goal is that all users can easily retrieve open government resources, and the words provided in this element improve search and discovery. Also, we want to ensure that a user can understand what an item is about without having to download it first. |

### 3.3.6 Keywords

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **8** | **Keywords English** | ~~The commonly used word(s), or formalized word(s) or phrase(s) used to describe the subject of the~~ **~~dataset, in English~~**  Commonly-used words or phrases which describe the dataset / information asset, in English | M | Repeatable | Free text (CharacterString) | [http://purl.org/dc/elements/1.1/subject  Label=Keywords](http://purl.org/dc/elements/1.1/subject%20Label=Keywords) | <http://purl.org/dc/terms/subject> |

| **Comment** | **Response/Action** |
| --- | --- |
| Since most tools are capable of doing full text searches, this element and values as optional rather than mandatory. | We are unable to support full text searches. SOLR can search for terms in the metadata record. Since the ‘Description’ fields are not always robust, we will ask departments to add additional search terms to support effective search. |
| Are you going to standardize the translation of keywords?  Some people might choose different French/English equivalents for things. | The Open Government Portal will be unable to support this, however we will provide guidance for departments, for example to use the [GC Core Subject Thesaurus](http://www.thesaurus.gc.ca/recherche-search/mtwdk.exe?k=these&l=60&s=scat&w=90) or [Répertoire de Vedettes-Matiére](https://rvmweb.bibl.ulaval.ca/en/a-propos). |
| Is this meant to accommodate individual departments’ controlled vocabularies, or similar?  Will there be a process that ensures that there is some control over keywords generated (e.g. nudge to choose from existing and create new only if nothing suitable exists)? | No, the objective of this field is to support ‘free text’, or to reflect the natural language of users’ search terms. |
| The description of this element should be updated to be more search and audience specific. | Noted, more guidance will be provided in the Open Government Common Core Metadata Application Profile (to come) |

### 3.3.7 Subject

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **10** | **Subject** | ~~The topic of the content of the~~ **~~dataset~~**  The topic of the content of the dataset / information asset | M | R | [Controlled Vocabulary Government of Canada Core Subject Thesaurus - Subject Categories](http://www.thesaurus.gc.ca/recherche-search/mtwdk.exe?k=these&l=60&s=scat&w=90) | [http://purl.org/dc/elements/1.1/subject  Label=Subject](http://purl.org/dc/elements/1.1/subject%20Label=Subject) | <http://purl.org/dc/terms/subject> |

| **Comment** | **Response/Action** |
| --- | --- |
| Are we using the full thesaurus or just the subject categories? | We are using the 19 top level “Subject Categories” |
| Is there a way someone picks the relevant subject categories and then it gives a list of the subject terms from the CST that may be relevant? The ‘keywords’ list should be formalized, so that if you had the subject field using the subject category, then in the keyword (or another label) field they pick actual terms from the CST. | Although not feasible at the moment, this suggestion will be evaluated at a later time. |
| If these elements are limited in usage to only GC-created datasets, then using GC Core Subject Thesaurus should be okay – but if expanding to additional levels of government, then it may not be sufficient.  If this element set is to be extended to resources other than government datasets, it may be wise to consider another subject vocabulary. | The GC Core Subject Thesaurus is a standard used to identify all GC assets. If there are gaps in thesauri used by those outside of the GC, subject mappings will be conducted. Also, we would encourage departments to populate this element by using words from their own thesauri. |
| Currently, both “keywords” and “subject” are mapped to dcterms:subject. “keywords” should be mapped to dcat:keyword [1]. “subject” can stay as-is, or be mapped to dcat:theme [2] which is a sub-property of dcterms:subject, to be better aligned with DCAT. | Noted, we will align definitions with the proper Dublin core mapping. |
| Is the [GC Core Subject Thesaurus](http://www.thesaurus.gc.ca/default.asp?lang=En&n=B1B9013A-1) maintained and developed to meet the business needs of the portals? Are there communications between TBS and LAC to make sure that there is / will be continuing development / maintenance work to be done in the GC CORE so that it can meet the GC business needs with regard to the Open Government Portals? | It is still being maintained, however the maintenance and review of the GC Core Subject Thesaurus is currently under review. It is TBS’s objective that the process to update the thesaurus will be improved to better align with user and GC business needs. |

### 3.3.8 Geographic Region

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Obli.** | **Occ.** | **Encoding Scheme** | **DC Elements** | **DC Terms** |
| **New** | **Spatial** | The range of spatial applicability of a dataset / information asset. This element could include either a bounding box, or GeoJson string. | O | Rep. | Free text (BoundingBox or GeoJson string) |  |  |
| **11** | **Geographic Region Name** | ~~The geographic area of the~~ **~~dataset~~**  The geographic region of a dataset / information asset. This element will be a region name. | ~~M~~  O | Rep. | [Controlled Vocabulary  Standard Geographical Classification (SGC) 2011](http://www23.statcan.gc.ca/imdb/p3VD.pl?Function=getVD&TVD=116940) | <http://purl.org/dc/elements/1.1/coverage> | <http://purl.org/dc/terms/spatial> |

| **Comment** | **Response/Action** |
| --- | --- |
| Clarify the description, as this does not indicate if you mean the geographic region relating to the data in the dataset or the geographic region where the dataset was created. | The description has been updated. |
| This could be optional, as it may not always have a geographic area. This should either be optional on the common core or moved to each of the extensions (and may be optional or mandatory depending on if it is for geospatial, etc.). As well, information may not have a geographic region at all. | This element was made optional. The default setting of ‘Canada’ that would be applied if this element was mandatory would not improve search and discovery. |
| The Standard Geographical Classification is very high level. If you only target this, it is of very limited use considering there is no other geospatial field. I suggest using a pure geospatial format (geojson or wkt) or using the locations as formatted by GeoNamesgeonames. The SGC can represent regions, provinces/territories, census divisions, and census subdivisions, which is decent. However, it’s hard to get agreement across governments on how to represent this field. In our first Open Data Working Group report, we discuss the options [2]. The US allows four different options to satisfy different publishers’ requirements [4]: bounding box, point, GML, or GeoNames. | We have added an additional element to capture a bounding box, or GeoJson string, which is supported by CKAN. We will work towards combining these two elements in the future. |

