

Odesi Metadata Best Practices Guide

Version 4.0.1

Based on Dataverse 5.x and DDI 2.x (Codebook)



February 2024

Table of Contents

About this Guide	3
Content Overview	3
Acknowledgements	4
Changes in v. 4.0.1.	4
Using this Guide to Edit Metadata in Borealis	5
Sections of this Guide	5
Other Sections	6
Finding the Metadata Sections in Borealis	7
Metadata Sections in Borealis	10
Citation Metadata	10
Geospatial Metadata	23
Social Science and Humanities Metadata	25
Terms of Use Metadata	29
Terms of Access Metadata	31
File Description Metadata	33
Variable Metadata	36
Appendix A: Controlled Vocabularies	41
English Controlled Vocabularies	41
Vocabulaires contrôlés en français	45
Appendix B: Glossary	49
References	53

About this Guide

Content Overview

The Odesi Metadata Best Practices Guide (v.4.0.1) (BPG) contains detailed guidance and best practices for curating metadata in the [Odesi repository](#). These best practices are informed by library and archival practices and by the FAIR Guiding Principles for Data Management and Stewardship.¹

The BPG v. 4.0.1. introduces minor changes to v. 4.0, which was updated in October 2023 to reflect major changes to the Odesi platform since Odesi's back-end migration from Nesstar to [Borealis, the Canadian Dataverse Repository](#). The new BPG updates and supersedes the [Odesi Best Practices Document v. 3.1](#) (February 2019). The BPG also borrows largely from the [Dataverse North Metadata Best Practices Guide v. 3.0](#) (November 2021).

The updated contents are intended to support both novice and experienced users creating metadata for Odesi data collections with the new Dataverse interface. With this purpose in mind, the document's structure has been overhauled and restructured with a new logic, to reflect how editors will encounter the metadata fields in practice. As experienced Odesi users will recall, previous versions of the guide were structured according to the five metadata sections in DDI Codebook (Document Description, Study Description, Data Files Description, Variable Description, and Other Documentation). In contrast, the present guide is structured according to the metadata sections of Dataverse 5.x and the Data Curation Tool, the new back-end for Odesi: Citation, Geospatial, Social Science & Humanities, Terms of Use, Terms of Access, File Description, Variables.

Despite all these changes, the Odesi platform and the new Best Practices Guide continue to support Odesi's DDI Codebook (2.x) metadata profile. Alongside metadata field descriptions users will find cross-references to the XML tag in Borealis' DDI metadata file.

Dataverse metadata is primarily based on the domain-agnostic DataCite² metadata schema, a set of core metadata properties designed to support citation and retrieval of digital objects. Dataverse also implements DataCite's [Digital Object Identifier](#) (DOI) standard by assigning a DOI to all datasets upon publication. Over time, Dataverse has sought to extend its interoperability by incorporating metadata standards from various domains, including DDI Codebook for the social, economic, behavioral, and health sciences.

To date, Borealis coverage of DDI-C metadata is not complete, and the new Odesi Best Practices have been adjusted for work in the Borealis and Data Curation Tool (DCT) interfaces

¹ Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

² "The DataCite Metadata Schema is a list of core metadata properties chosen for an accurate and consistent identification of a resource for citation and retrieval purposes, along with recommended use instructions" (<https://schema.datacite.org/>).

in their present versions. Best practices for variable-level metadata are described in this document, and the [Odesi Deposit Guide](#) describes in more detail the process of ingesting and creating variable-level metadata using the DCT.

Community efforts are underway to improve support of the DDI-C metadata schema in Dataverse, which is the collaborative output of an international, community-driven project to develop and maintain this tool as free, open source software. Like the metadata standards and controlled vocabularies applied in this guide, the software platforms that support Odesi are always being improved, and new versions or updates are to be expected. Whenever those changes affect Odesi Best Practices, Scholars Portal and the MarkIt! Program will update this and other guidance materials accordingly.

Acknowledgements

We would like to thank and acknowledge the Odesi Migration & Redevelopment Working Group, and the Odesi MarkIt! Program supervisors, for their time and support in creating this guide. With special thanks to our student Hannah Brown (University of Toronto) who created the first draft of this guide. We would like to thank everyone in the Odesi Training Materials Working Group for writing and editing this guide: Alexandra Cooper (Queen's University), Hafsah Hujaleh (University of Toronto), Amber Leahey (Scholars Portal), Guinsly Mondesir (Scholars Portal), Alicia Urquidi Diaz (Scholars Portal), and Odesi student Broghan White (University of Toronto). Project funding was provided by OCUL and Compute Ontario.

Changes in v. 4.0.1.

Controlled vocabularies. Added *Data Documentation – Methodology* to [Type of File](#) (English) and *Documentation des données – Méthodologie* to [Type de fichier](#) (French).

Using this Guide to Edit Metadata in Borealis

Sections of this Guide

This guide aims to present the content in a way that supports the active process of curating metadata using Borealis. As much as possible, the structure of this document reflects the way users will encounter the metadata fields in practice. In contrast with previous versions, the present guide is structured according to the metadata sections of Dataverse 5.x and the Data Curation Tool (DCT), the new back-end for Odesi. The new metadata sections are Citation, Geospatial, Social Science & Humanities, Terms of Use, Terms of Access, File Description, and Variables. Field relationships to the DDI-C schema is indicated in the XML Tag column, which shows the element tag (and attribute, if applicable) for each field.

The metadata sections in Borealis are summarized in Table 1. For each section, this guide presents the metadata in a table with five columns. The table format and columns are described below in Table 2.

Table 1. Metadata Sections in Borealis

Citation	The Citation Metadata block includes the core metadata needed to publish a dataset in a Borealis repository.
Geospatial	Metadata that describes the spatial extent (e.g., location information) of data in a dataset.
Social Science & Humanities	The Social Science and Humanities block contains metadata that describes survey data and related datasets in the social sciences and humanities.
Terms of Use	Metadata that describes the terms of use of the dataset.
Terms of Access	Metadata that describes the terms of access of the dataset.
File Description	This section consists of metadata items describing the characteristics and contents of file(s) that comprise the dataset.
Variables	This section contains a rich set of elements with detailed descriptive information about the variables in the study that produced the dataset.

Table 2. Structure of the Metadata Tables in this Guide.

Column	Description
Field	The label used for this field in the Dataverse interface.
Description	Field definition based on Odesi Best Practices. This may differ from the description in Borealis. This column can include tips and notes about the field's meaning or usage. Usage examples are provided in this field in italics. Any applicable controlled vocabularies are referenced as well.
Usage	Whether a field is required or optional. <ul style="list-style-type: none">• Required [R] - Fields have been designated by Odesi as required to enable data discoverability and reuse.• Optional [O] - Optional fields provide added information about a dataset. The recommended best practice is to enter optional metadata whenever fields are applicable and the information is available.
Repeatable	Whether a field allows multiple entries (repeatable, or R) or not (NR).
XML tag	This corresponds to the XML tag for this metadata field in the Borealis DDI metadata file. In this guide we follow DDI's typographic convention representing tags in angle brackets.

We provide an example of real metadata usage for the majority of fields used for Odesi datasets. These examples are drawn from dataset records in Odesi, including: [General Social Survey, Cycle 23, 2009 \[Canada\]: Victimization, Incident File](#); [General Social Survey, Cycle 34, 2019 \[Canada\]: Victimization, Main File](#) and [Canadian Tobacco Use Monitoring Survey, 2009: Cycle 1, Person File](#). The MarkIt! Program group has also created Borealis templates to create metadata for specific collections.

In using this guide, you will markup datasets using Borealis DDI metadata fields and Odesi's Best Practices that are described in more detail throughout this guide. Please note that, as long as the information is accurate, duplication of information is expected. For example, the **Author** of a dataset may also be the **Producer**. In other cases, fields' coverage may appear to overlap, as with **Funder** vs. **Grant agency**. Metadata information may be the same in multiple fields, and this kind of duplication works to create records that are as complete and thorough as possible.

Other Sections

[Appendix A: Controlled Vocabularies](#) defines and references the controlled vocabularies used for Odesi metadata, in English and in French. [Appendix B: Glossary](#) defines some of the most relevant metadata and markup-related terms used in this guide. The [References](#) section lists the works that were cited and consulted while writing this guide.

Finding the Metadata Sections in Borealis

Finding the right place to edit a metadata field in Borealis can be difficult at first, because metadata fields are distributed across different sections of the site.

One important thing to keep in mind at the beginning is that most metadata fields can't be seen or edited until after a new record has been created and saved.

Once a dataset has been created and saved, the record will display four tabs: **Files**, **Metadata**, **Terms** and **Versions** (see Figure 1a). You can choose a section using the tabs or, alternatively, you can use the drop-down menu on the top right-hand side of the record (Figure 1b) to select a metadata section to edit.



Figure 1a.

Use the tabs to choose which sections to edit: **Files**, **Metadata** or **Terms**.

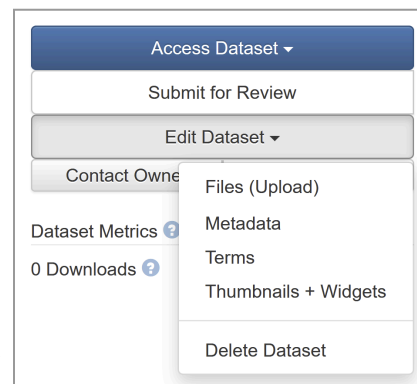


Figure 1b.

Alternatively, you can use the **Edit Dataset** drop-down to select a metadata section to edit.

Citation, Geospatial, and Social Science Metadata

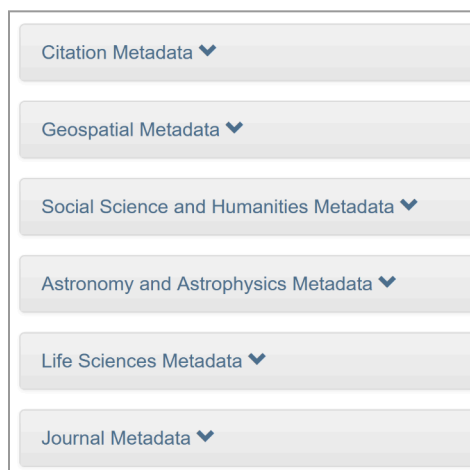


Figure 2.

