SECTION 2.0 - STUDY DESCRIPTION (Source Document, Data Collection)

The Study Description consists of information about the data collection, study, or compilation that the DDI-compliant documentation file describes. This section includes information about how the study should be cited, who collected or compiled the data, who distributes the data, keywords about the content of the data, summary (abstract) of the content of the data, data collection methods and processing, etc. Note that some content of the Study Description's Citation -- e.g., Responsibility Statement -- may be identical to that of the Documentation Citation. This is usually the case when the producer of a data collection also produced the print or electronic codebook for that data collection.

Source:DDI Codebook

DTD Numbers	Tags
2.0	<stdyDscr>
2.1	<citation>
2.1.1	<titlStmt>
2.1.1.1	<titl>
2.1.1.2	<subTitl>
2.1.1.3	<altTitl>
2.1.1.4	<parTitl>
2.1.1.5	<IDNo>
2.1.2	<rspStmt>
2.1.2.1	<AuthEnty>
2.1.2.2	<othID>
2.1.3	<prodStmt>
2.1.3.1	<producer>
2.1.3.2	<copyright>
2.1.3.3	<prodDate>
2.1.3.4	<prodPlace>
2.1.3.5	<software>
2.1.3.6	<fundAg>
2.1.3.7	<grantNo>
2.1.4	<distStmt>
2.1.4.1	<distrbtr>
2.1.5	<serStmt>
2.1.5.1	<serName>
2.1.5.2	<serInfo>
2.1.6	<verStmt>
2.1.6.1	<version>
2.1.6.2	<verResp>
2.1.6.3	<notes>
2.1.7	<biblCit>
2.1.8	<holdings>

2.2	<stdyInfo>	
2.2.1	<subject>	
2.2.1.1	<keyword>	
2.2.1.2	<topcClas>	
2.2.2		
2.2.3	<sumDscr>	
2.2.3.1	<timePrd>	
2.2.3.2	<collDate>	
2.2.3.3	<nation>	
2.2.3.4	<geogCover>	
2.2.3.5	<geogUnit>	
2.2.3.6	<anlyUnit>	
2.2.3.7	<universe>	
2.2.3.8	<dataKind>	
2.2.4	<notes>	
2.3	<method>	
2.3.1	<dataColl>	
2.3.1.1	<timeMeth>	
2.3.1.2	<dataCollector>	
2.3.1.3	<frequenc>	
2.3.1.4	<sampProc>	
2.3.1.6	<collMode>	
2.3.1.8	<source>	
2.3.1.9	<collSitu>	
2.3.1.12	<weight>	
2.3.2	<notes>	
2.3.3	<anlyInfo	
2.3.3.1	<respRate>	
2.4	<dataAccs>	
2.4.1	<setAvail>	
2.4.1.1	<accsPlac>	
2.4.2	<useStmt>	
2.4.2.3	<restrctn>	
2.4.2.4	<contact>	
2.4.2.5	<citReq>	
2.4.2.7	<conditions>	
2.4.2.8	<disclaimer>	
2.5	<othrStdyMat>	
2.5.1	<relMat>	
2.5.2	<relStdy>	
2.5.3	<relPubl>	

Description of tags and working examples

2.0   <stdyDscr>   Study Description

- Mandatory
- Repeatable

- Attributes: ID, xml:lang, source, access

Description: The Study Description consists of information about the data collection, study, or compilation that the DDI-compliant documentation file describes. This section includes information about how the study should be cited, who collected or compiled the data, who distributes the data, keywords about the content of the data, summary (abstract) of the content of the data, data collection methods and processing, etc. Note that some content of the Study Description's Citation -- e.g., Responsibility Statement -- may be identical to that of the Documentation Citation. This is usually the case when the producer of a data collection also produced the print or electronic codebook for that data collection

2.1 <citation> Citation

- Mandatory
- Repeatable
- Attributes: ID, xml:lang, source, MARCURI

Description: Citation for the data collection described by the marked-up documentation. This element encodes the bibliographic information describing the data collection, including title information, statement of responsibility, production and distribution information, series and version information, text of a preferred bibliographic citation, and notes (if any). A MARCURI attribute is provided to link to the MARC record for this citation.