### 3.3.9 Date Published

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **12** | **Date Published** | ~~Date on which the~~ **~~dataset~~** ~~was created~~  The date of issuance (e.g. publication) of the dataset or information asset. | M | Single | [W3C Date and Time Formats (YYYY-MM-DD)](http://www.w3.org/TR/NOTE-datetime) | <http://purl.org/dc/elements/1.1/date> | <http://purl.org/dc/terms/created> |

| **Comments** | **Response/Action** |
| --- | --- |
| Does this when the record was published on the OG Portal? | No, that element would be ‘Date Released’—this element captures when the item was first published. |
| Should we add ‘date Published’ at a resource level? | This has been added to the elements applied to each resource. |
| Since it is possible for a dataset to be created on one day and published on another, the description is confusing.   Choose either use Date created or Date published as the metadata element OR have TWO metadata elements…one for Date created and another for Date published. | The description has been updated for consistency. |
| Are we trying to capture when it was created or when it was published?  a. If it is published, we have always used Date issued in the past (term name: Issued). Its description is ‘Date of formal issuance (e.g. publication) of the resource.’ This would align it with the Standard on Metadata if we used issued and it would better describe the published date.  b. If we are trying to capture when it was actually originally created, then we should use Date Created (term name: created). | a. The definition has been updated.  b. This property should be set using the first known date of issuance. |

### 3.3.10 Date Modified

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Notes** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **13** | **Date Modified** | ~~Date on which the~~ **~~metadata record~~** ~~was modified~~  The date on which the dataset / information asset was changed, updated or modified. | ~~M~~  O | Single | W3C Date no less than YYYY-MM-DD up to a fully qualified Date time | [W3C Date and Time Formats (YYYY-MM-DD)](http://www.w3.org/TR/NOTE-datetime) | <http://purl.org/dc/elements/1.1/date> | htttp://purl.org/dc/terms/modified |

| **Comments** | **Response/Action** |
| --- | --- |
| This element should be optional, as not everything will be modified | The obligation of this element has been changed to ‘Optional’. |
| Since a dataset could be modified more than once, perhaps this should be a Repeatable value | We won’t track the history of all updates made, rather, this field will capture the latest update made. The ‘Frequency’ field will identify how often updates are made. |
| The resource itself would require a date modified as well. So there should be a Resource – Date Modified added | This was considered; however, it was not added due to difficulty in maintaining. A resource could potentially change every hour, and if a department did not edit this date at a resource level to reflect the change, the element would be ineffective and irrelevant. |
| As a single date, is this intended to include only the last date of modification of the metadata record for the dataset – i.e. a list of all dates modified would not be possible and is not the intention? Or could this element be repeated? | The value of this property indicates a change to the actual dataset, not a change to the catalog record. An absent value may indicate that the dataset has never changed after its initial publication, that the date of last modification is not known, or that the dataset is continuously updated. |

### 3.3.11 Date Contributed

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **14** | **Date ~~Contributed~~ Released** | ~~Date on which the~~ **~~metadata record~~** ~~was contributed to the Open Government Portal~~  The date on which the metadata record was made available on the Open Government Portal. | M | Single | [W3C Date and Time Formats (YYYY-MM-DD)](http://www.w3.org/TR/NOTE-datetime) | <http://purl.org/dc/elements/1.1/date> | <http://purl.org/dc/terms/dateSubmitted> |

| **Comments** | **Response/Action** |
| --- | --- |
| Is the objective of this field to capture when an asset was made available on the portal? As the datasets may have been developed sometime before and are not just being made available, the term name, available may be useful. It could be a better fit than “Date Submitted”.  I understand "Date Contributed" is in a more neutral tone; however, "Date Submitted" is more to the point. | We felt that ‘Date Submitted’ captures when the record was submitted to CKAN, rather to the portal. Our objective is to use this element to prioritize recently added resources; therefore, ‘Date Released’ is more applicable. |
| DCAT uses dcterms:issued instead of dcterms:dateSubmitted [3] for the date on which a metadata record is contributed | We will align to this term. |

### 3.3.12 Catalogue Type

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **15** | **~~Catalogue Type~~**  **Portal Type** | ~~The GC Open Government Catalogue that the~~ **~~dataset~~** ~~belongs to (i.e. the Open Data Portal or Open Information Portal)~~  The portal to which the metadata record belongs (Open Data or Open Information). | M | Single | Controlled Vocabulary   GC Defined ~~(i.e. Non-Geospatial, Geospatial, Open Maps,~~ Open Information) (Open Data, Open Information) | <http://purl.org/dc/elements/1.1/type> | <http://purl.org/dc/terms/type> |

| **Comments** | **Response/Action** |
| --- | --- |
| You mention “open information” as a possible value, but isn’t that for documents and publications?  If so, the rest of the element set is often referred to as if it was describing data rather than a publication, and there are other descriptors related to publications that aren’t treated here (e.g. departmental or Government of Canada identifiers, number in series, periodical frequency – though that could be covered by “Frequency”) | This ‘Common Core Metadata Element Set’ will be used to describe all Open Government Resources. It is noted that the definitions should be revised to ensure that the element describes both datasets and information assets. |
| The controlled list of values needs to be shared to better understand the purpose of this element. If it’s solely for the purpose of identifying whether something was a catalogued as “Open Information” vs. “Open Data”, there is little value since users will not use this to search | This element will be used as a facet to allow users to refine their search to either a dataset (Open Data Portal) or an information asset (Open Information Portal). We have also added a ‘Collection Type’ facet, to allow users to search a specific collection, and the additional metadata elements captured in the domain specific extension (i.e. open maps, open publications). |
| The GC Defined CV and definition are inconsistent. If the definition is correct, but the encoding is incorrect, then dcterms:isPartOf should be used instead of dcterms:type | Noted, there will be better consistency between the controlled list and the ‘Catalogue Type’ facets that are developed |
| The term "Catalogue" makes it very much like the traditional library; based on the description, would you consider using Portal Type? | We will use the term ‘Portal’ to avoid confusion and be consistent. |