The **Metadata** tab in Borealis.

The **Metadata** tab in Borealis holds the **Citation**, **Geospatial**, **Social Science and Humanities**, **Astronomy and Astrophysics**, and the **Life Sciences**, or fields to describe journal articles associated with a dataset.

Click on a section title to expand the view and edit the metadata fields in that section (see Figure 2).

File Description Metadata

File description metadata can be edited under the **Files** tab. Some file information is created at file upload, but it is possible to edit and add further metadata after the file has been saved in Borealis. The options to edit file metadata fields can be accessed through drop-down menus, as shown in Figure 3 and described below.

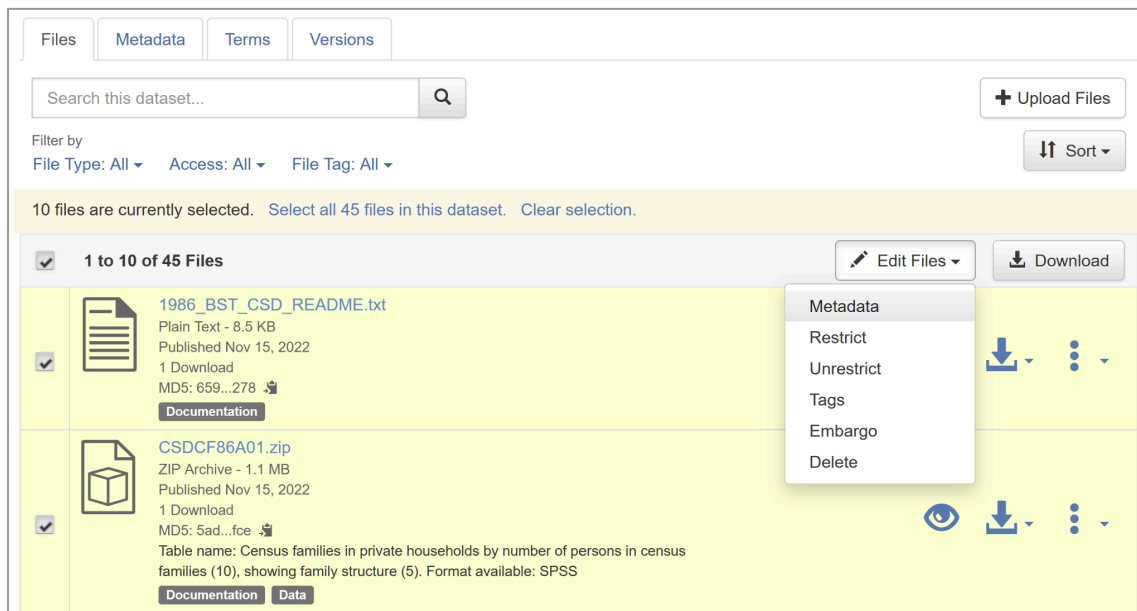


Figure 3. File metadata can be edited under the **Files** tab in Borealis.

To edit metadata for multiple files, use the checkboxes to select the files and click on the **Edit Files** button on the top right of the list (as shown on Figure 3). Metadata for individual files can be edited using the three blue dots. The blue dots are located on the right-hand side of the screen, next to each file item on the list.

Figure 4 shows the file editing interface, where **File Name**, **File Path** and **Description** can be edited. Further **File Options** to add **Tags** and **Provenance** information, can be displayed by clicking on the three blue dots on the right-hand side of the screen.

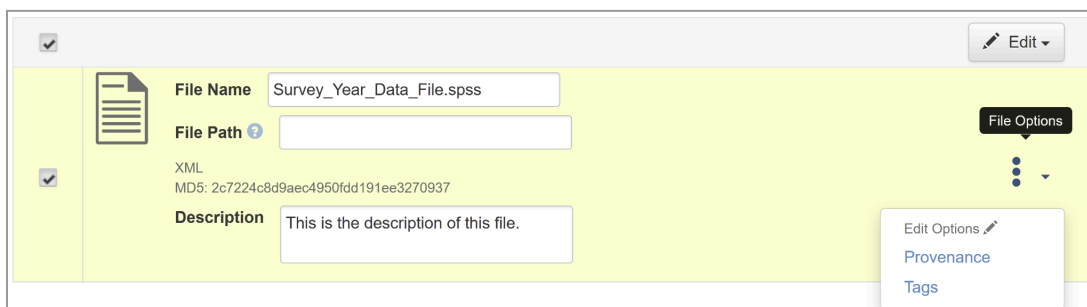


Figure 4. Editing metadata for a file.

Terms Metadata

Terms metadata can be edited on the **Terms** tab (see figure 1a). To view all editing options, click on the **Edit Terms Requirements** button on the top-right side of the tab. This will display three sections: **Dataset Terms**, **Restricted Files + Terms of Access**, and **Guestbook**.

By default, Borealis hides the Terms of Use metadata fields and sets the default dataset terms to CC0 1.0. To edit the Terms of Use metadata, select the option **Custom Dataset Terms** from the drop-down menu, as shown in figure 5. This will display all the relevant fields.

The **Terms of Access** metadata fields are found under the **Restricted Files + Terms of Access** section.

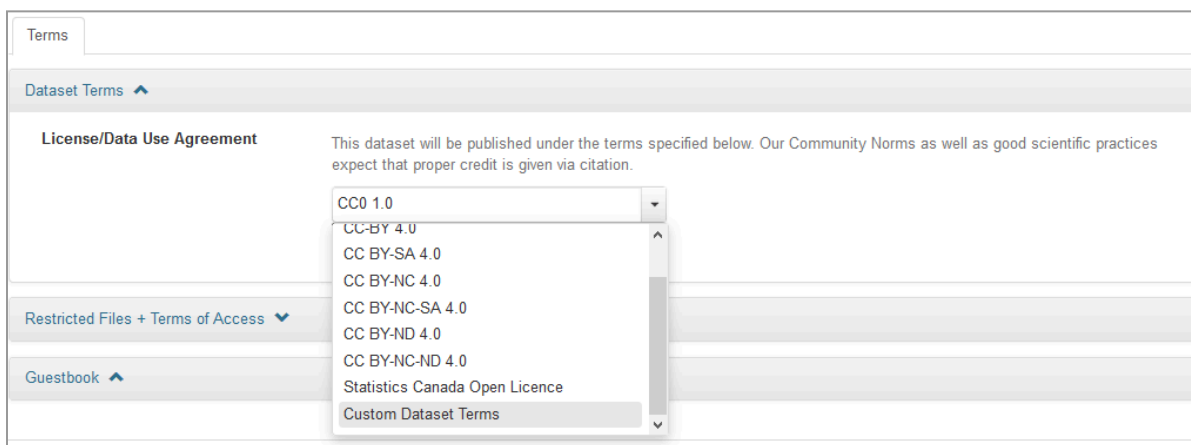


Figure 5.

To edit Terms of Use metadata, select the **Custom Dataset Terms** option from the drop-down menu.

Editing Variable Metadata

All variable metadata fields used in Odesi are described in the [Variable Metadata section](#) of this guide. System generated fields are also described, as some variable metadata fields are generated automatically during tabular file ingest.

In Borealis, all other variable metadata is entered and displayed using the Data Curation Tool. A separate guide, the [Odesi Deposit Guide](#) provides detailed information and instructions on using this tool to edit variable metadata.

Metadata Sections in Borealis

Citation Metadata

The Citation Metadata block includes the core metadata needed to publish a dataset in a Borealis repository. The required fields in this section are used to create the citation for the dataset.

Citation Metadata				
Field	Description	Usage	Repeatable	XML tag
Title	<p>Title of marked up document. A full title should indicate the geographic scope of the data collection as well as the time period covered.</p> <p><i>Example: General Social Survey, Cycle 23, 2009 [Canada]: Victimization, Incident File.</i></p>	R	NR	<titl>
Subtitle	<p>A subtitle is a secondary title used to amplify or state certain limitations of the main title.</p> <p>Tip: The subtitle is not included in the auto-generated citation. If you want the subtitle to be included in the citation, then you must add it to the Title field in addition to the Subtitle field.</p> <p><i>Example: Victimization, Incident File</i></p>	O	NR	<subTitl>

Citation Metadata				
Field	Description	Usage	Repeatable	XML tag
Alternative Title	<p>A title by which the work is commonly referred, a translation or an abbreviation of the title.</p> <p>Tip: Acronym, short form, or translation of full title. This field is repeatable. Whenever possible, include both the abbreviation and the translation (as separate Alternative Title entries).</p> <p><i>Example (abbreviation): GSS, Cycle 23, 2009: Victimization, Incident File</i></p> <p><i>Example (translation): Enquête Sociale Générale, Cycle 23, 2009 [Canada]: Victimisation, Dossier Principal.</i></p>	O	R	<altTitl>
Other ID	<p>Unique string or number uniquely assigned to this dataset. This can refer to the producer's or archive's number for the marked-up document, or to an ID generated by a centralized agency (DOI, handle, other).</p> <p>Note: Borealis assigns DOIs to all published records. For Statistics Canada surveys: The catalog number refers to the microdata file. If there is no catalog number available, use the Record Number.</p>			
Other ID (agency)	<p>Name of agency which generated this identifier.</p> <p><i>Example: DOI</i> <i>Example: Statistics Canada</i></p>	O	R	<citation> <titlStmnt> <IDNo> (agency)
Other ID (Identifier)	<p>The identifier that is uniquely assigned to this dataset.</p> <p>For Statistics Canada surveys: Use the format <i>Acronym-CatalogueNumberOrRecordNumber-language-year</i></p> <p><i>Example (DOI): doi:10.5683/SP2/ILLIXV</i> <i>Example (Statistics Canada survey): accs-3312-E-b2020</i></p>	O	R	<IDNo>

