**Project Plan: Applying DCAT and Schema.org to Open.Canada.ca**

Open Government Secretariat

Treasury Board of Canada Secretariat

Government of Canada

July 22, 2015

Contents

[1. Introduction 3](#_Toc425332322)

[2. DCAT 3](#_Toc425332323)

[3. Schema.org 6](#_Toc425332324)

[4. Implementation 7](#_Toc425332325)

[4.1. Timelines 8](#_Toc425332326)

[5. Metadata Mapping 9](#_Toc425332327)

[5.1. Catalog - dcat:Catalog 9](#_Toc425332328)

[5.2. Catalogue Record - dcat:CatalogRecord 11](#_Toc425332329)

[5.3. Dataset - dcat:Dataset 12](#_Toc425332330)

[5.4. Distribution - dcat:Distribution 14](#_Toc425332331)

[6. Outstanding Questions 15](#_Toc425332332)

# Introduction

During the consultation on the Open Government Common Core Metadata Element Set, it was suggested that the Government of Canada explore the possibility of applying RDF and HTML vocabularies to the metadata used to describe open government resources on open.canada.ca. We have taken this suggestion, and have developed a plan to implement both DCAT and Schema.org, to the common core element set. This document outlines our proposed approach.

The objective of applying DCAT and Schema.org elements to the Open Government Common Core Metadata Element Set is to improve search engine optimization and increase the discoverability of resources. It will also make metadata machine readable, which increases its use, and allows applications to easily consume, parse and interpret it. Since common vocabularies are key to communication, and make metadata more valuable, applying RDF and HTML will enable open.canada.ca to support federation. Finally, as DCAT is implemented by both Data.gov and data.gov.uk, open.canada.ca would align with G8 Best Practices.

# DCAT

Data Catalog Vocabulary (DCAT) is an RDF vocabulary, closely based on Dublin Core, which is applied to open data portals, such as data.gov and data.gov.uk, and datasets within. The [original DCAT vocabulary](http://vocab.deri.ie/dcat) was developed at DERI, refined by the [eGov Interest Group](http://www.w3.org/egov/), and then finally standardized by the [Government Linked Data (GLD)](http://www.w3.org/2011/gld/) Working Group (<http://www.w3.org/TR/vocab-dcat/>) .

RDF (Resource Description Framework) is used as a general method for the description of information embedded in web resources, as well as a standard for data exchange on the web. RDF has a linking structure that applies URIs (Unique Resource Identifiers) to define the relationship between items. This approach allows data to be mixed, shareable and exposed throughout the web, and ultimately improves the search and discovery of resources. DCAT specifically identifies a standard way to publish machine readable metadata about a datasets and specifies properties to add to existing metadata element sets in order to support the interoperability between open data portals throughout the web. These properties define a common set of elements, in six main classes, as well as the appropriate RDF property to apply. CKAN will conform to DCAT if it has RDF description of the catalogue itself and its datasets and distributions, and it is expressed using appropriate classes and properties from DCAT (except where no class exists). A scheme will still be DCAT compliant if it includes additional non DCAT metadata fields (i.e. Open Government Common Core Metadata Element Set Specific).