2.1.1 <titlStmt> Title Statement

- Mandatory
- Not Repeatable
- Attributes: ID, xml:lang, source

Description: Title statement for the data collection. Contains the following sub-elements:

2.1.1.1 <titl> Title

- Mandatory
- Not Repeatable
- Attributes: ID, xml:lang, source

Description: Contains the full authoritative title of the data collection. The data collection title will in most cases be identical to the title for the marked-up document (1.1.1.1) and the source document (1.4.1.1). A full title should indicate the geographic scope of the data collection as well as the time period covered.

Source: STC Online Catalogue, BiblioCat (STC's Library Catalogue)

Example:

<titl>**Survey of Household Spending, 2002 [Canada]**</titl>

<titl>**Survey of Consumer Finances, 1998 [Canada]: Census Families, 1997 Income**</titl>

<titl>**Domestic Violence Experience in Omaha, 1986-1987 [Nebraska, USA]**</titl>

<titl>**Youth Smoking Survey, 2002 [Ontario, Canada]**</titl>

<titl>**Canadian Community Health Survey, Cycle 3.1, HSS1 File, 2005**</titl>

2.1.1.2 <subTitl>   Subtitle

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: A subtitle is a secondary title used to amplify or state certain limitations of the main title. It may repeat information already in the main title.

Source: STC Online Catalogue, BiblioCat (STC's Library Catalogue)

Example:

<subTitl>**Census Families, 1997 Income** <subTitl>

<subTitl>**HSS1 File, Health Utility Index (HUI), Fruit and Vegetable Consumption and Labour Force (long form), 2005** <subTitl>

2.1.1.3 <altTitl>   Alternate Title

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: The alternative title may be the title by which a data collection is commonly referred to or it may be an abbreviation for the title.

Source: STC Online Catalogue, BiblioCat (STC's Library Catalogue)

Example:

<altTitl>**SHS 2001**</altTitl>

<altTitl>**SCF 1998**</altTitl>

<altTitl>**CCHS, 3.1, HSS1, 2005**</altTitl>

2.1.1.4 <parTitl>   Parallel Title

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: The title translated into another language.

Source: STC Online Catalogue, BiblioCat (STC's Library Catalogue)

Example:

<parTitl>**Enquête sur les dépenses des ménages, 2002 [Canada]**</parTitl>

2.1.1.5 <IDNo>   Identification Number

- Optional
- Repeatable

- Attributes: ID, xml:lang, source, agency

Description: Unique string or number (producer's or archive's number) for the data collection. An "agency" attribute is supplied.

Source: STC Online Catalogue, BiblioCat (STC's Library Catalogue)

Example:

<IDNo agency="**Statistics Canada**">**62M0004XCB_2001_E**</IDNo>

<IDNo agency="**Statistics Canada**">**82M0020XCB_2004_Annual_HH_E**</IDNo>

<IDNo agency="**Statistics Canada**">**82M0013XCB_2004_Cycle2.1_E**</IDNo>

2.1.2   \<rspStmt\>   Responsibility Statement

- Optional
- Not Repeatable
- Attributes: ID, xml:lang, source

Description: Responsibility for the data collection. Contains the following sub-elements:

2.1.2.1 \<AuthEnty\>   Author Entity

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, affiliation

Description: 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 attribute if available. Invert the first and last names and use commas.

Example:

<AuthEnty affiliation="**Special Surveys Division, Statistics Canada" Boyle, Eve**</AuthEnty>

2.1.2.2 \<othID\>   Other Identification

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, type, role, affiliation

Description: 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. For example, the name of the person who cleaned the data collection might be cited here, using the role and affiliation attributes.

Example:

<othId affiliation="**SSD**">**Guy, Dominique**</othId>

2.1.3   <prodStmt>   Production Statement

- Optional
- Not Repeatable
- Attributes: ID, xml:lang, source

Description: Production statement for the data collection. Contains the following elements:

2.1.3.1 <producer>   Producer

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, abbr, affiliation, role

Description: 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. Use the role attribute to distinguish different stages of involvement in the production process, such as original producer.