### 3.3.13 Level of Government

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **16** | **~~Level of Government~~**  **Jurisdiction** | ~~This element will indicate which level of government has contributed the data. This element will be added to the common core when the Open Data Portal federates National data~~  The extent or range of judicial, law enforcement, or other authority. For example, the level of government has contributed the dataset or open information asset. | M | Single | Controlled Vocabulary  GC (i.e. Federal, Provincial, Municipal) |  | http://purl.org/dc/terms/Jurisdiction |

| **Comments** | **Response/Action** |
| --- | --- |
| This value is only meaningful if the portal indexes data that is not just at the federal level. The element is made tangibly useful in a search facet. If the above is not implemented, there is little value to having this metadata element | Note, this element will be used to facet resources from various levels of government when metadata records are federated from other portals. For all resources provided by GC Departments and Agencies, this element will be defaulted to ‘Federal’. |
| There is no property dcterms:jurisdiction for "Level of Government”. There is just a class dcterms:Jurisdiction (capitalized). Classes cannot be used as properties. You can use dcterms:coverage for this instead, whose value can be a Jurisdiction, Location or Period.- | Noted, this update has been made. |

### 3.3.14 Licence

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **17** | **Licence** | ~~The limitations or constraints on the use of the~~ **~~dataset~~**~~.~~  The licence applied to the dataset / information asset. It may outline the limitations or constraints on the use of the item. | M | Single | Set text (for all GC resources) Open Government Licence - Canada | <http://purl.org/dc/elements/1.1/rights> | <http://purl.org/dc/terms/license> |

| **Comments** | **Response/Action** |
| --- | --- |
| For older items in the “open information” collection, are they really going to get “OGL” by default?  What about 3rd-party materials in older publications? | All resources found on the Open Government Portal will be released under the open government licence. It will be a department’s responsibility to ensure this. |
| Are all the metadata descriptions "open" under the license also? – Yes, all metadata records will be reusable under the OGL.  There are some existing predefined value list for licence : <http://clipol.org/> It would help to have a defined set of values (and try to limit as much a possible proliferation). The alternative is to use the canonical URL of the licence instead of a value list. | This will be considered. |
| This indicates that only one form of licence is allowed for one set of dataset. Would there be a scenario that the various "resources" attached a different form of licence? | All resources will need to be released under the OGL; if there are conflicts, they will not be in scope and not included on the Open Government Portal. |

### 3.3.15 Resource Title

Note for all ‘resource’ elements: These elements are not repeatable within a resource, but multiple resources can be added to a record

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **18** | **~~Resource~~ Title English** | ~~The name by which the cited~~ **~~resource~~** ~~is known~~  An English name given to the resource. | M | Single | Free text (CharacterString) |  |  |

| **Comments** | **Response/Action** |
| --- | --- |
| The definitions of resource title and Title should be clearer.  Clarify the distinction between "known" and "formally known"? | The definition of this element has been updated to reflect this |
| If the same dataset is available in multiple formats, the metadata record and each individual resource sub-record may have the same title, or they may all be different.  Just something to think about related to searching/indexing | We will ensure that the metadata documentation provides clear guidance on naming conventions. |
| A resource was defined as an individual data item (a component piece of a dataset). It seemed that only datasets (the aggregates) were being described, but this element makes me wonder if individual data items within datasets are also being described | Yes, a resource could be both an aggregate of a dataset (i.e. a specific region or year), or an individual data item (i.e. in a specific format). |

### 3.3.16 Resource Type

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **20** | **~~Resource~~ Content Type** | The nature or genre of the content of the **resource.** (For example: dataset, supporting documentation, data dictionary, API or App). | M | Single | Controlled Vocabulary  GC Defined | <http://purl.org/dc/elements/1.1/type> | <http://purl.org/dc/terms/type> |

| **Comments** | **Response/Action** |
| --- | --- |
| There should be no overlap of terms here. We could use this to add the fields needed for each extension. So, if they choose geospatial, then a set of fields comes up for geospatial. We aren’t quite sure on the controlled vocabulary to be used here, but something that captures the geospatial data, raw data, publications (etc.) and maintains uniqueness between the terms. | The field you are referring to is ‘Catalogue Type,’ which would be used to allow users to facet on a collection. This ‘resource type’ field defines what ‘content type’. The definition of this element will be updated to avoid confusion. |
| “Dataset” is included as an example, but this cannot be true, since a resource was defined as an individual data item. A dataset would be an aggregate of resources (individual data items). Correct? So, resource is at the item-level, while dataset is at the aggregate level - they cannot mean the same thing. | We will clarify the definition. To be clear, this element defines the specific content type of the individual resource, and not the ‘collection’ or dataset. Since multiple resources can be added to a single record, we will need to support the description of a specific type. |
| In this element set, "Resource Type" is designated as a "single" occurrence. This metadata element is similar to the use of "document type" in EDRMS - when allowing just "single" occurrence, it diminishes the objectivity of this data field. | Note: this element has a single occurrence for each ‘resource’ added; however, multiple resources can be added. Also note that we will leverage the Controlled Vocabulary used by the Web Renewal Initiative to capture ‘Content Type’ consistently. |

### 3.3.17 Resource Format

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Obligation** | **Occurrence** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **21** | **~~Resource~~ Format** | ~~The physical or digital manifestation of the~~ **~~resource~~**  The file format, physical medium, or dimensions of the resource. | M | Single | Controlled Vocabulary  GC Defined | <http://purl.org/dc/elements/1.1/format> | <http://purl.org/dc/terms/format> |

| **Comments** | **Response/Action** |
| --- | --- |
| Which controlled vocabulary do you plan to use? | This will be GC Derived. |
| Is there a commonly-accepted, well-understood existing definition of format that you could use instead? | We have applied the Dublin Core definition to ensure consistency |
| What is the vocabulary for "Resource - Format”? If the vocabulary terms are meant to be human-readable, there should be an additional field “Resource - Media Type”, using dcat:mediaType [5], whose encoding is an IANA media type [6]. This makes the data more machine-readable. The "Resource - Media Type” can be automatically set based on "Resource - Format”. | This information will be captured in the element ‘Content Type’ |
| What is the scope of collection of the open data portal?  a. does it include digital data made available on Web only?  b. does it include formats other than Web data? That is to say, is there a chance that the dataset record is a dummy record created to relate the information on the dataset whereas the physical object is hosted/stored somewhere?  c. If so, the "physical" and "digital" resource would ideally be listed as 2 separate data fields, or the controlled vocabulary needs to nested properly to reflect the different formats. Physical format, and Digital format - does it refer to "file format" or does it refer to DVD, USB keys? | a. Yes  b. No, all resources must be available for download. The metadata record in the Open Government Portal will provide direct access to it.  c. Noted. If the scope expands to capture physical items, this will be explored. |