**The six main classes outlined in DCAT are:**

1. **Catalog :** [**dcat:Catalog**](http://www.w3.org/ns/dcat#Catalog)
   * A catalog is a curated collection of metadata about (datasets). Typically, a web based data catalog is represented a single instance of this class
   * <http://www.w3.org/TR/2013/WD-vocab-dcat-20130312/#class-catalog>
2. **Catalogue Record :** [**dcat:CatalogRecord**](http://www.w3.org/ns/dcat#CatalogRecord)
   * This class is optional and not all catalogs will use it. It exists for catalogs where a distinction is made between metadata about a dataset and metadata about the dataset's entry in the catalog.
   * <http://www.w3.org/TR/2013/WD-vocab-dcat-20130312/#class-catalog-record>
3. **Dataset :** [**dcat:Dataset**](http://www.w3.org/ns/dcat#Dataset)
   * This class represents the actual dataset as published by the dataset publisher.
   * <http://www.w3.org/TR/2013/WD-vocab-dcat-20130312/#class-dataset>
4. **Distribution :**  [**dcat:Distribution**](http://www.w3.org/ns/dcat#Distribution)
   * This class represents an accessible form of a dataset as for example a downloadable file, an RSS feed or a web service that provides the data. Note, this represents a general availability of a dataset it implies no information about the actual access method of the data, i.e. whether it is a direct download, API, or some through Web page.
   * <http://www.w3.org/TR/2013/WD-vocab-dcat-20130312/#class-distribution>
5. [**Concept**](http://www.w3.org/TR/2013/WD-vocab-dcat-20130312/#class-distribution) **and Concept Scheme -** [**skos:ConceptScheme**](http://www.w3.org/2004/02/skos/core#ConceptScheme)**,** [**skos:Concept**](http://www.w3.org/2004/02/skos/core#Concept)
   * It is recommended to use either skos:inScheme or skos:topConceptOf on every skos:Concept used to classify datasets to link it to the concept scheme it belongs to. This concept scheme is typically associated with the catalog using dcat:themeTaxonomy
   * Concept Scheme is the the knowledge organization system (KOS) used to represent themes/categories of datasets in the catalog.
   * <http://www.w3.org/TR/2013/WD-vocab-dcat-20130312/#class-category-and-category-scheme>
6. **Organization / Person :** [**foaf:/Person**](http://xmlns.com/foaf/0.1/Person)**,** [**foaf:Organization**](http://xmlns.com/foaf/0.1/Organization)
   * [foaf:Person](http://xmlns.com/foaf/0.1/Person) for people and [foaf:Organization](http://xmlns.com/foaf/0.1/Organization) for government agencies or other entities.
   * <http://www.w3.org/TR/2013/WD-vocab-dcat-20130312/#class-organization-person>

# Schema.org

Schema.org, developed by Bing, Yahoo and Google, aims to create and support a common set of schemas, or HTML tags, for structured data mark-up on web page in order to be recognized by major search providers. These HTML tags enable the marking-up of web pages with metadata about itself. Search engine spiders can then better understand a web page, and therefore provide richer results by helping to surface content clearly and prominently in search results. Schema.org provides a single place for users to find the appropriate tags to apply, and essentially how to mark up their pages with reasonable syntax and style consistency across types, rather than applying various HTML vocabularies.

The [Government of Canada’s Standard on Web Interoperability](http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=25875&section=text) was developed with the objective of ensuring the re-use, and portability of Government of Canada Web content across websites and Web applications, platforms, and devices, as well as to ensure that web content can be found, retrieved, and exchanged. It also outlines requirements for applying HTML data, used to embed and extract machine-readable content in Web pages, and assists in the retrieval, extraction and re-use of content by web browsers and search engines. It specifically indicates that the HTML vocabulary Schema.org must be applied. It states:

Departmental Web Managers, Functional Specialists, and equivalents are responsible for:

1. Ensuring that HTML data is applied to Web pages to facilitate retrieval and re-use of content, as follows:

1.1 HTML data syntax: RDFa 1.1 Lite or later

1.2 Primary HTML data vocabulary: Schema.org

Therefore, in order to implement DCAT for the Open Government Common Core Metadata Element Set, we must also apply HTML tags specified by Schema.org.

Schema.org has various schemas, which are a set of ‘Types’, with associated properties or elements for each, and are arranged in a hierarchy. For a full list of all of these types, visit : http://schema.org/docs/full.html . The ‘types’ that are applicable to the Open Government Portal include : [DataCatalog](http://schema.org/DataCatalog) (a collection of datasets), [Dataset](http://schema.org/Dataset) (a body of structured information describing some topic(s) of interest) and [DataDownload](http://schema.org/DataDownload) (a dataset in downloadable form) which correspond to Catalog, Dataset and Distribution in DCAT. See the mapping below in section 5 to see how these properties align with the Open Government Common Core Metadata Element Set.

# Implementation

By leveraging best practices from [data.gov](http://data.gov) (see: <https://project-open-data.cio.gov/v1.1/schema/>) and [data.gov.uk](http://data.gov.uk) (see <http://datagovuk.github.io/guidance/dcat_fields.html> ), The Government of Canada proposes to apply DCAT and Schema.org to open.canada.ca, by completing the following steps :

 1. Conduct metadata mapping of OGCCMES and DCAT and Schema.org - completed (see: Section 5)

* Note, we will limit the application of DCAT and Schema.org to metadata records describing datasets only, and not open information resources. The field ‘Portal Type’ will identify which records belong to the Open Data Portal, rather than Open Information Portal.

2. Consult on mapping (July 23-Aug4), and project plan

* Gather feedback and make updates.