Example:

<producer abbr="**ICPSR**" affiliation="**Institute for Social Research**">**Inter-university Consortium for Political and Social Research**</producer>

<producer abbr="**SSD**" affiliation="**Special Surveys Division**">**Statistics Canada**</producer>

2.1.3.2 <copyright>   Copyright

- Optional
- Not Repeatable
- Attributes: ID, xml:lang, source

Description: Copyright statement for the data collection. Copyright for data collection (2.1.3.2) maps to Dublin Core Rights. Inclusion of this element is recommended element.

Source: IMDB

Example:

<copyright>**Copyright© Statistics Canada, Data Liberation Initiative, 2002**</copyright>

2.1.3.3 <prodDate>   Production Date

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, date

Description: Date the data collection was produced (not distributed or archived). The ISO standard for dates (YYYY-MM-DD) is recommended for use with the date attribute. Production date for data collection (2.1.3.3) maps to Dublin Core Date element.

Example:

<prodDate date="**2003-11-31**">**November 31, 2003**</prodDate>

2.1.3.4 <prodPlac>   Production Place

- Optional
- Repeatable
- Attributes: ID, xml:lang, source,

Description: Address or the archive or agency that produced the data collection.

Example:

<prodPlac>**Ottawa, Ontario: Carleton University**</prodPlac>

<prodPlac>**Ottawa, Ontario: Statistics Canada**</prodPlac>

2.1.3.5 <software>   Software

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, version, date

Description: Identifies the software used in creating or storing the data collection. A "version" attribute permits specification of the software version number. The "date" attribute is provided to enable specification of the date (if any) for the software release. The ISO standard for dates (YYYY-MM-DD) is recommended for use with the date attribute.

Example:

<software version="1.0">**Arbortext Adept Editor**</software>

2.1.3.6 <fundAg>   Funding Agency

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, abbr, role

Description: The source(s) of funds for production of the data collection. If different funding agencies sponsored different stages of the production process, use the role attribute to distinguish them.

Example:

<fundAg>**Environics, [Ottawa, Canada]**</fundAg>

2.1.3.7 <grantNo>   Grant Number

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, agency, role

Description: The grant/contract number of the project that sponsored the data collection effort. If more than one, indicate the appropriate agency using the "agency" attribute. If different funding agencies sponsored different stages of the production process, use the "role" attribute to distringuish the grant numbers.

Example:

<grantNo>**Award for the Prettiest Research, Grant #28745B**</grantNo>

2.1.4   <distStmt>   Distributor Statement

- Optional
- Not Repeatable
- Attributes: <u>ID, xml:lang, source</u>

Description: Distribution statement for the data collection.

2.1.4.1 <distrbtr>   Distributor

- Optional
- Repeatable
- Attributes: <u>ID, xml:lang, source</u>, abbr, affiliation, URI

Description: 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. A URI attribute is included to provide an URN or URL to the ordering service or download facility on a website.

Example:

<distrbtr abbr="**DLI**" affiliation="**Statistics Canada**"
URI="**http://www.statcan.ca/english/Dli/dli.htm**">**Data Liberation Initiative**</distrbtr>

2.1.5   <serStmt>   Series Statement

- Optional
- Not Repeatable
- Attributes: <u>ID, xml:lang, source</u>, URI

Description: Series statement for the data collection. The URI attribute is provided to point to a central Internet repository of series information.

2.1.5.1 <serName>   Name of Series

- Optional
- Repeatable
- Attributes: <u>ID, xml:lang, source</u>, abbr

Description: The name of the data series to which the collection belongs*.*

Source: IMDB, Microdata User Guide

Example:

<serName abbr="**SHS**"> **Survey of Household Spending** </serName>

2.1.5.2 <serInfo>   Series Information

- Optional
- Repeatable

- Attributes: ID, xml:lang, source

Description: Contains a history of the data series and a summary of those features that apply to the data series as a whole.

Source: IMDB, Microdata User Guide

Example:

<serInfo>**Conducted since 1997, the Survey of Household Spending provides Information about the spending habits, dwelling characteristics and household equipment of Canadian households. Annual from 1997 – 1999, it has become a biennial survey as of 2000**.</serInfo>

2.1.6   <verStmt>   Version Statement

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: Version statement for the data collection. Contains the following sub-elements:

2.1.6.1 <version>   Version

- Optional
- Not Repeatable
- Attributes: ID, xml:lang, source, date, type

Description: Also known as release or edition. If there have been substantive changes in the data collection since its creation, this statement should be used. The ISO standard for dates (YYYY-MM-DD) is recommended for use with the date attribute.