### 3.3.18 Resource – Access URL

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **22** | **~~Resource~~**  **~~Access~~ Download URL** | ~~The electronic location where the~~ **~~resource~~** ~~can be found.~~  The location for online access to the distribution of the resource | M | Single | Free text (CharacterString) Website URL | dcat:downloadURL |  |

| **Comments** | **Response/Action** |
| --- | --- |
| We are concerned about using a URL and feel a persistent identifier should be used | Since the Open Government Portal contains a registry that houses metadata and not the resource itself, it is difficult to ensure that URI assigned. We are not able to manage this because departments are responsible for storing and maintaining their own resources. However, we will provide guidance to encourage departments to assign URIs, and even DOIs, to their resources. |
| Do you want to have people use DOIs or perhaps develop something specific to the government? | Yes, the GC does use and apply DOIs. There will be a separate field in the domain-specific extensions (i.e. Open Data, Open Information) to capture this. |
| Resource - Access URL” should use dcat:accessURL | We have decided to apply ‘dcat:downloadURL’. This is because DCAT notes : accessURL, and not downloadURL, when it is definitely not a download or when you are not sure whether it is. If the distribution(s) are accessible only through a landing page (i.e. direct download URLs are not known), then the landing page link SHOULD be duplicated as accessURL on a distribution. Therefore, since the Open Government Portal will always provide direct access to the dataset, we will use ‘dcatLdownloadURL’, as access will not be made only through a landing page. However, we will add the element ‘landing’ page to allow departments to provide a link the page where the data or other contextual information resides. |

### 

### 3.3.19 Resource – Language

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Obligation** | **Occurrence** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **23** | **~~Resource~~ Language** | The language of the content of the **resource** | M | Single | ISO 639-2/T | <http://purl.org/dc/elements/1.1/language> | <http://purl.org/dc/terms/language> |

| **Comments** | **Response/Action** |
| --- | --- |
| Why is ISO 639-2/B being used when the Government of Canada has chosen to use ISO 639-2/T in our standards as the terminology code uses the proper name while the bibliographic code uses an abbreviation of the English name of the language? If we used ISO 639 2/B, then we’d use the abbreviation ‘fre’ instead of the ISO 639-2/T version ‘fra’ | Thank you for pointing this error out. We are currently implementing, [ISO 639-2/**T**](http://loc.gov/standards/iso639-2/php/code_list.php) not ISO 639-2/**B** |

### 3.3.20 Frequency

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **24** | **Frequency** | ~~The frequency with which changes and additions are made to the~~ **~~dataset~~** ~~after the initial resource is completed.~~  The frequency with which changes and additions are made to the dataset / information asset after the initial resource is completed | M | Single | [Drop down : Controlled vocabulary North American Profile ISO:19115 napMD\_MaintenanceFrequencyCode](http://nap.geogratis.gc.ca/metadata/register/registerItemClasses-eng.html#IC_102) | dcterms:accrualPeriodicity. |  |

| **Comments** | **Response/Action** |
| --- | --- |
| Will the frequency always be known? Perhaps this should be optional to include situations where frequency of updates is not known at time of creation, or include a status of "Unknown" to the list of controlled vocabulary to be used in the drop down list for this metadata value | If the frequency is unknown, departments can select ‘Unknown’ from the controlled vocabulary. (The complete list includes: [continual](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_532), [daily](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_533), [weekly](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_534), [fortnightly](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_535), [monthly](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_536), [quarterly](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_537), [biannually](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_538), [annually](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_539), [asNeeded](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_540), [irregular](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_541), [notPlanned](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_542), [unknown](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_543), [semimonthly](http://nap.geogratis.gc.ca/metadata/register/registerItems-eng.html#RI_544) |
| Frequency is very important to add, because it will indicate to the user if / when they can expect updates to a dataset. | Agreed |
| Frequency should use dcterms:accrualPeriodicity | Noted. |
| What is the Purpose/function of this metadata field? - If it changes daily, weekly, monthly, bi-weekly, annually etc. what does it mean? | The purpose is to capture the frequency with which modifications are made to the dataset after it is first published. |
| Date Modified is available, why is it necessary to have this field? | Date modified captures the update made, where this element captures the expected frequency of updates, if known. |

### 3.3.21 Time Period Coverage Start and End Date

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Dublin Core Elements** | **Dublin Core Terms** |
| **~~25~~** | **~~Time Period Coverage Start~~** | ~~The start date or time interval(s) covered by the~~ **~~dataset~~**~~. (YYYY-MM-DD)~~ | ~~O~~ | ~~Single~~ | ~~free text, YYYY-MM-DD~~  ~~IO: 8601~~ |  |  |
| **~~26~~** | **~~Time Period Coverage End~~** | ~~The end date or time interval(s) covered by the~~ **~~dataset~~**~~. (YYYY-MM-DD)~~ | ~~O~~ | ~~Single~~ | ~~free text, YYYY-MM-DD~~ |  |  |
|  | **Temporal** | The range of temporal applicability of the dataset / information asset. For example, the start and end date covered by the data. | O | Single | Free-text :  W3C Date and Time Formats (YYYY-MM-DD) |  |  |

| **Comments** | **Response/Action** |
| --- | --- |
| Time Period” should use dcterms:temporal . ISO 8601 uses a solidus (forward slash /) to separate start and end dates. So “YYYY-MM-DD/YYYY-MM-DD”. To lend more support to using ISO 8601 for temporal extent, it’s what the US does: <https://project-open-data.cio.gov/v1.1/schema/#temporal> | Noted, we will follow this convention. |

# 4.0 Outcome and Next Steps

## 4.1 Final Core and Implementation

The Open Government Common Core Metadata Element Set has been completed (See Appendix B.)