Citation Metadata				
Field	Description	Usage	Repeatable	XML tag
Author	<p>The person, corporate body, or agency responsible for the data collection's substantive and intellectual content. Repeat the element for each author, and use the affiliation field if available. Invert first and last name and use commas.</p> <p>Tip: The author entity is based on information from the dataset's User's Guide (usually on the title page). Tip: The Author of a dataset may also be the Producer.</p>			
Name	<p>Formatting note: For personal names, use the format last name, first name</p> <p><i>Example: Social and Aboriginal Statistics Division.</i> <i>Example (if the division/unit is unknown): Statistics Canada</i></p>	R	R	<AuthEnty>
Affiliation	<p>The name of the entity affiliated with the author, e.g. an organization's name.</p> <p>Formatting note: use the format Name of Institution/University. Name of Data Centre/Division</p> <p><i>Example: Statistics Canada</i></p>	O	R	<AuthEnty> (affiliation)
Point of Contact	Names, affiliations, and addresses of institutions responsible for the work.			
Name	<p>The contact's family name, given name, or the name of the organization.</p> <p><i>Example: Odesi</i></p>	R	R	<contact>
Affiliation	<p>The organization with which the contact is affiliated.</p> <p>Tip: use the full official name of the organization; avoid abbreviations.</p> <p><i>Example: Scholars Portal</i></p>	O	R	<contact> (affiliation)
Email	<p>Email address for the contact institution. This will not be displayed.</p> <p><i>Example: odesi@scholarsportal.info</i></p>	R	R	<contact> (email)

Citation Metadata				
Field	Description	Usage	Repeatable	XML tag
Description	A summary describing the purpose, nature, and scope of the data collection, special characteristics of its contents, major subject areas covered, and what questions the PIs attempted to answer when they conducted the study. A listing of major variables in the study is important here.			
Text	<p>A summary describing the purpose, nature, and scope of the dataset.</p> <p><i>Example: The purpose of this survey is to better understand how Canadians perceive crime and the justice system and their experiences of victimization. This survey is the only national survey of self-reported victimization which provides data on criminal victimization for the provinces and territories. As not all crimes are reported to the police for a variety of reasons, the survey provides an important complement to officially recorded crime rates. It measures both crime incidents that come to the attention of the police and those that are unreported. It also helps to understand why some people choose whether or not to report a crime to the police. Results from this survey will be used by police departments, all levels of government, victim and social service agencies, community groups and researchers in universities to study Canadians' perceptions of the level of crime around them and their attitudes toward the criminal justice system; to profile victims of crimes; and to study characteristics of criminal incidents.</i></p>	R	R	<abstract>
Date	<p>The date in which the abstract was created. The date attribute follows the ISO convention of YYYY-MM-DD.</p> <p>Tip: In cases where a dataset contains more than one description (e.g. one supplied by the data producer and another prepared by the data repository where the data are deposited), the date attribute is used to distinguish between the two descriptions.</p>	O	R	<abstract> (date)

Citation Metadata				
Field	Description	Usag e	Repeatable	XML tag
Subject	<p>Domain-specific Subject Categories defined by Dataverse. For most Odesi material, the subject will be “Social Sciences”. Depending on the dataset you can add multiple subjects to describe the contents, such as CCHS (“Social Sciences”, “Medicine, Health and Life Sciences”).</p> <p>Tip: Be aware that this field is distinct from Topic Classification and Keyword.</p> <p><i>Example: Social Sciences</i></p> <p>→ See Appendix A: Subject (Borealis) for the controlled vocabulary terms for this field in Borealis (en français).</p>	R	NR	<keyword>
Keyword	<p>Words or phrases that describe salient aspects of a data collection's content. Can be used for building keyword indexes and for classification and retrieval purposes. Use this field to indicate the broad substantive topic(s) that the data cover. Use controlled vocabularies whenever possible, referencing the Vocabulary and Vocabulary URL for each keyword entered.</p> <p>Tip: Be aware that this field is distinct from Subject and Topic Classification.</p> <p>Odesi specific: Apply the Library and Archives Canada Thesauri whenever possible (en français).</p>			
Value	<p>Key terms that describe important aspects of the dataset.</p> <p><i>Example: Victims</i> <i>Example: Crime</i> <i>Example: Justice System</i> <i>Example: Law Enforcement</i></p>	R	R	<keyword>
Vocabulary	<p>For the specification of the keyword controlled vocabulary in use, such as LCSH, MeSH, or others.</p> <p>Tip: Use this field when the Keyword Value is from a controlled vocabulary.</p> <p><i>Example: Government of Canada Core Subject Thesaurus</i> <i>Example: Library and Archives Canada Thesauri</i></p>	O	R	<keyword> (vocab)

Citation Metadata				
Field	Description	Usage	Repeatable	XML tag
Vocabulary URL	<p>Enter an absolute URL³ that points to a location where the keyword vocabulary web site is found, such as http://www.my.org.</p> <p>Tip: Use this field when the Keyword Value is from a controlled vocabulary.</p> <p><i>Example: https://canada.multites.net/cst/index.htm</i> <i>Example: https://en.thesaurus.gc.ca/index.htm</i></p>	O	R	<keyword> (vocabURI)
Topic Classification	<p>Domain-specific subject categories defined by Odesi.</p> <p>Tip: Use title case.</p> <p>Odesi specific: Odesi usage of this field differs from general DV usage. Be consistent, keeping in mind that these categories are used as facets to filter the Browse results in the Odesi search site.</p> <p><i>Example: Crime and Justice</i></p> <p>→ See Appendix A: Topic Classification for Odesi’s controlled vocabulary terms for this field (en français).</p>	R	R	<topcClas>
Related Publication	Bibliographic and access information about articles and reports based on the data in this collection.			
Citation	<p>The full bibliographic citation for this related publication.</p> <p><i>Example: Garnett, Holly Ann, and Keir, Mark V. 2022. “Where Regulation Counts: Provincial Regulation of Municipal Elections in Canada.” Canadian Public Administration 65, 333– 351. doi: 10.1111/capa.12462</i></p>	O	R	<relPub> <citation> <biblCit>

³ An absolute URL is a complete address including its prefix (<https://>), as seen in provided examples.

Citation Metadata				
Field	Description	Usag e	Repeatable	XML tag
ID Type	<p>The type of digital identifier used for this publication (e.g., Digital Object Identifier (DOI)).</p> <p>Note: This field is a drop-down list; select one.</p> <p>Tip: Use a persistent identifier where possible (e.g., DOI, handle, ISBN). If only a URL is available, select the type “url” and enter the URL into both the ID Number and URL fields.</p> <p><i>Example: DOI</i> <i>Example: handle</i></p>	O	R	<relPubl> <citation> <titlStmt> <IDNo> (agency)
ID Number	<p>The identifier for the selected ID type.</p> <p>Tip: If using ID Type “url”, enter the URL into both this field and the URL field.</p> <p><i>Example: doi:10.1111/capa.12462</i> <i>Example: hdl:10864/11104</i></p>	O	R	<relPubl> <citation> <titlStmt> <IDNo>
URL	<p>Link to the publication web page (e.g., journal article page, archive record page, or other).</p> <p>Tip: This field is only used for display purposes in the Dataverse interface. For interoperability, ensure that the ID Type and ID Number fields are also completed.</p> <p><i>Example: https://doi.org/10.1111/capa.12462</i> <i>Example: https://hdl.handle.net/10864/11104</i></p>	O	R	<relPubl> <ExtLink> (uri)
Notes	<p>Additional important information about the dataset.</p> <p><i>Example: This survey was administered online. Mode of interview has been found to impact results, therefore it is not recommended that these results are compared with other survey results where the interview mode was telephone based.</i></p>	O	R	<stdyInfo> <notes>

Citation Metadata				
Field	Description	Usage	Repeatable	XML tag
Language	<p>Language of the dataset.</p> <p>Note: This field is a drop-down list. Choose <i>English</i> or <i>French</i> based on the dataset and markup language.</p> <p>Odesi specific: Be consistent, keeping in mind that Language is among the facets used to filter the Browse results in the Odesi search site.</p>	R	R	<slang> (custom Odesi tag)
Producer	<p>The producer of the data collection is the person or organization with the financial or administrative responsibility for the physical processes whereby the data collection was brought into existence.</p> <p>Tip: The Producer of a dataset may also be the Author.</p>			
Name	<p>Producer name.</p> <p><i>Example: Social and Aboriginal Statistics Division</i> <i>Example (if division/unit unknown): Statistics Canada</i></p>	R	R	<producer>
Affiliation	<p>The organization with which the producer is affiliated.</p> <p><i>Example: Statistics Canada</i></p>	O	R	<producer> (affiliation)
Abbreviation	<p>The abbreviation by which the producer is commonly known. (ex. IQSS, ICPSR).</p> <p><i>Example: SASD</i></p>	O	R	<producer> (abbr)
Production Date	<p>Date the data collection was produced (not distributed or archived).</p> <p>Tip: Date when dataset was finalized and ready for analysis or distribution.</p> <p><i>Example: 2010-09-28</i></p>	O	NR	<prodDate>

Citation Metadata				
Field	Description	Usag e	Repeatable	XML tag
Production Place	Address of the archive or organization that produced the work. Tip: Include an institution if relevant. No abbreviations should be used. <i>Example: Ottawa, Ontario: Statistics Canada</i>	O	NR	<prodPlac>
Contributor	Statements of responsibility not recorded in the title and statement of responsibility areas. Indicate here the persons or bodies connected with the work, or significant persons or bodies connected with previous editions and not already named in the description. Tip: The contributor type Funder may overlap with Grant agency.			
Type	<i>Example: Data Collector</i> → See Appendix A: Contributor Type (Borealis) for the controlled vocabulary terms for this field in Borealis (en français).	O		<othID> (type)
Name	<i>Example: EKOS Research Associates</i>	O		<othID>
Grant Information	The grant/contract number of the project that sponsored the effort. <i>Example: Social Sciences and Humanities Research Council (SSHRC)</i>			
Grant Agency	Grant Number Agency. Tip: This is the grant/funding institution that issued and assigned the Grant Number to the sponsored research effort. Also note that the Grant Agency may also be the Contributor (type: Funder) . <i>Example: Social Sciences and Humanities Research Council (SSHRC)</i>	O	R	<grantNo> (agency)
Grant Number	The grant or contract number of the project that sponsored the effort. <i>Example: '4-567-0M'</i>	O	R	<grantNo>