Example:

<version type="**edition**" date"="**2002-12-15**">**DLI Version 2.0**</version>

2.1.6.2 <verResp>   Version Responsibility

- Optional
- Not Repeatable
- Attributes: ID, xml:lang, source, affiliation

Description: Used to indicate the organization and person responsible for the version of the data collection.

Example:

<verResp>**Special Surveys Division**</verResp>

2.1.6.3 <notes>   Notes

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, type, subject, level, resp,

Description: Used to indicate additional information regarding the version or the version responsibility statement for the data collection, in particular to indicate what makes a new version different from its predecessor. "Notes" sections appear in several places in the DTD. The attributes for notes permit a controlled vocabulary to be developed (type and subject), the level of the DTD to which the note refers to be identified (study, file, variable, etc.), and the author of the note to be indicated (resp).

Example:

<notes resp="**Jane Doe**">**1997 Data files and documentation revised. The update standardizes the files with the 2000 file. Include data of latest change (YYYY-MM-DD) as well as full name of person which made the change.**</notes>

2.1.7   <biblCit>   Bibliographic Citation

- Optional
- Not Repeatable
- Attributes: ID, xml:lang, source, format

Description: Complete bibliographic reference containing all of the standard elements of a citation that can be used to cite the data collection. The "format" attribute is provided to enable specification of the particular citation style used, e.g. APA, MLA, Chicago, etc.

Example:

<biblCit format="**APA**">**Statistics Canada. 2004. *User Guide, Public-use Microdata File, Survey of Household Spending, 2002.* Statistics Canada Catalogue no. 62M0004XCB. Ottawa. 103 p. http://www.statcan.ca/english/Dli/Metadata/shs/2002/shs2002gid.pdf (accessed August 16, 2005).**</biblCit>

2.1.8   <holdings>   Holdings Information

- Optional
- Not Repeatable
- Attributes: ID, xml:lang, source, location, callno, URI

Description: Information concerning either the physical or electronic holdings of the cited work. Attributes include: location—The physical location where a copy is held; callno—The call number for a work at the location specified; and URI—A URN or URL for accessing the electronic copy of the cited work.

Example:

<holdings location="**DRC University of Guelph**" URI="**http://nesstar.tdr.uoguelph.ca/CONSUM/SHS/2002/DOCS/**">**DRC XML Codebook**</holdings>

2.2   <stdyInfo>   Study Scope

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: This section contains information about the data collection's scope across several dimensions, including substantive content, geography, and time.

2.2.1.   <subject>   Subject

- Optional
- Not Repeatable
- Attributes: <u>ID, xml:lang, source</u>

Description: Subject information describing the data collection's intellectual content.

2.2.1.1 <keyword>   Keywords

- Optional
- Repeatable
- Attributes: <u>ID, xml:lang, source</u>, vocab, vocabURI

Description: 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. A controlled vocabulary can be employed. Maps to Dublin Core Subject. The vocab attribute is provided for specification of the controlled vocabulary in use, e.g., LCSH, MeSH, etc. The vocabURI attribute specifies the location for the full controlled vocabulary.

Source: STC Online Catalogue, BiblioCat (STC's Library Catalogue)

Example:

```
<keyword>Accomodation</keyword>
<keyword>Computers</keyword>
<keyword>Data collection</keyword>
<keyword>Data processing</keyword>
<keyword>Educational attainment</keyword>
<keyword>Internet use</keyword>
<keyword>Lifestyle</keyword>
<keyword>Physical activity</keyword>
<keyword>Questionnaires</keyword>
<keyword>Social surveys</keyword>
```

2.2.1.2 <topcClas>   Topic Classification

- Optional
- Repeatable
- Attributes: <u>ID, xml:lang, source</u>, vocab, vocabURI

Description: The classification field indicates the broad substantive topic(s) that the data cover.