This element set will be shared with the CKAN developers at Statistics Canada, who will begin implementing the core. They will build this core in a single, common CKAN instance, and then migrate records from both the Open Data and Open Information Portal, which currently have been developed on separate CKAN instances. Records from other catalogues, including GeoGratis, and Open Maps (from the Federal Geospatial Platform’s catalogue ‘GeoNetwork’) will then be migrated. For a mapping of the Open Government Common Core to the metadata schemes applied to existing collections in CKAN, see Appendix C.

## 4.2 DCAT and Schema.org

During the development of the single CKAN instance, we will also be implementing DCAT and Schema.org. As mentioned, we received a lot of feedback recommending that we follow this best practice. For more information, see the ‘DCAT and Schema.org Implementation Project Plan’ (Appendix D).

## 4.3 Metadata Mapping

## In order to support interoperability, and potential metadata federation, we would like to encourage Provinces, Territories, and municipalities, participate in an ‘Open Canada Metadata Mapping’ exercise. This activity will allow us to explore how the metadata elements sets that have been developed to describe open data, and information resources, align or not. You can add your scheme directly on the Github page, or submit it to [open-ouvert@tbs-sct.gc.ca](mailto:open-ouvert@tbs-sct.gc.ca). Periodically, all submissions will be collected, and all participating jurisdictions metadata mapping will be added to a single mapping document. **Appendix A**

**Draft Open Government Common Core Metadata Element Set – 2015-06-08**

Notes :

- This Metadata Element Set supports the creation of a single bilingual record. Therefore, most free texts fields require an English and French input. The rest of the elements will have a French / or English equivalent system generated.

- The elements indicated in gray are used to describe each 'resource'. According to CKAN, a resource ‘represents individual data items in a dataset. The Open Data Registry enables departments to group resources together if the content is similar in some way. For example, a dataset may contain multiple resources that represent the same underlying data, but in different formats (i.e. csv and xls files) or language (i.e. English, French or Bilingual). Also, a dataset may have other resources such as supplementary documentation (i.e. data dictionary) or other information that outlines how the data was collected. A user can add as many resources that relate to the dataset that is being entered.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Name** | | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | | **Notes** | **Dublin Core Elements** | **Dublin Core Terms** |
| **1** | **Metadata Identifier** | | A unique phrase or string which uniquely identifies the **metadata record** within the catalogue. | M | Single | System generated | |  | <http://purl.org/dc/elements/1.1/identifier> | <http://purl.org/dc/terms/identifier> |
| **2** | **Title English** | | The English name by which the **dataset** is formally known | M | Single | Free text | |  | <http://purl.org/dc/elements/1.1/title> | <http://purl.org/dc/terms/title> |
| **3** | **Title French** | | The French name by which the **dataset** is formally known | M | Single | Free text | |  |  |  |
| **4** | **Publisher** | | An entity (GC Department or Agency) primarily responsible for publishing the **resource** | M | Single | [Controlled Vocabulary  Government of Canada Registry of Applied Titles](http://www.tbs-sct.gc.ca/fip-pcim/reg-eng.asp) | |  | <http://purl.org/dc/elements/1.1/publisher> | <http://purl.org/dc/terms/publisher> |
| **5** | **Creator** | | An [individual] entity primarily responsible for making the **dataset** | O | Repeatable |  | |  | <http://purl.org/dc/elements/1.1/creator> | <http://purl.org/dc/terms/creator> |
| **6** | **Description English** | | A brief narrative summary of the **dataset**’s contents, in English | M | Single | Free text (CharacterString) | |  | <http://purl.org/dc/elements/1.1/description> | <http://purl.org/dc/terms/description> |
| **7** | **Description French** | | A brief narrative summary of the **dataset**’s contents, in French | M | Single | Free text (CharacterString) | |  | <http://purl.org/dc/elements/1.1/description> | <http://purl.org/dc/terms/description> |
| **8** | **Keywords English** | | The commonly used word(s), or formalized word(s) or phrase(s) used to describe the subject of the **dataset, in English** | M | Repeatable | Free text (CharacterString) | |  | [http://purl.org/dc/elements/1.1/subject  Label=Keywords](http://purl.org/dc/elements/1.1/subject%20Label=Keywords) | <http://purl.org/dc/terms/subject> |
| **9** | **Keywords French** | | The commonly used word(s), or formalized word(s) or phrase(s) used to describe the subject of the **dataset, in French** | M | Repeatable | Free text (CharacterString) | |  | [http://purl.org/dc/elements/1.1/subject  Label=Keywords](http://purl.org/dc/elements/1.1/subject%20Label=Keywords) | <http://purl.org/dc/terms/subject> |