Citation Metadata				
Field	Description	Usag e	Repeatable	XML tag
Distributor	<p>The organization designated by the author or producer to generate copies of a particular data collection including any necessary editions or revisions. Names and addresses may be specified, and other archives may be co-distributors.</p> <p>Note: If there is more than one affiliation/organization that is a distributor, this tag is repeated.</p>			
Name	<p>Distributor name.</p> <p><i>Example: Data Liberation Initiative</i></p>	O	R	<distrbtr>
Affiliation	<p>The organization with which the distributor contact is affiliated.</p> <p><i>Example: Statistics Canada</i></p>	O	R	<distrbtr> (affiliation)
Abbreviation	<p>The abbreviation by which this distributor is commonly known (e.g., IQSS, ICPSR).</p> <p><i>Example: DLI</i></p>	O	R	<distrbtr> (abbr)
URL	<p>Distributor URL points to the distributor's web presence, if appropriate. Enter an absolute URL where the distributor's web site is found, such as https://www.my.org.</p> <p><i>Example: https://www.statcan.gc.ca/en/microdata/dli</i></p>	O	R	<distrbtr> (URL)
Distribution Date	<p>Date that the work was made available for distribution/presentation.</p> <p>Tip: This is the date the dataset was made available by your institution (the secondary distributor), not the date of release by the Producer.</p> <p><i>Example: 2009</i></p>	O	NR	<distDate>
Depositor	<p>The name of the institution (and person, where applicable) who provided this work to the archive storing it.</p> <p><i>Example: Queen's University Library</i></p>	O	NR	<depositr>

Citation Metadata				
Field	Description	Usage	Repeatable	XML tag
Deposit Date	Date that the data was deposited with the archive that originally received it. Tip: This is the date the dataset was made available by your institution (the secondary distributor), not the date of release by the Producer . <i>Example: 2010-01-01 (YYYY-MM-DD)</i>	O	NR	<depDate>
Time Period Covered	The time period to which the data refer. This item reflects the time period covered by the data, not the dates of coding or making documents machine-readable or the dates the data were collected. Also known as <i>span</i> .			
Start	Start date of the time period covered by the data, not the dates of coding or making documents machine readable or the dates the data were collected. <i>Example: 2009-02</i>	O	R	<timePrd> (start)
End	End date of the time period covered by the data, not the dates of coding or making documents machine readable or the dates the data were collected. <i>Example: 2009-11</i>	O	R	<timePrd> (end)
Date of Collection	Contains the date(s) when the data were collected.			
Start	Date when the data collection started. <i>Example: 2009-11</i>	O	R	<collDate> (start)
End	Date when the data collection ended. <i>Example: 2009-12</i>	O	R	<collDate> (end)

Citation Metadata				
Field	Description	Usag e	Repeatable	XML tag
Data Type	<p>The type of data included in the file: survey data, census/enumeration data, aggregate data, clinical data, event/transaction data, program source code, machine-readable text, administrative records data, experimental data, psychological test, textual data, coded textual, coded documents, time budget diaries, observation data/ratings, process-produced data, etc.</p> <p><i>Example: Microdata</i> <i>Example: Survey data</i></p> <p>→ See Appendix A: Data Type for Odesi’s controlled vocabulary terms for this field (en français).</p>	R	R	<dataKind>
Series	Series statement for the data collection.			
Name	<p>The name of the data series to which the collection belongs.</p> <p><i>Example: General Social Survey</i></p>	R	NR	<serName>
Information	<p>Contains a history of the data series and a summary of those features that apply to the data series as a whole.</p> <p><i>Example: The two primary objectives of the General Social Survey (GSS) are: to gather data on social trends in order to monitor changes in the living conditions and well being of Canadians over time; and to provide information on specific social policy issues of current or emerging interest.</i></p>	O	NR	<serInfo>
Holdings	<p>Information about the physical or electronic holdings of the cited work.</p> <p><i>Example: https://doi.org/10.5072/FK2/13X8U5</i></p>	O	R	<holdings>
Software	Information about the software used to generate the dataset.			
Name	<p>Name of software used to generate the dataset.</p> <p><i>Example: SPSS</i></p>	O	R	<software>

Citation Metadata				
Field	Description	Usage	Repeatable	XML tag
Version	Version of the software used to generate the dataset. <i>Example: 26</i>	O	R	<software> (version)
Related Material	Describes materials related to the study description, such as appendices, additional information on sampling found in other documents, etc. Can take the form of bibliographic citations. May consist of a single URI or a series of URIs comprising a series of citations/references to external materials which can be objects as a whole (journal articles) or parts of objects (chapters or appendices in articles or documents). Tip: These tend to be secondary materials or accompanying documentation (e.g. technical guides, websites, etc.) linked to the study description. <i>Example: Canada Yearbook, 2007</i>	O	R	<relMat>
Related Datasets	Information on the relationship of the current data collection to others (e.g., predecessors, successors, other waves or rounds) or to other editions of the same file. This would include the names of additional data collections generated from the same data collection vehicle plus other collections directed at the same general topic. Can take the form of bibliographic citations. <i>Example: Cycle 3.1, 2005, Public Use Microdata File (PUMF)</i>	O	R	<relStdy>
Data Sources	Used to list the book(s), article(s), serial(s), and/or machine-readable data file(s)—if any—that served as the source(s) of the data collection. <i>Example: Survey respondents</i>	O	R	<dataSrc>
Origin of Sources	For historical materials, information about the origin(s) of the sources and the rules followed in establishing the sources should be specified. May not be relevant to survey data. <i>Example: ICPSR</i>	O	NR	<srcOrig>

Geospatial Metadata

Metadata that describes the spatial extent (e.g., location information) of data in a dataset. Geospatial metadata can describe maps, GIS files, or other location-based data. Any dataset that has a spatial extent—that relates to a location(s) or geographic area(s)—should include geospatial metadata in addition to the general citation metadata block.

At a minimum, provide place names to describe locations in your data. Use [GeoNames.org](https://www.geonames.org) to confirm these terms. Alternate names (e.g., in other languages) may be added. If applicable, enter the bounding box coordinates to allow the data to be findable with map-based search tools.

Geospatial Metadata				
Field	Description	Usage	Repeatable	XML tag
Geographic Coverage	Information on the geographic coverage of the data. Includes the total geographic scope of the data. Tip: For consistency, use the GeoNames database to check the form and spelling of place names.			
Country/ Nation	The country or nation that the dataset is about. Note: This field is a drop-down list of ISO-3166 country names. If the dataset covers multiple countries, list all of them. <i>Example: Canada</i>	R	R	<geogCover>
State/Province	The state or province that the dataset is about. Use GeoNames for correct spelling and avoid abbreviations. Tip: If using this field, also include Country to disambiguate. <i>Example: Manitoba</i>	O	R	<geogCover>

Geospatial Metadata				
Field	Description	Usage	Repeatable	XML tag
City	<p>The name of the city that the dataset is about. Use GeoNames for correct spelling and avoid abbreviations.</p> <p>Tip: If using this field, also include Country AND, if possible, State/Province to disambiguate.</p> <p><i>Example: Winnipeg</i></p>	O	R	<geogCover>
Other	<p>Other information on the geographic coverage of the data.</p> <p>Tip: Use for geographical names that are not a country, state/province, or city, e.g., regions, water bodies, astronomy names. If applicable, disambiguate by including City AND/OR State/Province AND/OR Country.</p> <p><i>Example: Kingston, Frontenac and Lennox and Addington Health Unit</i></p>	O	R	<geogCover>
Geographic Unit	<p>Lowest level of geographic aggregation covered by the dataset, e.g., village, county, region.</p> <p>Tip: Include the lowest geographic level that can be analyzed in the dataset (e.g., if the geography is province, CMA, and CSD, enter CSD).</p> <p><i>Example: Census metropolitan area (CMA)</i></p> <p>→ See Appendix A: Geographic Unit for Odesi’s controlled vocabulary terms for this field (en français).</p>	R	R	<geogUnit>

Social Science and Humanities Metadata

The Social Science and Humanities block contains metadata that describes survey data and related datasets in the social science and humanities.

Social Science and Humanities Metadata				
Field	Description	Usage	Repeatable	XML tag
Unit of Analysis	Basic unit of analysis or observation that the file describes: individuals, families/households, groups, institutions/organizations, administrative units, etc. <i>Example: Individuals</i>	O	R	<anlyUnit>
Universe	A description of the population covered by the data in the file; the group of persons or other elements that are the object of the study and to which the study results refer. Age, nationality, and residence commonly help to delineate a given universe, but any of a number of factors may be involved, such as age limits, sex, marital status, race, ethnic group, nationality, income, veteran status, criminal convictions, etc. The universe may consist of elements other than persons, such as housing units, court cases, deaths, countries, etc. In general, it should be possible to tell from the description of the universe whether a given individual or element (hypothetical or real) is a member of the population under study. Also known as universe of interest, population of interest, and target population. Tip: This field refers to the universe of the study. The universe for each individual variable is entered in the Variable Metadata section's Universe field. <i>Example: Canadian population aged 15 and over and not residing in institutions.</i>	O	R	<universe>
Time Method	The time method or time dimension of the data collection. <i>Example: This is a sample survey with a cross-sectional design.</i>	O	NR	<timeMeth>

Social Science and Humanities Metadata				
Field	Description	Usage	Repeatable	XML tag
Data Collector	<p>The entity (individual, agency, or institution) responsible for administering the questionnaire or interview or compiling the data. This refers to the entity collecting the data, not to the entity producing the documentation.</p> <p><i>Example: Social and Aboriginal Statistics Division</i></p>	O	NR	<dataCollector>
Frequency	<p>If the data collected include more than one point in time, indicate the frequency with which the data were collected.</p> <p><i>Example: Quinquennial (5 year)</i></p>	O	NR	<frequenc>
Sampling Procedure	<p>The type of sample and sample design used to select the survey respondents to represent the population. May include reference to the target sample size and the sampling fraction.</p> <p><i>Example: In the ten provinces, households were selected for the survey by Random Digit Dialing.</i></p>	O	NR	<sampProc>
Collection Mode	<p>The method used to collect the data; instrumentation characteristics.</p> <p><i>Example: Data collection was conducted by Computer Assisted Telephone Interviewing (CATI) methods in the ten provinces and by a combination of CATI and Computer Assisted Person Interviewing (CAPI) methods in the territories.</i></p>	O	NR	<collMode>
Type of Research Instrument	<p>The type of data collection instrument used. "Structured" indicates an instrument in which all respondents are asked the same questions/tests, possibly with precoded answers. If a small portion of such a questionnaire includes open-ended questions, provide appropriate comments. "Semi-structured" indicates that the research instrument contains mainly open-ended questions. "Unstructured" indicates that in-depth interviews were conducted.</p> <p><i>Example: Semi-structured questionnaire</i></p>	O	NR	<resInstru>