 3. Apply mapping to an RDF template

* Develop XML RDF file to use as a template
* A phased approach may be used, as we may add basic, common fields first and update overtime. For example, we may add [dcat:dataset](http://www.w3.org/ns/dcat#dataset) and dcat:Distribution tags first , then add [dcat:Catalog](http://www.w3.org/ns/dcat#CatalogRecord) and [dcat:CatalogRecord](http://www.w3.org/ns/dcat#CatalogRecord)

4. Test RDF template

* Validate file

5. Host RDF file externally

* This will avoid potential prefix / namespace collisions,
* The RDF will be at a different URL than the dataset pages. The dataset pages are language-specific and include the [schema.org](http://schema.org) markup for search. A single RDF page for all languages of each dataset will be developed

6. Apply Schema.org to metadata record template on open.canada.ca

* The HTML markup will be applied to the CKAN record

 7. Deployment

* The new pages will be added to open.canada.ca as part of regular updates to the new open data site based on CKAN 2.3

 8. Evaluation / Revisions

* We will continue to update and revise the application of DCAT and Schema.org overtime,
* We will also continue to evaluated other RDF and HTML Vocabularies to apply to metadata extensions to the Open Government Common Core

## Timelines

|  |  |  |
| --- | --- | --- |
| Task | Responsibility | Deadline |
| 1. Finalize the Open Government Common Core Metadata Element Set (OGCCMES) | TBS | Thursday July 16 |
| 2. Develop mapping to OGCCMES, DCAT and Schema.org (see section 5) | TBS and StatsCan | Monday July 20 |
| 3. Share mapping, project plan, and consult on approach and proposed implementation | TBS, StatsCan, and Stakeholders | Tuesday July 21-August 4 |
| 4. Gather feedback, make revisions, complete mapping | TBS | Thursday August 6 |
| 5. Submit final mapping for implementation | TBS | Tuesday August 12 |
| 6. Implementation | StatsCan | August / September |
| 7. Documentation | TBS and StatsCan | August / September |
| 8. Deployment | StatsCan | October |
| 9. Evaluation / Revision | TBS, StatsCan, Stakeholders | Ongoing |

# Metadata Mapping

Below outlines how the GC proposes to implement DCAT by apply relevant elements to the Open Government Common Core Metadata Element Set. The mapping below also captures the HTML tags to apply from Schema.org. Note, we are unsure how to implement ‘Concept’, ‘Concept Scheme’ and ‘Organizations and Persons’.