| **10** | **Subject** | | The topic of the content of the **dataset** | M | Repeatable | [Controlled Vocabulary Government of Canada Core Subject Thesaurus - Subject Categories](http://www.thesaurus.gc.ca/recherche-search/mtwdk.exe?k=these&l=60&s=scat&w=90) | | Note, the English controlled list will be mapped to the French, so only one input is required. | [http://purl.org/dc/elements/1.1/subject  Label=Subject](http://purl.org/dc/elements/1.1/subject%20Label=Subject) | <http://purl.org/dc/terms/subject> |
| **11** | **Geographic Region** | | The geographic area of the **dataset** | M | Repeatable | [Controlled Vocabulary  Standard Geographical Classification (SGC) 2011](http://www23.statcan.gc.ca/imdb/p3VD.pl?Function=getVD&TVD=116940) | |  | <http://purl.org/dc/elements/1.1/coverage> | <http://purl.org/dc/terms/spatial> |
| **12** | **Date Published** | | Date on which the **dataset** was created | M | Single | [W3C Date and Time Formats (YYYY-MM-DD)](http://www.w3.org/TR/NOTE-datetime) | |  | <http://purl.org/dc/elements/1.1/date> | <http://purl.org/dc/terms/created> |
| **13** | **Date Modified** | | Date on which the **metadata record** was modified | M | Single | [W3C Date and Time Formats (YYYY-MM-DD)](http://www.w3.org/TR/NOTE-datetime) | |  | <http://purl.org/dc/elements/1.1/date> | htttp://purl.org/dc/terms/modified |
| **14** | **Date Contributed** | | Date on which the **metadata record** was contributed to the Open Government Portal | M | Single | [W3C Date and Time Formats (YYYY-MM-DD)](http://www.w3.org/TR/NOTE-datetime) | |  | <http://purl.org/dc/elements/1.1/date> | <http://purl.org/dc/terms/dateSubmitted> |
| **15** | **Catalogue Type** | | The GC Open Government Catalogue that the **dataset** belongs to (i.e. the Open Data Portal or Open Information Portal) | M | Single | Controlled Vocabulary   GC Defined (i.e. Non-Geospatial, Geospatial, Open Maps, Open Information) | |  | <http://purl.org/dc/elements/1.1/type> | <http://purl.org/dc/terms/type> |
| **16** | **Level of Government** | | This element will indicate which level of government has contributed the data. This element will be added to the common core when the Open Data Portal federates National data | M | Single | Controlled Vocabulary  GC (i.e. Federal, Provincial, Municipal) | | This field will be automatically populated with Federal when records are eventually harvested from other jurisdictions |  | <http://purl.org/dc/terms/jurisdiction> |
| **17** | **Licence** | | The limitations or constraints on the use of the **dataset**. | M | Single | Set text (for all GC resources) Open Government Licence - Canada | | When other collections are added from provinces and municipalities, this field will include a controlled list of all versions of the OGL | <http://purl.org/dc/elements/1.1/rights> | <http://purl.org/dc/terms/license> |
| **18** | **Resource - Title English** | | The name by which the cited **resource** is known | M | Single | Free text (CharacterString) | | Note : This element is not repeatable within a resource, but multiple resources can be added to a record |  |  |
| **19** | **Resource - Title French** | | The name by which the cited **resource** is known | M | Single | Free text (CharacterString) | | Note : This element is not repeatable within a resource, but multiple resources can be added to a record |  |  |
| **20** | **Resource Type** | | The nature or genre of the content of the **resource.** | M | Single | Controlled Vocabulary  GC Defined (dataset, supporting documentation, data dictionary, API or App) | | Note : This element is not repeatable within a resource, but multiple resources can be added to a record | <http://purl.org/dc/elements/1.1/type> | <http://purl.org/dc/terms/type> |
| **21** | **Resource - Format** | | The physical or digital manifestation of the **resource** | M | Single | Controlled Vocabulary  GC Defined | | Note : This element is not repeatable within a resource, but multiple resources can be added to a record | <http://purl.org/dc/elements/1.1/format> | <http://purl.org/dc/terms/format> |
| **22** | **Resource - Access URL** | | The electronic location where the **resource** can be found. | M | Single | Free text (CharacterString) Website URL | | Note : This element is not repeatable within a resource, but multiple resources can be added to a record |  |  |
| **23** | **Resource -Language** | | The language of the content of the **resource** | M | Single | [ISO 639-2/B](http://loc.gov/standards/iso639-2/php/code_list.php) | | Note : This element is not repeatable within a resource, but multiple resources can be added to a record | <http://purl.org/dc/elements/1.1/language> | <http://purl.org/dc/terms/language> |
| **Considerations** | | | | | | | | | | |
| **24** | | **Frequency** | The frequency with which changes and additions are made to the **dataset** after the initial resource is completed. | M | Single | [Drop down : Controlled vocabulary North American Profile ISO:19115 napMD\_MaintenanceFrequencyCode](http://nap.geogratis.gc.ca/metadata/register/registerItemClasses-eng.html#IC_102) |  | |  |  |
| **25** | | **Time Period Coverage Start** | The start date or time interval(s) covered by the **dataset**. (YYYY-MM-DD) | O | Single | free text, YYYY-MM-DD |  | |  |  |
| **26** | | **Time Period Coverage End** | The end date or time interval(s) covered by the **dataset**. (YYYY-MM-DD) | O | Single | free text, YYYY-MM-DD |  | |  |  |