Social Science and Humanities Metadata				
Field	Description	Usage	Repeatable	XML tag
Characteristics of Data Collection Situation	Description of noteworthy aspects of the data collection situation. Includes information on factors such as cooperativeness of respondents, duration of interviews, number of call-backs, etc. <i>Example: Data collection for this reference period: 2009-02-02 – 2009-11-30. Responding to this survey is voluntary. Data are collected directly from survey respondents.</i>	O	NR	<collSitu>
Weighting	The use of sampling procedures might make it necessary to apply weights to produce accurate statistical results. Describes the criteria for using weights in analysis of a collection. If a weighting formula or coefficient was developed, the formula is provided, its elements are defined, and it is indicated how the formula was applied to the data. <i>Example: The respondent weights are adjusted to represent the individuals who did not respond to the survey. Adjustment factors are computed separately by province based on a non-response model using frame information. The person weights coming from the household sample and the targeted respondent sample are pooled together. The person weights are calibrated so that the sum of the weights match demographic population counts at the region level by age group and by gender. The weights are also calibrated to demographic counts for large Census Metropolitan Areas. Variance estimation is based on a resampling method called the bootstrap.</i>	O	NR	<weight>
Response Rate	The percentage of sample members who provided information, if available. <i>Example: A response rate of 37% was obtained for youth aged 15 to 24, and 48% for adults aged 25 and over; an overall response rate of 42% was obtained for the CTNS.</i>	O	NR	<respRate>
Estimates of Sampling Error	Measure of how precisely one can estimate a population value from a given sample. <i>Example: Non-response errors</i>	O	NR	<EstSmpErr>

Social Science and Humanities Metadata				
Field	Description	Usage	Repeatable	XML tag
Other Forms of Data Appraisal	<p>Other issues pertaining to the data appraisal. Describe issues such as response variance, nonresponse rate and testing for bias, interviewer and response bias, confidence levels, question bias, or similar.</p> <p><i>Example: The first type of error treated was errors in questionnaire flow, where questions which did not apply to the respondent (and should therefore not have been answered) were found to contain answers. In this case a computer edit automatically eliminated superfluous data by following the flow of the questionnaire implied by answers to previous, and in some cases, subsequent questions.</i></p>	O	NR	<anlyInfo>
Notes	<p>General notes about this dataset.</p> <p>Tip: Includes notes related to fields in the Social Science and Humanities section.</p> <p><i>Example: This dataset contains two data files</i></p>			
Type	Type of note	O	NR	<notes> (type)
Subject	Note subject	O	NR	<notes> (subject)
Text	Text for this note	O	NR	<notes>

Terms of Use Metadata

Terms of Use Metadata				
Field	Description	Usage	Repeatable	XML tag
License/Data Use Agreement	<p>This section specifies the terms or license under which the dataset is published. Terms of Use and Licenses are formal documents that outline how data can be used once it has been downloaded.</p> <p>The drop-down menu list of data licenses that can be selected for a dataset. If the correct license is not listed, select Custom Dataset Terms.</p>	R	NR	<dataAccs> <notes> (type=DVN:TOU; level="dv")
Terms of Use	<p>Outlines how this data can be used once it has been downloaded.</p> <p>Note: This Terms of Use field and those below will only appear if users use the drop-down menu to opt out of a Creative Commons (CC) preset license.</p> <p><i>Example: This data is to be used for teaching and research purposes only. It is not to be used for any profit-making purposes.</i></p>	O	NR	<dataAccs> <notes> (type=DVN:TOU; level="dv")
Restrictions	<p>Any restrictions on access to or use of the collection such as privacy certification or distribution restrictions should be indicated here. These can be restrictions applied by the author, producer, or disseminator of the data collection. If the data are restricted to only a certain class of user, specify which type.</p> <p><i>Example: The data may be used for personal, academic research or teaching purposes only.</i></p>	O	NR	<restrctn>

Terms of Use Metadata				
Field	Description	Usage	Repeatable	XML tag
Citation Requirements	<p>Text of requirement that a data collection should be cited properly in articles or other publications that are based on analysis of the data.</p> <p><i>Example: The publishing of analysis and results from research using any of the data products is permitted in research communications such as scholarly papers, journals and the like. The authors of these communications are required to cite the Angus Reid Group (Ipsos-Reid) as the source of the data, and to indicate that the results or views expressed are those of the author/authorized user and are not those of the Angus Reid Group (Ipsos-Reid). Permission to include extracts of these data in textbooks must be obtained from Ipsos-Reid.</i></p>	O	NR	<citReq>
Depositor Requirements	<p>Information regarding user responsibility for informing archives of their use of data through providing citations to the published work or providing copies of the manuscripts.</p>	O	NR	<deposReq>
Conditions	<p>Indicate any additional information that will assist the user in understanding the access and use conditions of the data collection.</p> <p><i>Example: These data may be used by current students, faculty and staff of the University of Guelph for academic research only.</i></p>	O	NR	<conditions>
Disclaimer	<p>Information regarding responsibility for uses of the data collection.</p> <p><i>Example: These data are provided “as is”; the provider makes no representations or warranties, either expressed or implied, as to the appropriateness and fit for a particular purpose.</i></p>	O	NR	<disclaimer>

Terms of Access Metadata

Terms of Access Metadata				
Field	Description	Usage	Repeatable	XML tag
Request Access	If checked, users can request access to the restricted files in this dataset.	R	NR	n/a
Terms of Access for Restricted Files	Information on how and if users can access restricted files in this dataset. <i>Example: Statistics Canada Licence Agreement.</i>	R	NR	<dataAccs> <notes> (type=DVN:TOA; level="dv")
Data Access Place	Location where the data collection is currently stored. <i>Example: Toronto, Ontario, Canada</i>	O	NR	<setAvail> <accsPlac>
Original Archive	Archive from which the data collection was obtained; the originating archive. Note: This tag applies only to non-Statistics Canada data. <i>Example: ICPSR</i>	O	NR	<origArch>
Contact for Access	Name of institution(s) responsible for the work and their address. Tip: this field can be used for contacting the access authority, if this is required or different from the main dataset contact. <i>Example: Odesi</i>	O	NR	<useStmt> <contact>
Availability Status	Statement of collection availability. An archive may need to indicate a collection is unavailable because it is embargoed for a period of time, because it has been superseded, because a new edition is imminent, etc. <i>Example: Original data is restricted due to sensitivity</i>	O	NR	<avlStatus>

Terms of Access Metadata				
Field	Description	Usage	Repeatable	XML tag
Size of Collection	<p>Summary of the number of physical files that exist in a dataset, recording the number of files that contain data and noting whether the collection contains machine readable documentation and/or other supplementary files and information, such as code, data dictionaries, data definition statements, or data collection instruments.</p> <p><i>Example: This dataset contains 1 datafile in the 4 following formats: ASCII, SPSS, SAS, and Stata.Syntax (command code) files for SPSS, SAS, and Stata are also included along with 2 documentation files (questionnaire and codebook),</i></p>	O	NR	

File Description Metadata

This section consists of metadata items describing the characteristics and contents of file(s) that comprise the dataset. When a dataset contains multiple files, each file loaded to Borealis will require separate metadata.

File Description Metadata				
Field	Description	Usage	Repeatable	XML tag
File Description	Information about data files loaded to Borealis. For full details and instructions on file loading see the Odesi Deposit Guide .			
File Name	<p>Name of the file.</p> <p>Tip: Borealis pre-fills this field. Note that file names must be consistent with OtherID and contain the format extension in the name. All data files that are not SPSS must be double-zipped prior to upload (for more details and instructions on file loading see the Odesi Deposit Guide).</p> <p>For Statistics Canada surveys: Use the format <i>Acronym-CatalogueNumberOrRecordNumber-language-year-subset-Aggregate DataFileType</i></p> <p><i>Example (Statistics Canada surveys): gss-12M0018-E-2009-c-23-vif_F1.tab</i></p>	R	NR	<fileName>
Type of File	The type of file is generated automatically by Borealis, using the format extension (e.g. PDF, etc.) from the name of the file that is loaded.	System Generated		<fileType>
Description (type of file)	<p>Use this field to describe the type of file loaded. Files can be either documentation files or data files.</p> <p>Tip: This field refers to the format of the file. This is distinct from Type of Data.</p> <p><i>Example: Data in comma separated value format.</i></p> <p>→ See Appendix A: Type of File for Odesi’s controlled vocabulary terms for this field (en français).</p>	R	NR	<otherMat> <txt>

File Description Metadata				
Field	Description	Usage	Repeatable	XML tag
File Dimensions	Dimensions of the overall file.			
Number of cases / Record Quantity	Number of cases or observations in the entire file. Used for rectangular files only.	System Generated		<caseQty>
Number of variables per record	Number of variables in the entire file. This is to be used for rectangular files only.	System Generated		<varQty>
Notes	In Borealis, this field is used for the Universal Numeric Footprint (UNF), a unique signature of the semantic content of a digital object.		System Generated	<fileDscr> <notes>
Level		System Generated		<fileDscr> <notes> (level)
Type		System Generated		<fileDscr> <notes> (type)
Subject		System Generated		<fileDscr> <notes> (subject)
Other Material	Information about all other (not tabular) files.			
ID		System Generated		<otherMat> (ID)
URI		System Generated		<otherMat> (URI)

File Description Metadata				
Field	Description	Usage	Repeatable	XML tag
Level	<p>The "level" attribute is used to clarify the relationship of the other materials to components of the study.</p> <p>Suggested values for level include specifications of the item level to which the element applies: e.g.,</p> <ul style="list-style-type: none"> • level= data; • level=datafile; • level=studydsc; • level=study. 	System Generated		<otherMat> (level)
File Name	Name of the file.	System Generated		<otherMat> <labl>
Notes	In Borealis, this field is used for MIME type (media type).	System Generated		<otherMat> <notes>
Level		System Generated		<otherMat> <notes> (level)
Type		System Generated		<otherMat> <notes> (type)
Subject		System Generated		<otherMat> <notes> (subject)

Variable Metadata

This section contains a rich set of elements with detailed descriptive information about response and analysis units, question text, forward progression and backflow, interviewer instructions, universe, valid and invalid data ranges, derived variables, summary statistics, etc. The section has two sub-sections: A sub-section describing variable groups and one describing individual variables within the data file. System generated fields are produced during tabular file ingest.