## Catalog - [dcat:Catalog](http://www.w3.org/ns/dcat#Catalog)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Property** | **RDF Property** | **Definition** | **Range** | **Mapping to OG CC MES** | **Schema.org** |
| **title** | [dct:title](http://purl.org/dc/terms/title) | A name given to the catalog | [rdfs:Literal](http://www.w3.org/2000/01/rdf-schema#Literal) | **Hard Coded :** Government of Canada Open Government Portal |  |
| **description** | [dct:description](http://purl.org/dc/terms/description) | A free-text account of the catalog | [rdfs:Literal](http://www.w3.org/2000/01/rdf-schema#Literal) | **Hard Coded :** This catalog contains metadata records describing the … |  |
| **release date** | [dct:issued](http://purl.org/dc/terms/issued) | Date of formal issuance (publication) of the catalog | rdfs:Literal encoded using the relevant ISO 8601 Date and Time compliant string and typed using the appropriate XML Schema datatype [xmlschema-2] | **Hard Coded :** 2013-06-19 |  |
| **update / modification date** | [dct:modified](http://purl.org/dc/terms/modified) | Most recent date on which the catalog was changed, updated or modified | rdfs:Literal encoded using the relevant ISO 8601 Date and Time compliant string and typed using the appropriate XML Schema datatype [xmlschema-2] | **Hard Coded :** (date we update to CKAN 2.3?) |  |
| **language** | [dct:language](http://purl.org/dc/terms/language) | The language of the catalog, this refers to the language used in the textual metadata describing title, description..etc. | dct:LinguisticSystem  Resources defined by the Library of Congress (1, 2) SHOULD be used.  If a ISO 639-1 (two-letter) code is defined for language, then its corresponding IRI SHOULD be used; if no ISO 639-1 code is defined, then IRI corresponding to the ISO 639-2 (three-letter) code SHOULD be used. | **Hard Coded :** English / French |  |
| **homepage** | [foaf:homepage](http://xmlns.com/foaf/0.1/homepage) | The homepage of the catalog. | foaf:homepage is an inverse functional property (IFP) which means that it should be unique and precisely identify the catalog. This allows smushing various descriptions of the catalog when different URIs are used. | **Hard Coded:** open.canada.ca |  |
| **publisher** | [dct:publisher](http://purl.org/dc/terms/publisher) | The entity responsible for making the catalog online. | [Resources of type foaf:Agent are recommended as values for this property.](http://xmlns.com/foaf/0.1/Agent) | **Hard Coded:** Government of Canada, Treasury Board of Canada Secretariat |  |
| **spatial/**  **geographic** | [dct:spatial](http://purl.org/dc/terms/spatial) | spatial/geographic: The geographical area covered by the catalog. | [dct:Location](http://purl.org/dc/terms/Location) | **Hard Coded:** Bounding box for Canada, or Ottawa ? |  |
| **themes** | [dcat:themeTaxonomy](http://www.w3.org/ns/dcat#themeTaxonomy) | The knowledge organization system (KOS) used to classify catalog's datasets. | [skos:ConceptScheme](http://www.w3.org/2004/02/skos/core#ConceptScheme) | **Hard Coded:**  Information and Communications  Government and Politics |  |
| **license** | [dct:license](http://purl.org/dc/terms/license) | This links to the license document under which the catalog is made available and not the datasets. Even if the license of the catalog applies to all of its datasets and distributions, it should be replicated on each distribution. | [dct:LicenseDocument](http://purl.org/dc/terms/LicenseDocument) | **Hard Coded :** Open Government Licence - Canada <linkto:http://open.canada.ca/en/open-government-licence-canada> |  |
| **rights** | [dcat:dataset](http://www.w3.org/ns/dcat#dataset) | This describes the rights under which the catalog can be used/ reused and not the datasets. Even if theses rights apply to all the catalog datasets and distributions, it should be replicated on each distribution. | [dct:LicenseDocument](http://purl.org/dc/terms/LicenseDocument) | n/a |  |
| **dataset** | [dcat:dataset](http://www.w3.org/ns/dcat#dataset) | A dataset that is part of the catalog. | [dcat:Dataset](http://www.w3.org/ns/dcat#Dataset) | Describe dcat:dataset elements |  |
| **catalog record** | [dcat:record](http://www.w3.org/ns/dcat#record) | A catalog record that is part of the catalog. | [dcat:CatalogRecord](http://www.w3.org/ns/dcat#CatalogRecord) | Describe to dcat:record elements |  |

## Catalogue Record - [dcat:CatalogRecord](http://www.w3.org/ns/dcat#CatalogRecord)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Property** | **RDF Property** | **Definition** | **Range** | **DCAT : Usage Note** | **Mapping to OG CC MES** | **Schema.org Mapping** |
| **title** | [dct:title](http://purl.org/dc/terms/title) | A name given to the record. | [rdfs:Literal](http://www.w3.org/2000/01/rdf-schema#Literal) |  | **Hard Coded :** Open Data Portal / Open Information Portal | [Thing>name](http://schema.org/name) |
| **description** | [dct:description](http://purl.org/dc/terms/description) | free-text account of the record. | [rdfs:Literal](http://www.w3.org/2000/01/rdf-schema#Literal) |  | **Hard Coded :**  The Open Data Portal provides access…. | [Thing >description](http://schema.org/description) |
| **listing date** | [dct:issued](http://purl.org/dc/terms/issued) | The date of listing the corresponding dataset in the catalog. | rdfs:Literal encoded using the relevant ISO 8601 Date and Time compliant string and typed using the appropriate XML Schema datatype [xmlschema-2] | This indicates the date of listing the dataset in the catalog and not the publication date of the dataset itself. | Date Released |  |
| **update / modification date** | [dct:modified](http://purl.org/dc/terms/modified) | update/modification date : Most recent date on which the catalog entry was changed, updated or modified. | rdfs:Literal encoded using the relevant ISO 8601 Date and Time compliant string and typed using the appropriate XML Schema datatype [xmlschema-2] | This indicates the date of last change of a catalog entry, i.e. the catalog metadata description of the dataset, and not the date of the dataset itself. | **Hard Coded :**  Date of CKAN 2.3 launch | [CreativeWork> dateModified](http://schema.org/dateModified) |
| **primary topic** | [foaf:primaryTopic](http://xmlns.com/foaf/0.1/primaryTopic) | Links the catalog record to the dcat:Dataset resource described in the record. |  | foaf:primaryTopic property is functional: each catalog record can have at most one primary topic i.e. describes one dataset. | Subject |  |