# Appendix B

**Final Open Government Common Core Metadata Element Set – 2015-07-22**

|  |
| --- |
| **Catalog Metadata** |
| **Asset Metadata** |
| **Distribution (Resource) Metadata** |

M – Mandatory, O – Optional

S – Single, R- Repeatable

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Metadata Type** | **Name** | **Description** | **Oblig.** | **Occur.** | **Encoding Scheme** | **Comments** |
| **1** | **Catalog Metadata** | **Portal Type** | The portal to which the metadata record belongs (Open Data or Open Information) | M | S | Controlled Vocabulary   OGS Defined | This element will be system generated. |
| **2** | **Collection Type** | The collection (domain specific) to which the metadata record belongs to, for example 'Open Maps, Non-Spatial Data, Open Publications etc.) | M | S | Controlled Vocabulary   OGS Defined | This element will be system generated, and determined by identifying the source of the metadata record. |
| **3** | **Metadata Scheme** | The name of the metadata scheme (including profile name) used. | M | S | Free text (URL) | This element will be system generated to provide a URL to the Open Government Metadata Application Profile, and related information (including machine readable version of the schema) located on the open data portal |
| **4** | **Metadata Scheme Version** | The version of the metadata scheme (version of the profile) used. | M | S |  | This element will be system generated. |
| **5** | **Dataset / Asset Metadata** | **Metadata Record Identifier** | A unique phrase or string which identifies the metadata record within the Open Government Portal | M | S | Free text | This element will be system generated, unless a department has already captured a unique ID for the record in their catalogue. If so, then they can map their Unique ID to this field. |
| **6** | **Title English** | The English name by which the dataset or information asset is known | M | S | Free text | This element should be populated with plain language, and include sufficient detail to support search and discovery. Avoid using acronyms, or table numbers. |
| **7** | **Title French** | The French name by which the dataset or information asset is formally known | M | S | Free text | This element should be populated with plain language, and include sufficient detail to support search and discovery. Avoid using acronyms, or table numbers. |
| **8** | **Publisher** | An entity (GC Department or Agency) primarily responsible for publishing the dataset or information asset | M | R | [Controlled Vocabulary  Government of Canada Registry of Applied Titles](http://www.tbs-sct.gc.ca/fip-pcim/reg-eng.asp) |  |
| **9** | **Creator** | An [individual] entity primarily responsible for making the dataset or information asset | O | R | Free text |  |
| **10** | **Contact Email** | The contact person's email for the dataset or information asset | M | S | Free text | This will be defaulted to 'open-ouvert@tbs-sct.gc.ca', however departments will be encourage to populate this with their own generic email accounts |
| **11** | **Description English** | An account of the dataset or information asset, in English. A description may include but is not limited to: an abstract, a table of contents, or a free-text account of the resource. | M | S | Free text (CharacterString) | This element should be populated with a human-readable description of the dataset / information asset with enough detail to allow the user to easily understand what the item is about. |
| **12** | **Description French** | An account of the dataset or information asset, in French. A description may include but is not limited to: an abstract, a table of contents, or a free-text account of the resource. | M | S | Free text (CharacterString) | This element should be populated with a human-readable description of the dataset / information asset with enough detail to allow the user to easily understand what the item is about. |
| **13** | **Keywords English** | Commonly used words or phrases which describe the dataset or information asset, in English | M | R | Free text (CharacterString) | This element should be populated with plain language terms or phrases that match the terms expected to be used by target, defined audiences to support search |
| **14** | **Keywords French** | Commonly used words or phrases which describe the dataset or information asset, in French | M | R | Free text (CharacterString) | This element should be populated with plain language terms or phrases that match the terms expected to be used by target, defined audiences to support search |
| **15** | **Subject** | The topic of the content of the dataset or information asset | M | R | [Controlled Vocabulary Government of Canada Core Subject Thesaurus - Subject Categories](http://www.thesaurus.gc.ca/recherche-search/mtwdk.exe?k=these&l=60&s=scat&w=90) | Note, the English controlled list will be mapped to the French, so only one input is required. |
| **16** | **Spatial** | The range of spatial applicability of a dataset or information asset. This element could include either a bounding box, or GeoJson string. | O | R | Free text (CharacterString) | Bounding Box coordinates, or a GeoJson string are acceptable values |
| **17** | **Geographic Region Name** | The geographic region of a dataset or information asset. This element will be a region name. | O | R | [Controlled Vocabulary  Standard Geographical Classification (SGC) 2011](http://www23.statcan.gc.ca/imdb/p3VD.pl?Function=getVD&TVD=116940) |  |
| **18** | **Temporal** | The range of temporal applicability of the dataset or information asset. For example, the start and end date covered by the data. | O | S | Free-text :  W3C Date and Time Formats (YYYY-MM-DD) | This element will be expressed by YYYY-MM-DD / YYYY-MM-DD to represent the start - end date of coverage. If the end date is unknown, leave end date blank. |
| **19** | **Frequency** | The frequency with which changes and additions are made to the dataset or information asset after the initial resource is completed | M | S | [Controlled Vocabulary :  North American Profile ISO:19115 napMD\_MaintenanceFrequencyCode](http://nap.geogratis.gc.ca/metadata/register/registerItemClasses-eng.html#IC_102) | This element indicates the frequency of changes and additions made to the resource after the initial completion. If the frequency is unknown, select 'unknown' from the code list. |
| **20** | **Date Published** | The date of issuance (e.g., publication) of the dataset or information asset. | M | S | Free-text :  W3C Date and Time Formats (YYYY-MM-DD) | W3C Date no less than YYYY-MM-DD up to a fully qualified Date time |
| **21** | **Date Modified** | The date on which the dataset or information asset was changed, updated or modified. | O | S | Free-text :  W3C Date and Time Formats (YYYY-MM-DD) | W3C Date no less than YYYY-MM-DD up to a fully qualified Date time |
| **22** | **Date Released** | The date on which the metadata record was made available on the Open Government Portal | M | S | Free-text :  W3C Date and Time Formats (YYYY-MM-DD) | W3C Date no less than YYYY-MM-DD up to a fully qualified Date time |
| **23** | **Homepage URL** | A program or homepage that can be navigated to gain additional or contextual information about the dataset / information asset. | O | S | Free-text (URL) | This is not used to capture access to the resources, but rather background information. |
| **24** | **Related Items** | A related dataset or information asset in the Open Government Portal. | O | R | Free-text (URL) | Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system. Therefore, this field will be implemented by adding the unique i.d. of the related record in the Open Government Portal. |
| **25** | **Jurisdiction** | The extent or range of judicial, law enforcement, or other authority. For example, the level of government has contributed the dataset or open information asset. | M | S | Controlled Vocabulary  OGS Defined (i.e. Federal, Provincial, Municipal) | This field will be automatically populated with 'Federal' for all records provided by the Government of Canada. Then, the appropriate level of government will be selected for those records that are contributed from other portals, via federation. |
| **26** | **Licence** | The licence applied to the dataset or information asset. It may outline the limitations or constraints on the use of the item. | M | S | Set text (for all GC resources) Open Government Licence - Canada | This element will be automatically populated with 'Open Government Licence - Canada'. When other collections are added from provinces and municipalities, this field will include a controlled list of all versions of the OGL |
| **27** | **Distribution (Resource) Metadata** | **Unique Identifier** | A unique phrase or string which identifies the resource. | O | S | Free text (CharacterString) |  |
| **28** | **Title English** | An English name given to the resource. | M | S | Free text (CharacterString) | Free text. Enter the title applied to the resource. Refine using file formats, the name of the service, or year of coverage, and other additional terms as required. |
| **29** | **Title French** | A French name given to the resource. | M | S | Free text (CharacterString) | Free text. Enter the title applied to the resource. Refine using file formats, the name of the service, or year of coverage, and other additional terms as required. |
| **30** | **Date Published** | Date of formal issuance (e.g., publication) of the resource. | O | S | Free-text :  W3C Date and Time Formats (YYYY-MM-DD) |  |
| **31** | **Content Type** | The nature or genre of the content of the resource. (For example; dataset, supporting documentation, data dictionary, API or App). | M | S | Controlled Vocabulary | Select from the drop down list, controlled vocabulary. Note the full 'Content Type' controlled vocabulary is currently being developed by the GC Web Renewal team. |
| **32** | **Format** | The file format, physical medium, or dimensions of the resource. | M | S | Controlled Vocabulary | Select from the drop down list, controlled vocabulary. The controlled vocabulary has been developed for the Open Government Portal. Additions can be made by sending a request to open-ouvert@tbs-sct.gc.ca. |
| **33** | **Size** | The [estimated] size of a distribution in bytes. | O | S |  | Enter the size of the file in bytes. |
| **34** | **Language** | The language of the resource. | M | S | Controlled vocabulary : ISO 639-2/T | Select from the drop down list, controlled vocabulary. |
| **35** | **Download URL** | The location for online access to the distribution of the resource | M | S | Free-text (URL) | Enter the electronic location where the resource can be found. Note, this element should be populated with the URL that provides direct access, and not a homepage or interface to access it. Departments are also encouraged to use persistent identifiers for the URLs. |