All variable metadata fields used in Odesi are described below. In Borealis, most variable metadata is entered and displayed using the Data Curation Tool. A separate guide provides detailed information and instructions on using this tool to edit variable metadata: [Odesi Deposit Guide](#). Some of these fields may have associated [DDI Controlled Vocabularies \(en français\)](#).

Variable Metadata				
Field	Description	Usage	Repeatable	XML tag
Variable Group	A group of variables that may share a common subject, arise from the interpretation of a single question, or are linked by some other factor. Note: this field is editable via creating groups in the Data Curation Tool	O	R	<varGrp>
ID	To uniquely identify the Variable Group	System Generated		<varGrp> (ID)
List	Listing of variables that make up the Variable Group	System Generated		<varGrp> (list)
Label	A short description of the parent element, that is, the group. In the variable label, the length of this phrase may depend on the statistical analysis system used <i>Example: SRV: Survey Related Variables</i>	O	R	<varGrp> <labl>
Variable	This element describes all of the features of a single variable in a social science data file. There are a number of attributes within this tag that allows us to specify items such as the name of the variable, what the weight variable is, etc.	System Generated		<var>
Variable ID	To uniquely identify the variable.	System Generated		<var> (ID)

Variable Metadata				
Field	Description	Usage	Repeatable	XML tag
Interval	Indicates whether the variable is continuous or discrete, ordinal or nominal.	System Generated		<var> (interval)
Files	Indicates the file where the variable can be found	System Generated		<var> (files)
Weight	Indicates whether the variable is the weight variable or not. Tip: This field is editable via selecting “is weight” on the variable you wish to declare, in the Data Curation Tool. <i>Example: wgt</i>	O	R	<var> (wgt)
Name	Name of the variable	System Generated		<var> (name)
Location	This is an empty element containing only the attribute “fileid” (an IDREF link to the fileDscr element for the file that this location is within).	System Generated		<var> <location> (fileid)
Label	A short description of the parent element. In the variable label, the length of this phrase may depend on the statistical analysis system used (e.g., some versions of SAS permit 40-character labels, while some older versions of SPSS permit 120 characters), although the DDI itself imposes no restrictions on the number of characters allowed.	System Generated		<var><labl>

Variable Metadata				
Field	Description	Usage	Repeatable	XML tag
Universe	<p>The class of individuals that share a unit type and typically have other characteristics in common (beyond time and geography). For example a unit type might be 'humans' and a universe might be 'humans who are nurses'. A population is a universe at a given place and time.⁴</p> <p>Tip: This field refers to the universe for each individual question. This field can be edited on Borealis via the Data Curation Tool. The study's universe is entered in the Social Science and Humanities metadata section's Universe field.</p> <p><i>Example: Respondents who consumed an alcoholic beverage in the last 30 days (DRR_110 = 1, 2, 3, 4, 5).</i></p>	O	R	<var> <universe>
Summary Statistics	One or more statistical measures that describe the responses to a particular variable and may include one or more standard summaries, e.g., minimum and maximum values, median, mode, etc.	System Generated		<var> <sumStat>
Summary Statistics Type	Denotes the type of statistics being shown: mean, median, mode, valid cases, invalid cases, minimum, maximum, or standard deviation.	System Generated		<var> <sumStat> (type)
Category	A description of a particular response. The attribute "missing" indicates whether this category group contains missing data or not.	System Generated		<catgry>
Category Value	The explicit response.	System Generated		<catgry> <catValu>

⁴ Source: [DDI Glossary \(Draft, 2023\)](#).

Variable Metadata				
Field	Description	Usage	Repeatable	XML tag
Category Label	A short description of the response. In the variable label, the length of this phrase may depend on the statistical analysis system used (e.g., some versions of SAS permit 40-character labels, while some older versions of SPSS permit 120 characters). Although the DDI itself imposes no restrictions on the number of characters allowed. A “level” attribute is included to permit coding of the level to which the label applies, i.e. record group, variable group, variable, category group, category, nCube group, nCube, or other study-related materials.	System Generated		<catgry> <labl>
Category Group Statistics Type	May include frequencies, percentages, or cross-tabulation results which define the category; often appears in a table. The attribute “type” indicates the type of statistics presented – frequency, percent, or crosstabulation.	System Generated		<catgry> <catStat> (type)
Category Group Statistics Type Weighted		System Generated		<catgry> <catStat> (wgt)
Variable Format Type	The technical format of the variable in question. <i>Example: numeric</i>	System Generated		<catgry> <varFormat> (type)
Variable Format Type Schema	indicates if the variable is character or numeric	System Generated		<catgry> <varFormat> (schema)
Notes	This field is used for UNF (Universal Numeric Fingerprint) in Borealis. A UNF is a unique signature of the semantic content of a digital object.	System Generated		<catgry> <notes>

Variable Metadata				
Field	Description	Usage	Repeatable	XML tag
Question	<p>The question element may have mixed content. The element itself may contain text for the question, with the sub-elements being used to provide further information about the question. Alternatively, the question element may be empty and only the sub-elements used. The element has a unique question ID attribute which can be used to link a variable with other variables where the same question has been asked. This would allow searching for all variables that share the same question ID perhaps because the questions were asked several times in a panel design.</p> <p>Tip: Information for this field and its subfields can be obtained from the dataset Questionnaire files. This field and its sub-fields can be edited on Borealis via the Data Curation Tool.</p>	O	R	<var><qstn>
PreQuestion Text	<p>Text describing a set of conditions under which a question might be asked.</p> <p><i>Example: The following question collects information in accordance to the Employment Equity Act and its Regulations and Guidelines to support programs that promote equal opportunity for everyone to share in the social, cultural, and economic life of Canada.</i></p>	O	NR	<var> <qstn> <preQTxt>
Literal Question Text	<p>Text of the actual, literal question asked.</p> <p><i>Example: To begin with, thinking of the issues presently confronting Canada, which one do you feel should receive the greatest attention from Canada's leaders? What other issues do you think are important for Canada right now?</i></p>	R	NR	<var> <qstn> <qstnLit>
PostQuestion Text	<p>Text describing what occurs after the literal question has been asked.</p>	O	NR	<var><qstn> <postQTxt>
Interviewer Instructions	<p>Specific instructions to the individual conducting an interview.</p> <p><i>Example: 1st Mention ONE ONLY</i></p> <p><i>Example: [FOR ALL QUESTIONS THAT ASK TO INSERT COUNTRY, INSERT THIS TEXT:]</i></p>	O	NR	<var> <qstn> <ivulnstr>

Appendix A: Controlled Vocabularies

Controlled vocabularies contribute to making research data more FAIR (findable, accessible, interoperable and reusable) by defining a set of standardized terminology. Controlled vocabularies enable dataset discovery through the Odesi portal, powering search filters and facets.

The controlled vocabularies used in Odesi markup are listed below. Some of the controlled vocabularies used in Odesi were created for Odesi collections specifically, and others are generic controlled vocabularies recommended in DDI and/or Borealis (Dataverse) best practices.

English Controlled Vocabularies

Odesi-Specific

Geographic Unit

*Collection: Statistics Canada Public Use
Microdata (PUMF)*

- Arctic
- Census agglomeration (CA)
- Census consolidated subdivision (CCS)
- Census division (CD)
- Census metropolitan area (CMA)
- Census subdivision (CSD)
- City
- Community type
- Cluster
- Country
- District
- Economic region
- Fishing region
- Forward sortation area (FSA)
- Health region
- Major urban centre
- Municipality
- National capital region
- Non-census metropolitan area
- Province
- Public health unit
- Region
- Rural area
- State
- Subprovincial region
- Territory
- Tourist region
- Type of inhabited area
- Urban area
- U.S. Region

Collection: Statistics Canada Aggregate Data

- Census consolidated subdivision (CCS)
- Census division (CD)
- Census metropolitan area (CMA)
- Census subdivision (CSD)
- Country
- Economic region
- Electoral district
- Forward sortation area (FSA)
- Health region
- Municipality
- Polling station
- Postal area
- Province
- Region
- State
- Territory
- Urban area

Data Type

Collection: Statistics Canada Public Use Microdata (PUMF)

- Administrative records data
- Census data
- Simulation data
- Survey data
- Synthetic data
- Microdata

Topic Classification

- Agriculture
- Business and Financial
- Census of Population
- Communications and Information
- Consumer Surveys
- COVID-19
- Crime and Justice
- Demographics and Population
- Education
- Elections and Politics
- Geography
- Government Finances and Economic Indicators
- Health
- Labour and Employment
- Natural Resources and Environment
- Public Opinion Polls
- Social Surveys
- Trade
- Travel

Collection: Statistics Canada Aggregate Data

- Administrative records data
- Aggregate data
- Census data
- Geographic data
- Simulation data
- Survey data
- Synthetic data

Type of File

- Data Documentation – Questionnaire
- Data Documentation – Codebook
- Data Documentation – Methodology
- Data Documentation – User Guide
- Data in comma separated value format
- Data in STATA7 format
- Data in STATA format
- Data in SAS format or SAS auxiliary file
- Data in SPSS format
- Data in SPSS syntax format or program file
- Others

Other Controlled Vocabularies

Subject (Borealis)

- Agricultural Sciences
- Arts and Humanities
- Astronomy and Astrophysics
- Business and Management
- Chemistry
- Computer and Information Science
- Earth and Environmental Sciences
- Engineering
- Law
- Mathematical Sciences
- Medicine, Health and Life Sciences
- Physics
- Social Sciences
- Other

Keyword

Use [Library and Archives Canada thesauri](#).

Geographic Coverage

State/Province

City

Use [GeoNames](#) for correct spelling and avoid abbreviations.

Country/Nation

Use ISO-3166.