Source: STC Online Catalogue

Example:

<topcClas>**Mass Political Behaviour**</topcClas>

2.2.2   Abstract

- Optional

Example:

```
<timePrd event="start" date="1997-01-01">January 1, 1997</timePrd>
<timePrd event="end" date="1997-12-31">December 31, 1997</timePrd>
```

2.2.3.2 <collDate>   Date of Collection

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, date, event, cycle

Description: Contains the date(s) when the data were collected. Use the event attribute to specify "start", "end", or "single" for each date entered to distinguish between, for example, the first day of collection (start), only day of collection (single), and last day of collection (end). The ISO standard for dates (YYYY-MM-DD) is recommended for use with the date attribute. The "cycle" attribute permits specification of the relevant cycle, wave, or round of data. Maps to Dublin Core Coverage. Inclusion of this element in the codebook is recommended.

Source: IMDB, Microdata User Guide

Example:

```
<collDate event="start" date="1998-04-19">April 19, 1998</collDate>
<collDate event="end" date="1998-05-01">May 1, 1998</collDate>
```

2.2.3.3 <nation>   Country

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, abbr

Description: Indicates the country or countries covered in the file. Attribute "abbr" may be used to list common abbreviations; use of ISO country codes is recommended. Maps to Dublin Core Coverage element. Inclusion of this element is recommended.

Source: IMDB, Microdata User Guide

Example:

```
<nation abbr="CAN">Canada</nation>
```

2.2.3.4 <geogCover>   Geographic Coverage

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: Information on the geographic coverage of the data. Include the total geographic scope of the data, and any additional levels of geographic coding provided in the variables. Maps to Dublin Core Coverage. Inclusion of this element is recommended.

Source: IMDB, Microdata User Guide

Example:

<geogCover>**Canada**</geogCover>
<geogCover>**Provinces**</geogCover>
<geogCover>**Excluding Territories**</geogCover>

2.2.3.5 <geogUnit>   Geographic Unit

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: Lowest level of geographic aggregation covered by the data.

Source: IMDB, Microdata User Guide

Example:

<geogUnit>**Census Metropolitan Area**</geogUnit>

2.2.3.6 <analyUnit>   Unit of Analysis

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, unit

Description: Basic unit of analysis or observation that the file describes: individuals, families/households, groups, institutions/organizations, administrative units, etc. The "unit" attribute is included to permit the development of a controlled vocabulary for this element.

Source: IMDB, Microdata User Guide

Example:

<anlyUnit>**households**</anlyUnit>

<anlyUnit>**individuals**</anlyUnit>

2.2.3.7 <universe>   Universe

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, level, clusion

Description: 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. A "level" attribute is included to permit coding of the level to which universe applies, i.e., the study level, the file level (if different from study), or the variable level. The "clusion" attribute provides for specification of groups included (I) in or excluded (E) from the universe.

Source: IMDB, Microdata User Guide

Example:

<universe level="**study**" clusion="**I**"> **Included: The target population of the SHS is comprised of individuals living in private, Canadian households who are neither official representatives of foreign countries living in Canada and their families, nor residents of Indian reservations or Crown lands.**</universe>

<universe level="**study**" clusion="**E**">**Excluded: The following groups were excluded from the survey; those living on Indian reserves and crown lands; official representatives of foreign countries living in Canada and their families; members of religious and other communal colonies; members of the Canadian Armed Forces living in Military Camps; people living in residences for senior citizens; people living full time in institutions: for example, inmates of penal institutions and chronic care patients living in hospitals and nursing homes. The survey covers about 98% of the population in the ten provinces. In the Territories, coverage was restricted to 88% in the Yukon, 92% in the Northwest Territories, and 89% in Nunavut. Information was not gathered from persons temporarily living away from their families (for example, students at university), because it would be gathered from their families if selected. In this way, double counting of such individuals was avoided**</universe>

2.2.3.8 <dataKind> Kind of Data

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: 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.

Source: Microdata User Guide

Example:

<dataKind>**survey data**</dataKind >

2.2.4 <notes> Notes

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, type, subject, level, resp, sdatrefs

Description: Used to indicate additional information regarding the scope of a data collection. "Notes" sections appear in several places in the DTD. The attributes for notes permit a controlled vocabulary to be developed (type and subject), the level of the DTD to which the note refers to be identified (study, file, variable, etc.), and the author of the note to be indicated (resp).