# Appendix B

# Open Government Metadata Mapping

Below is a metadata mapping of the Final Open Government Common Core Metadata Element Set, to the existing schemas for collections in CKAN. The elements outside of the common core will be added to the domain specific metadata extensions.

|  |
| --- |
| **Mandatory Elements** |
| **Optional** |
| **Defaulted Text** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Common Core** | **Open (Publications) Information** | **Open Data-Non-Spatial (RAW)** | **Open Data-Geospatial** | **Open Maps** |
| **Portal Type** | Default : Open Information | Default: Open Data | Default: Open Data | Default: Open Data |
| **Collection Type** | Default: Open Publications | Default:Non-Spatial Data | Default: Geospatial Data | Default: Open Maps |
| **Metadata Scheme** | Open Government Metadata Element Set : Open Publications <linkto:tocome> | Open Government Metadata Element Set : Open Non -Spatial Data <linkto:tocome> | Open Government Metadata Element Set : Open Geospatial <linkto:tocome> | Open Government Metadata Element Set : Open Maps <linkto:tocome> |
| **Metadata Scheme Version** | 1.2 | 1.2 | 1.2 | 1.2 |
| **Metadata Record Identifier** | Metadata Identifier | Identifier | Identifier | Metadata URI |
| **Title English** | Title | Title English | Title English | Title English |
| **Title French** | Title | Title French | Title French | Title French |
| **Publisher** | GC Department or Agency | Publisher | Publisher | Organization Name (english and french) |
| **Creator** | Author |  |  | Individual Name |
| **Contact Email** | Default: open-ouvert@tbs-sct.gc.ca | Default: open-ouvert@tbs-sct.gc.ca | Default: Geoginfo@NRCan-RNCan.gc.ca | Default: Geoginfo@NRCan-RNCan.gc.ca |
| **Description English** | Description English | Description English | Description English | Abstract English |
| **Description French** | Description French | Description French | Description French | Abstract French |
| **Keywords English** | Keywords | Tags | Tags | Descriptive Keywords EN |
| **Keywords French** | Keywords | Tags | Tags | Descriptive Keywords FR |
| **Subject** | Subject | Subject | Topic category  (mapped to GC Core subject) | Topic category  (mapped to GC Core subject) |
| **Spatial** |  | Spatial | Spatial | Geographic Bounding Box |
| **Geographic Region Name** | Geographic Region | Geographic Region | Geographic Region |  |
| **Temporal** | Time Period Coverage Start / Time Period Coverage End | Time Period Coverage Start / Time Period Coverage End | Time Period Coverage Start / Time Period Coverage End | Temporal element |
| **Frequency** | Frequency of Serial | Frequency | Frequency | Maintenance and update frequency |
| **Date Published** | Date Published | Date Published | Date Published | Date Published |
| **Date Modified** | Date Modified | Date Modified | Date Modified | Date Modified |
| **Date Released** | Date Contributed | Portal Release Date | Portal Release Date | Date Contributed |
| **Homepage URL** | Homepage URL | Homepage URL | Homepage URL |  |
| **Related Items** | Other Language Edition |  |  | Aggregate dataset identifier |
| **Jurisdiction** | Default : Federal | Default : Federal | Default : Federal | Default : Federal |
| **Licence** | Licence  Default : Open Government Licence | Licence  Default : Open Government Licence | Licence  Default : Open Government Licence | Licence  Default : Open Government Licence |
| **Unique Identifier** |  |  |  | Identifier |
| **Resource - Title English** | Resource-Title English | Resource-Title English | Resource-Title English | Resource name english |
| **Resource - Title French** | Resource-Title French | Resource-Title French | Resource-Title French | Resource name french |
| **Date Published** |  |  |  |  |
| **Content Type** | File Type | Content Type | Content Type | Content Type |
| **Format** | Format | Format | Format | Format |
| **Size** | Size | Size | Size |  |
| **Language** | Language | Language | Language | Language |
| **Download URL** | Access URL / Info | URL | URL | Access URL |

Additional ‘domain specific elements’

|  |  |  |  |
| --- | --- | --- | --- |
| **Open (Publications) Information** | **Open Data-Non-Spatial (RAW)** | **Open Data-Geospatial** | **Open Maps** |
| Series Title | Data Series Name | Spatial Representation Type | Spatial Representation Type |
| Series Number | Issue Identification | Presentation Form - TO DELETE | Short Key |
| DOI | Digital Object Identifier | Browse Graphic File Name - TO DELETE | Metadata Record Language |
| Date Created |  |  | Character Set |
| ISBN |  |  | Parent Identifier |
| ISSN |  |  | Hierarchy Level |
| GC Catalogue Number |  |  | Metadata Contact English |
| Departmental Catalogue Number |  |  | Metadata Contact French |
| Weekly Checklist Number |  |  | Metadata Record Stamp |
| File Code |  |  | Locale |
| GC Docs Number |  |  | Contact Information English |
| LAC Identifier |  |  | Contact Information French |
| Amicus Identifier |  |  | Role |
| PWGSC Identifier |  |  | Status |
| Number of pages |  |  | Association Type |
| Former Frequency of Serial |  |  | Aggregate Dataset Identifier |
| Numeric and Chronological Designation |  |  | Reference System Information |
| File Type |  |  | Distributor Englih |
|  |  |  | Distributor French |

# Appendix D

DCAT and Schema.org Implementation Project Plan

See : [GCConnex](http://gcconnex.gc.ca/file/view/12376747/open-government-common-core-metadata-element-set)  and [Github](https://github.com/open-data/open-data)