Contributor Type (Borealis)

(See also: [DDI Controlled Vocabularies — Contributor Role](#))

Data Collector	Person/institution responsible for finding or gathering/collecting data under the guidelines of the author(s) or Principal Investigator (PI).
Data Curator	Person tasked with reviewing, enhancing, cleaning, or standardizing metadata and the associated data submitted for storage, use, and maintenance within a data centre or repository.
Data Manager	Person (or organization with a staff of data managers, such as a data centre) responsible for maintaining the finished resource.
Editor	A person who oversees the details related to the publication format of the resource Note: if the Editor is to be credited in place of multiple creators, the Editor's name may be supplied as Creator, with "(Ed.)" appended to the name.
Funder	Information about financial support (funding) for the resource being registered.
Hosting Institution	Typically, the organization that allows a resource to be available on the internet through the provision of its hardware/software/operating support.
Project Leader	Person officially designated as head of project team or sub- project team instrumental in the work necessary to development of the resource.
Project Manager	Person officially designated as manager of a project. Project may consist of one or many project teams and sub-teams.
Project Member	Person on the membership list of a designated project/project team.
Related Person	A person without a specifically defined role in the development of the resource, but who is someone the author wishes to recognize.
Researcher	A person involved in analyzing data or the results of an experiment or formal study. May indicate an intern or assistant to one of the authors who helped with research but who was not so "key" as to be listed as an author.
Research Group	Typically refers to a group of individuals with a lab, department, or division that has a specifically defined focus of activity.
Rights Holder	Person or institution owning or managing property rights, including intellectual property rights over the resource.
Sponsor	Person or organization that issued a contract or under the auspices of which a work has been written, printed, published, developed, etc.
Supervisor	Designated administrator over one or more groups/teams working to produce a resource, or over one or more steps of a development process.
Work Package Leader	A Work Package is a recognized data product, not all of which is included in publication. The package, instead, may include notes, discarded documents, etc. The Work Package Leader is responsible for ensuring the comprehensive contents, versioning, and availability of the Work Package during the development of the resource.
Other	Any person or institution making a significant contribution to the development and/or maintenance of the resource, but whose contribution is not adequately described by any of the other values for contributorType.

Contributor Type *table adapted from Dataverse and [DataCite documentation](#).*

DDI Controlled Vocabularies

The DDI Alliance maintains a set of controlled vocabularies for use with the DDI standard, some of which have multilingual versions available. The latest English versions of these controlled vocabularies can be found through the [CESSDA Vocabulary Service](#).

AggregationMethod	LanguageProficiency	TypeOfAddress
AnalysisUnit	ModeOfCollection	TypeOfConceptGroup
CharacterSet	NumericType	TypeOfFrequency
CommonalityType	ResponseUnit	TypeOfInstrument
ContributorRole	SamplingProcedure	TypeOfNote
DataSourceType	SoftwarePackage	TypeOfTelephone
Data Type	SummaryStatisticType	TypeOfTranslationMethod
DateType	TimeMethod	
GeneralDataFormat	TimeZone	

Vocabulaires contrôlés en français

Propres à l'Odesi

Unité géographique

Collection: Statistique Canada microdonnées

Arctique
Agglomération de recensement
Division de recensement
État
Municipalité
Pays
Région
Région agricole de recensement (RAR)
Région de la capitale nationale
Région économique
Région économique du recensement (RER)
Région géographique
Région géographique contigüe (grappe)
Région métropolitaine de recensement (RMR)
Région métropolitaine du Canada
Région sanitaire
Région sociosanitaire
Région rurale
Région touristique
Région urbaine
Subdivision de recensement
Subdivision de recensement unifiées (SRU)
Territoire

Collection: Statistique Canada agrégées

Subdivision de recensement unifiée (SRU)
Divisions de recensement
Région métropolitaine de recensement (RMR)
Subdivision de recensement (SDR)
Pays
Région économique (RE)
Circonscription électorale
Régions de tri d'acheminement censitaires (RTA)
Régions sociosanitaires
Municipalité
Code postal
Province
Région
Territoire
Régions Urbaines

Type de données

Collection: Statistique Canada microdonnées

Données des dossiers administratifs
Données de recensement
Données de simulation
Données d'enquête
Microdonnée

Collection: Statistique Canada agrégées

Données des dossiers administratifs
Données agrégées
Données de recensement
Données géographique
Données de simulation
Données d'enquête
Données synthétiques

Classification par sujets

Agriculture
Commerce
Communications et information
COVID-19
Démographie et population
Emploi et main-d'oeuvre
Enquêtes auprès des consommateurs
Enquêtes sociales
Finances publiques et indicateurs économiques
Géographie
Industries et Finance
Ressources naturelles et environnement
Santé
Voyages

Type de fichier

Documentation des données – Questionnaire
Documentation des données – Codebook
Documentation des données – Méthodologie
Documentation des données – Guide de référence
Données au format ASCII
Données au format de valeurs séparées par des virgules
Données au format STATA7
Données au format STATA
Données au format SAS
Données au format SPSS
Données au format de syntaxe SPSS
Autres

Autres vocabulaires contrôlés

Sujet (Borealis)

Sciences de l'agriculture
Arts et sciences humaines
Astronomie et astrophysique
Affaires et gestion
Chimie
Informatique et science de l'information
Sciences de la terre et de l'environnement
Génie
Droit
Sciences mathématiques
Médecine, santé et sciences de la vie
Physique
Sciences sociales
Autre

Couverture géographique

État / Province

Ville

Use [GeoNames](#) for correct spelling and avoid abbreviations.

Pays / Nation

Use ISO-3166.

Mot-clé

Servez-vous du [Thésaurus des sujets de base du gouvernement du Canada](#).

Collaborateur — Type (Borealis)

Chargé de la collecte de données	Person/institution responsible for finding or gathering/collecting data under the guidelines of the author(s) or Principal Investigator (PI).
Curateur des données	Person tasked with reviewing, enhancing, cleaning, or standardizing metadata and the associated data submitted for storage, use, and maintenance within a data centre or repository.
Gestionnaire des données	Person (or organization with a staff of data managers, such as a data centre) responsible for maintaining the finished resource.
Éditeur	A person who oversees the details related to the publication format of the resource Note: if the Editor is to be credited in place of multiple creators, the Editor's name may be supplied as Creator, with "(Ed.)" appended to the name.
Bailleur de fonds	Information about financial support (funding) for the resource being registered.
Établissement hôte	Typically, the organization that allows a resource to be available on the internet through the provision of its hardware/software/operating support.
Chef de projet	Person officially designated as head of project team or sub- project team instrumental in the work necessary to development of the resource.
Gestionnaire de projet	Person officially designated as manager of a project. Project may consist of one or many project teams and sub-teams.
Membre du projet	Person on the membership list of a designated project/project team.
Personne liée	A person without a specifically defined role in the development of the resource, but who is someone the author wishes to recognize.
Chercheur	A person involved in analyzing data or the results of an experiment or formal study. May indicate an intern or assistant to one of the authors who helped with research but who was not so "key" as to be listed as an author.
Groupe de recherche	Typically refers to a group of individuals with a lab, department, or division that has a specifically defined focus of activity.
Détenteur de droits	Person or institution owning or managing property rights, including intellectual property rights over the resource.
Commanditaire	Person or organization that issued a contract or under the auspices of which a work has been written, printed, published, developed, etc.
Superviseur	Designated administrator over one or more groups/teams working to produce a resource, or over one or more steps of a development process.
Chef d'un bloc de tâches	A Work Package is a recognized data product, not all of which is included in publication. The package, instead, may include notes, discarded documents, etc. The Work Package Leader is responsible for ensuring the comprehensive contents, versioning, and availability of the Work Package during the development of the resource.
Autre	Any person or institution making a significant contribution to the development and/or maintenance of the resource, but whose contribution is not adequately described by any of the other values for contributorType.

Contributor Type table adapted from *Dataverse* and [DataCite documentation](#).

Vocabulaires contrôlés par DDI

The DDI Alliance maintains a set of controlled vocabularies for use with the DDI standard, some of which have multilingual versions available. The latest French versions of these controlled vocabularies can be found through the [CESSDA Vocabulary Service](#).

Format général des données

Méthode temporelle

Mode de collecte

Type d'instrument

Unité d'analyse

Procédure d'échantillonnage

Appendix B: Glossary

Term	Definition
Bilingual	Related to more than one language. Bilingual features in Odesi include metadata creation, search, browsing and navigation in French and English, both of which are official languages in Canada.
Borealis	Borealis, the Canadian Dataverse Repository, is a bilingual , multidisciplinary, secure, Canadian research data repository , supported by academic libraries and research institutions across Canada. Borealis supports open discovery , management, sharing, and preservation of Canadian research data .
Codebook	A codebook describes the contents, structure, and layout of a data collection. A well-documented codebook "contains information intended to be complete and self-explanatory for each variable in a data file". ⁵ → <i>For the DDI metadata schema, see entry for DDI-Codebook.</i>
Collection	A set of datasets (e.g. surveys or public opinion polls) grouped together under the same topic (e.g. landslides), subject (e.g. Geography), or institution (e.g. Queen's University).
Controlled Vocabulary	A list of standardized terminology, words, or phrases, used for content filtering, analysis, selection and retrieval. ⁶ The Odesi Topic Categories are an example of a controlled vocabulary.
Data Curation Tool (DCT)	The Data Curation Tool (DCT) is a Dataverse extension developed by Scholars Portal. The DCT allows data owners and curators to view summary statistics for variables and to create and edit variable-level metadata for any tabular file in a data set.
Data Explorer	The Data Explorer is an extension of Dataverse that provides a graphical user interface (GUI) which lists the variables in a tabular data file and allows users to search, chart, and conduct cross tabulation analysis.
Dataset	Any organized collection of data in a computational format, defined by a theme or category that reflects what is being measured/observed/monitored. The presentation of the data in the application is enabled through metadata . ⁷
Data repository	Repositories preserve, manage, and provide access to many types of digital materials in a variety of formats. Materials in online repositories are curated to enable search, discovery, and reuse. There must be sufficient control for the digital material to be authentic, reliable, accessible and usable on a continuing basis. ⁸
Dataverse	Dataverse is an open source repository software built to share, preserve, cite, explore, and analyze research data. It was developed by the Institute for Quantitative Social Sciences (IQSS) at Harvard University, and it has a vibrant international community of collaborators and contributors. → <i>see also entry for Borealis</i>
Dataverse North	A Canadian research library initiative created to develop a community of practice for

⁵ What is a Codebook? <https://www.icpsr.umich.edu/web/ICPSR/cms/1983>

⁶ Controlled Vocabulary. [CODATA RDM Terminology](#).