Source: Microdata User Guide

Example:

<notes>**Data refer to previous year, known as the Reference Year.**</notes>

<notes>**Data on employment and income refer to the preceding year, although demographic data refer to the time of the survey.**</notes>

2.3   <method>   Methodology

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: This section describes the methodology and processing involved in a data collection.

2.3.1   <dataColl>   Data Collection Methodology

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: Information about the methodology employed in a data collection.

2.3.1.1 <timeMeth>   Time Method

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, method

Description: The time method or time dimension of the data collection. The "method" attribute is included to permit the development of a controlled vocabulary for this element.

Source: IMDB, Microdata User Guide

Example:

<timeMeth>**cross-section**</timeMeth>

<timeMeth>**pilot**</timeMeth>

<timeMeth>**longitudinal**</timeMeth>

<timeMeth>**time-series**</timeMeth>

<timeMeth>**trend study**</timeMeth>

<timeMeth>**panel survey**</timeMeth>

2.3.1.2 <dataCollector>   Data Collector

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, abbr, affiliation

Description: 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.

Source: IMDB, Microdata User Guide

Example:

<dataCollector>**For the CRIC surveys, Environics conducted some of them so the information for Environics would be entered here.**</dataCollector>

2.3.1.3 <frequenc>   Frequency of Data Collection

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, freq

Description: If the data collected include more than one point in time, indicate the frequency with which the data were collected. The "frequency" attribute is included to permit the development of a controlled vocabulary for this element.

Source: IMDB, Microdata User Guide

Example:

<frequenc>**Annual 1997-1999 Biennial 2000**</frequenc>

<frequenc>**Data is collected at only one point in time (per case).**</frequenc>

<frequenc>**occasional**</frequenc>

<frequenc>**one-time**</frequenc>

<frequenc>**annual**</frequenc>

<frequenc>**biennial**</frequenc>

<frequenc>**biannual**</frequenc>

2.3.1.4 <sampProc>   Sampling Procedure

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: 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.

Source: IMDB, Microdata User Guide

Example:

<sampProc>**Stratified random sample**</sampProc>

<sampProc>**Random digit dialling**</sampProc>

<sampProc>**Two stage selection process: stage 1 is the selection of households; and stage 2 is selection of respondents within each household.**</sampProc>

2.3.1.6 <collMode>   Mode of Collection

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: The method used to collect the data; instrumentation characteristics.

Source: IMDB, Microdata User Guide

Example:

<collMode>**paper questionnaire**</collMode>

<collMode>**computer assisted telephone interviews (CATI)**</collMode>

2.3.1.8 <sources>   Sources Statement

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: Description of sources used for the data collection.

2.3.1.9 <collSitu>   Characteristics of the Data Collection Situation

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: Used to describe noteworthy aspects of the data collection situation. Include information on factors such as cooperativeness of respondents, duration of interviews, number of call-backs, etc.

Source: IMDB, Microdata User Guide

Example:

<collSitu>**There were 1,194 respondents who answered questions in face-to-face interviews lasting approximately 75 minutes each.**</collSitu>

2.3.1.12   <weight>   Weighting

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: The use of sampling procedures may make it necessary to apply weights to produce accurate statistical results. Describe here the criteria for using weights in analysis of a collection. If a weighting formula or coefficient was developed, provide this formula, define its elements, and indicate how the formula is applied to data.

Source: IMDB, Microdata User Guide

Example:

<weight>**WEIGHT=Weight at household level** </weight>

<weight>**The principles behind the calculation of the weights for the HIUS are nearly identical to those for the LFS. However, this survey is a household-weighted survey, not a person-weighted survey. Also, further adjustments are made to the LFS sub-weights in order to derive a final weight for each record on the HIUS microdata file. 1) An adjustment to account for the use of a four-sixths sub-sample, instead of the full LFS sample. 2) An adjustment to account for the additional non-response to the supplementary survey, i.e., households that did not respond to the HIUS but did respond to the LFS (or for which previous month's LFS data was brought forward). Statistical techniques are used to group together records that are similar in terms of demographic variables obtained from LFS responses. The adjustment is made separately within all non-response groups created for each province. Household Internet Use Survey, 2003 – user guide 3) The final adjustment ensured that estimates produced for a province-household size group would agree with the known population totals for that province-household size group. The adjustments were made for household size groupings of one person, two people and three or more people. The resulting weight (WTHM) is the final weight which appears on the HIUS microdata file.** </weight>

2.3.2   <notes>   Notes

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, type, subject, level, resp

Description: Use to indicate any additional information about the methodology and processing involved in the collection.