## Dataset - [dcat:Dataset](http://www.w3.org/ns/dcat#Dataset)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Property** | **RDF Property** | **Definition** | **Range** | **DCAT : Usage Note** | **Mapping to OG CC MES** | **Schema.org Mapping** |
| **title** | [dct:title](http://purl.org/dc/terms/title) | A name given to the dataset. | [rdfs:Literal](http://www.w3.org/2000/01/rdf-schema#Literal) |  | Title English /  Title French | [Thing>name](http://schema.org/name) |
| **description** | [dct:description](http://purl.org/dc/terms/description) | free-text account of the dataset. | [rdfs:Literal](http://www.w3.org/2000/01/rdf-schema#Literal) |  | Description English / Description French | [Thing >description](http://schema.org/description) |
| **release date** | [dct:issued](http://purl.org/dc/terms/issued) | Date of formal issuance (e.g., publication) of the dataset. | rdfs:Literal encoded using the relevant ISO 8601 Date and Time compliant string and typed using the appropriate XML Schema datatype [xmlschema-2] | This property should be set using the first known date of issuance. | Date Published | [CreativeWork>datePublished](http://schema.org/datePublished) |
| **update / modification date** | [dct:modified](http://purl.org/dc/terms/modified) | Most recent date on which the dataset was changed, updated or modified. | rdfs:Literal encoded using the relevant ISO 8601 Date and Time compliant string and typed using the appropriate XML Schema datatype [xmlschema-2] | The value of this property indicates a change to the actual dataset, not a change to the catalog record. | Date Modified | [CreativeWork>dateModified](http://schema.org/dateModified) |
| **publisher** | [dct:publisher](http://purl.org/dc/terms/publisher) | An entity responsible for making the dataset available. |  | [See also : Organization / Person](http://www.w3.org/TR/vocab-dcat/#Class:_Organization.2FPerson) | Publisher | [CreativeWork> publisher](http://schema.org/publisher) |
| **frequency** | [dct:accrualPeriodicity](http://purl.org/dc/terms/accrualPeriodicity) | The frequency at which dataset is published. | dct:Frequency (A rate at which something recurs) |  | Frequency |  |
| **identifier** | [dct:identifier](http://purl.org/dc/terms/identifier) | A unique identifier of the dataset |  | The identifier might be used as part of the URI of the dataset, but still having it represented explicitly is useful. | Metadata Record Identifier |  |
| **spatial/**  **geographical coverage** | [dct:spatial](http://purl.org/dc/terms/spatial) | Spatial coverage of the dataset. | dct:Location (A spatial region or named place) |  | Spatial | [Dataset> spatial](http://schema.org/spatial) |
| **language** | [dct:language](http://purl.org/dc/terms/language) | The language of the dataset. | dct:LinguisticSystem  Resources defined by the Library of Congress (1, 2) SHOULD be used.  If a ISO 639-1 (two-letter) code is defined for language, then its corresponding IRI SHOULD be used; if no ISO 639-1 code is defined, then IRI corresponding to the ISO 639-2 (three-letter) code SHOULD be used. | • If the dataset is available in multiple languages, use multiple values for this property. If each language is available separately, define an instance of dcat:Distribution for each language and describe the specific language of each distribution using dct:language | Language | [CreativeWork> inLanguage](http://schema.org/inLanguage) |
| **temporal coverage** | [dct:temporal](http://purl.org/dc/terms/temporal) | The temporal period that the dataset covers. | dct:PeriodOfTime <linkto:http://purl.org/dc/terms/PeriodOfTime (An interval of time that is named or defined by its start and end dates) |  | Temporal | [Dataset> datasetTimeInterval](http://schema.org/datasetTimeInterval) |
| **theme/ category** | [dcat:theme](http://www.w3.org/ns/dcat#theme) | The main category of the dataset. A dataset can have multiple themes. | [skos:Concept](http://www.w3.org/2004/02/skos/core#Concept) | The set of skos:Concepts used to categorize the datasets are organized in a skos:ConceptScheme describing all the categories and their relations in the catalog. | Subject | [CreativeWork>genre](http://schema.org/genre) |
| **keyword/tag** | [dcat:keyword](http://www.w3.org/ns/dcat#keyword) | A keyword or tag describing the dataset. | [rdfs:Literal](http://www.w3.org/2000/01/rdf-schema#Literal) |  | Keywords English / Keyword French | [CreativeWork>keywords](http://schema.org/keywords) |
| **contact point** | [dcat:contactPoint](http://www.w3.org/ns/dcat#contactPoint) | Link a dataset to relevant contact information which is provided using VCard [vcard-rdf]. | [vcard:Kind](http://www.w3.org/2006/vcard/ns#Kind) |  | Contact Email |  |
| **dataset distribution** | [dcat:distribution](http://www.w3.org/ns/dcat#distribution) | Connects a dataset to its available distributions. | [dcat:Distribution](http://www.w3.org/ns/dcat#Distribution) |  | Describe dcat:distribution elements |  |
| **landing page** | [dcat:landingPage](http://www.w3.org/ns/dcat#landingPage) | A Web page that can be navigated to in a Web browser to gain access to the dataset, its distributions and/or additional information. | [foaf:Document](http://xmlns.com/foaf/0.1/Document) |  | Homepage URL | [Thing>sameAs](http://schema.org/sameAs) |