⁷ Dataset. CODATA RDM Terminology. <https://codata.org/rdm-terminology/dataset/>

⁸ Repository. CODATA RDM Terminology. <https://codata.org/rdm-terminology/repository/>

Term	Definition
	libraries using or interested in using Dataverse as a repository for research data in Canada.
DDI	The Data Documentation Initiative (DDI) is an international standard for describing statistical and social science data. Documenting data with DDI facilitates interpretation and understanding – both by humans and computers. The freely available international DDI standard describes data that result from observational methods in the social, behavioral, economic, and health sciences. The DDI standard has two separate versions, DDI-Codebook (up to version 2.x) and DDI-Lifecycle (version 3.x). Odesi metadata implements DDI-Codebook (v. 2.x).
DDI-Codebook	(also DDI-C, v. 2.x) is a light-weight metadata schema maintained by the DDI Alliance , which Odesi has adopted for its dataset markup. It is intended primarily to document simple survey data.
Discovery	Data discovery is the descriptive and technical processing of data and metadata , as well as the tools and infrastructure aimed at improving access and reuse of research data on the web. ⁹ This process enables centralized searching and query by end users.
Facet	Facets are metadata fields that are indexed (→ see indexing) to facilitate information retrieval. Facets usually reflect a set of categories for classification, which may serve as access points for users to discover relevant resources. In Odesi, facets include metadata fields that apply controlled vocabularies . Odesi facets allow users to narrow down a search or to filter Odesi search results using specific criteria (such as year, series, data producer, geographic coverage or topic category).
FAIR	The FAIR Principles are a framework that promotes the Findability, Accessibility, Interoperability, and Reuse of digital assets. ¹⁰
Harvester	A harvester collects and aggregates metadata (and sometimes also data) that is held in external repositories (e.g. Borealis) to make it available to the user via the Odesi Search platform.
Indexing	In a database, an index is a data structure that helps organize the information in a way that supports faster and more precise information discovery and retrieval. → see also entry for Facet
Markt! Program	The Markt! Program is the collective of individuals from OCUL libraries who maintain the Odesi data collections , including marking up and adding DDI metadata to datasets hosted in Odesi.
MarkLogic	NoSQL database server used to house, index and serve Odesi metadata to the Odesi Search site, to power data search and discovery in Odesi.
Markup	In the Odesi context, markup is the product and process of preparing datasets for publication using metadata standards , schemas and formats, and following the best practices adopted by Odesi. The markup process includes preparing data files for importing into Borealis, digitizing and converting data files to various formats, adding metadata tags to pre-determined fields and publishing datasets to Odesi using Borealis ,

⁹ Barsky, E., Brosz, J., & Leahey, A. (2016). Research Data Discovery and the Scholarly Ecosystem in Canada : A White Paper. doi:<http://dx.doi.org/10.14288/1.0307548>

¹⁰ Wilkinson, Mark D. et. al. (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. Scientific Data 3 (1):160018. doi:10.1038/sdata.2016.18.

Term	Definition
	the Canadian Dataverse Repository.
Metadata	<p>The documentation that accompanies and assists users in the interpretation of different types of data. Metadata may describe context, content and structure of digital and non-digital resources, and their management through time. In Odesi, metadata typically includes information about a study's context, methodology, its variables, and its related data and documentation files.</p> <p>→ see also entries for Metadata profile, Metadata schema, and Metadata standard</p>
Metadata profile	<p>A metadata profile is a subset of metadata elements and attributes used for a specific purpose by a specific community or organization. This guide describes the Odesi DDI profile, that is, the subset of DDI-Codebook metadata fields used in and supported by the Odesi infrastructure.</p>
Metadata schema	<p>Metadata schemas define the structure, usage and implementation of a set of rules to describe and document digital resources (such as datasets and files) in a human- and machine-readable format. A metadata schema that has been adopted by an organized user community (e.g. DDI Alliance) and/or validated by a standard-setting organization (e.g. ISO, OGC) is referred to as a metadata standard.</p>
Metadata standard	<p>A metadata standard is a schema (→ metadata schema) that has been adopted by an organized user community and/or validated by a standard-setting organization. Examples of metadata standards include DDI Codebook (metadata for social research data), ISO 15119 (geospatial resource metadata), and DataCite (metadata for research data).</p>
Nesstar	<p>Nesstar is a suite of software used by Scholars Portal from 2009-2024. That software was created by the Norwegian Social Science Data Services (also known as NSD). That software reached its end of life in 2022. The Nesstar suite includes Nesstar Server, Nesstar WebView and Nesstar Publisher.</p>
Nesstar Publisher	<p>Nesstar Publisher was an advanced data management program consisting of data and metadata conversion and editing tools for publishing data and its associated documentation.¹¹ As of 2023, Odesi will discontinue use of the Nesstar Publisher (see Nesstar) and all Odesi dataset markup and publishing will take place in Borealis, the Canadian Dataverse Repository and the Data Curation Tool.</p>
OCUL	<p>OCUL is a consortium of Ontario's 21 university libraries. Our strength lies in our commitment to work together to maximize our collective expertise and resources. We enhance information services in Ontario and beyond through collective purchasing and shared digital information infrastructure, collaborative planning, advocacy, assessment, research, partnerships, communications, and professional development.</p>
Odesi	<p>Odesi is a digital repository for social science data, including polling data. It is a web-based data exploration, extraction and analysis tool that serves dataset metadata in the DDI standard. (Spelling note: Use title case. Outdated spellings and typography <odesi> or ODESI are no longer in use).</p>
Odesi Search	<p>Odesi Search is a search engine developed at Scholars Portal. That interface lets the user search in our collections for surveys and public opinion polls.</p>

¹¹ Nesstar Publisher <https://ddialliance.org/tool/nesstar-publisher>

Term	Definition
Preservation	Preservation metadata is item-level information that describes the context and structure of a digital object. It provides background details pertaining to a digital object's provenance, authenticity, and environment. Preservation metadata is a specific type of metadata that works to maintain a digital object's viability while ensuring continued access by providing contextual information, usage details, and rights.
Research data	Research data refers to data that has been collected, observed, generated or created to validate original research findings. Although usually digital, research data also includes non-digital formats such as laboratory notebooks and diaries.
Scholars Portal	Scholars Portal is a service of the Ontario Council of University Libraries. Founded in 2002, Scholars Portal provides a shared technology infrastructure and shared collections for all 21 university libraries in the province. Through the Scholars Portal online services, Ontario's university students, faculty and researchers have access to an extensive collection of e-journals, e-books, social science and geospatial data. Scholars Portal also supports the online interlibrary loan platform for Ontario's universities, a virtual chat reference service, and other tools designed to aid and enhance academic research in Ontario.

References

- Barrett, Kathryn, Kara Handren, and Sabina Pagotto. 2020. La Francisation de Scholars Portal : Progrès Vers Le Bilinguisme, Deuxième Partie. *Open Shelf*.
<https://open-shelf.ca/200406-la-francisation-de-scholars-portal-progres-vers-le-bilinguisme-deuxieme-partie/>.
- Barsky, Eugene, John Brosz, and Amber Leahey. 2016. Research Data Discovery and the Scholarly Ecosystem in Canada : A White Paper. Ottawa : Canadian Association of Research Libraries.
doi:[10.14288/1.0307548](https://doi.org/10.14288/1.0307548).
- Bascik, Teresa, Philippe Boisvert, Alexandra Cooper, Martine Gagnon, Mark Goodwin, John Huck, Amber Leahey, Kelly Stathis, and Michael Steeleworthy. 2021. Dataverse North Metadata Best Practices Guide v 3.0. Zenodo. doi:[10.5281/ZENODO.5576411](https://doi.org/10.5281/ZENODO.5576411).
- DataCite. 2023. Create DOIs. *DataCite*. <https://datacite.org/create-dois/>.
- DataCite Metadata Working Group. 2021. DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs v4.4. Application/pdf. DataCite, 82 pages.
doi:[10.14454/3W3Z-SA82](https://doi.org/10.14454/3W3Z-SA82).
- Dataverse. 2023a. About The Project. <https://dataverse.org/about>.
- Dataverse. 2023b. Dataverse Documentation v. 6.0. *Dataverse.Org*.
<https://guides.dataverse.org/en/latest/>.
- DDI Alliance. 2014. DDI-Codebook 2.5 XML Schema Documentation.
https://ddialliance.org/Specification/DDI-Codebook/2.5/XMLSchema/field_level_documentation.html.
- DDI Alliance. n.d. Nesstar Publisher. *Data Documentation Initiative*.
<https://ddialliance.org/tool/nesstar-publisher>.
- DDI Controlled Vocabularies Group (CVG) and DDI Alliance. 2023. Controlled Vocabularies. *Data Documentation Initiative*. <https://ddialliance.org/controlled-vocabularies>.
- DDI Glossary Working Group and DDI Alliance. 2023. DDI Glossary (Draft). *Data Documentation Initiative*. <https://ddi-alliance.atlassian.net/wiki/spaces/DDI4/pages/2790359041/Glossary+work>.
- Fry, Jane, Alexandra Cooper, Susan Mowers, and Carys Carrington. 2019. Best Practices Document, Version 3.1. Zenodo. doi:[10.5281/ZENODO.5237350](https://doi.org/10.5281/ZENODO.5237350).
- ICPSR. 2023. What Is a Codebook? ICPSR. <https://www.icpsr.umich.edu/web/ICPSR/cms/1983>.
- RDM Terminology WG and CODATA. 2021. Research Data Management Terminology. CODATA, The Committee on Data for Science and Technology.
<https://codata.org/initiatives/data-science-and-stewardship/rdm-terminology-wg/rdm-terminology/>.
- Vierkant, Paul and DataCite. 2023. What We Do. *DataCite*. <https://datacite.org/what-we-do/>.
- Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino Da Silva Santos, Philip E. Bourne, et al. 2016. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data* 3 (1):160018. doi:[10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18).