## Distribution - [dcat:Distribution](http://www.w3.org/ns/dcat#Distribution)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Property** | **RDF Property** | **Definition** | **Range** | **DCAT : Usage Note** | **Mapping to OG CC MES** | **Schema.org Mapping** |
| **title** | [dct:title](http://purl.org/dc/terms/title) | A name given to the distribution. | [rdfs:Literal](http://www.w3.org/2000/01/rdf-schema#Literal) |  | Title English / Title French |  |
| **description** | [dct:description](http://purl.org/dc/terms/description) | free-text account of the distribution. | [rdfs:Literal](http://www.w3.org/2000/01/rdf-schema#Literal) |  | n/a |  |
| **release date** | [dct:issued](http://purl.org/dc/terms/issued) | Date of formal issuance (e.g., publication) of the distribution. | This property should be set using the first known date of issuance. |  | Date Published |  |
| **update/**  **modification date** | [dct:modified](http://purl.org/dc/terms/modified) | Most recent date on which the distribution was changed, updated or modified. | rdfs:Literal encoded using the relevant ISO 8601 Date and Time compliant string and typed using the appropriate XML Schema datatype [xmlschema-2] |  | n/a |  |
| **license** | [dct:license](http://purl.org/dc/terms/license) | This links to the license document under which the distribution is made available. | [dct:LicenseDocument](http://purl.org/dc/terms/LicenseDocument) |  | Licence | [CreativeWork> license](http://schema.org/license) |
| **rights** | [dct:rights](http://purl.org/dc/terms/rights) | Information about rights held in and over the distribution. |  | dct:license, which is a sub-property of dct:rights, can be used to link a distribution to a license document. However, dct:rights allows linking to a rights statement that can include licensing information as well as other information that supplements the licence such as attribution. | n/a |  |
| **download URL** | [dcat:downloadURL](http://www.w3.org/ns/dcat#downloadURL) | A file that contains the distribution of the dataset in a given format | [rdfs:Resource](http://www.w3.org/2000/01/rdf-schema#Resource) |  | Download URL | [Dataset>distribution](http://schema.org/distribution) |
| **byteSize** | [dcat:byteSize](http://www.w3.org/ns/dcat#size) | The size of a distribution in bytes. | rdfs:Literal typed as xsd:decimal. |  | Size |  |
| **media type** | [dcat:mediaType](http://www.w3.org/ns/dcat#mediaType) | The media type of the distribution as defined by IANA. | [dct:MediaTypeOrExtent](http://purl.org/dc/terms/MediaTypeOrExtent) |  | Content Type |  |
| **format** | [dct:format](http://purl.org/dc/terms/format) | The file format of the distribution. | [dct:MediaTypeOrExtent](http://purl.org/dc/terms/MediaTypeOrExtent) |  | Format |  |

# Outstanding Questions

How can DCAT and Schema.org both be applied, without any collisions?

Are there potential collisions when marking up data using both RDF and HTML?

When there are conflicting namespaces, and prefixes, which one should be applied?

Can an external DCAT file with RDF mark-up be applied to avoid collisions?

Are there risks associated with storing RDF externally?

How can DCAT and Schema.org be applied in CKAN?

Best practices from Data.gov?

Is the proposed mapping complete / accurate?