Example:

<notes resp="Jane Smith">**Additional information on derived variables has been added to this marked-up version of the documentation.**</notes>

2.3.3   <anlyInfo>   Data Appraisal Information

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: Information on data appraisal.

2.3.3.1 <respRate>   Response Rates

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: The percentage of sample members who provided information.

Source: IMDB, Microdata User Guide

Example:

<respRate>**The household response rate for the CTUMS, Cycle 1, conducted from February to June 2004 was 80.6%. The person response rate was 85.9%.**</respRate>

<respRate>**The overall response rate was 82%, although retail firms with an annual sales volume of more that $5,000,000 were somewhat less likely to respond.**</respRate>

2.4   <dataAccs>   Access to Data

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: This section describes access conditions and terms of use for the data collection. In cases where access conditions differ across individual files or variables, multiple access conditions can be specified. The access conditions applying to a study, file, variable group, or variable can be indicated by an IDREF attribute on the study (2.0), file (3.0), variable group (4.1), or variable (4.2) elements called "access".

2.4.1   <setAvail>   Data Collection Availability

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, media

Description: Information on availability and storage of the collection. The "media" attribute may be used in combination with any of the sub-elements. See Location of Data Collection.

Example:

<setAvail media="**On-line subscription service**">

2.4.1.1 <accsPlac>   Location of Data Collection

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, URI

Description: Location where the data collection is currently stored. Use the URI attribute to provide a URN or URL for the storage site or the actual address from which the data may be downloaded.

Example:

<accsPlac URI="**http://new.sourceoecd.org/**">**Access SourceOECD Subscription Service Here**</accsPlac>

2.4.2   <useStmt>   Use Statement

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: Information on terms of use for the data collection.

2.4.2.3 <restrctn>   Restrictions of Use

- Optional
- Not Repeatable
- Attributes: ID, xml:lang, source

Description: 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.

Example:

<restrctn>**The data may be used by current students, faculty and staff of the University of Guelph, the University of Waterloo and Wilfrid Laurier University for academic and personal research only.**</restrctn>

<restrctn>**For academic use only, that is, for teaching and non-commercial research. Data may not be re-disseminated. Other data users please contact CRIC.**</restrctn>

<restrctn>**For members of the Data Liberation Initiative community.**</restrctn>

2.4.2.4 <contact>   Access Authority

- Optional
- Not Repeatable
- Attributes: ID, xml:lang, source, affiliation, URI, email

Description: Contact person or organization (with full address and telephone number, if available) that controls access to a collection, if different from the data distributor.

Example:

<contact affil="**Statistics Canada**" URI=" **http://www.statcan.ca/english/Dli/dli.htm**> email="**hardlin@statcan.ca**"> **Linda Harding-DeVries, 613-951-4575**</contact>

2.4.2.5 <citReq>   Citation Required

- Optional
- Not Repeatable
- Attributes: ID, xml:lang, source

Description: Text of requirement that a data collection should be cited properly in articles or other publications that are based on analysis of the data.

Example:

<citReq>**All publications using CRIC data should identify the Centre for Research and Information on Canada as author, the respective survey title, as well as the year.**</citReq>

<citReq>**All publications using Statistics Canada data should identify Statistics Canada as author, the respective survey title, as well as the year.**</citReq>

2.4.2.7 <conditions>   Conditions of Use

- Optional
- Not Repeatable
- Attributes: ID, xml:lang, source

Description: Indicate any additional information that will assist the user in understanding the access and use conditions of the data collection.

Example:

<conditions> **Data Liberation Initiative Community**
 <ExtLink URI="**http://www.statcan.ca/english/Dli/caselaw/assess.htm**/>
 </conditions>

2.4.2.8 <disclaimer>   Disclaimer

- Optional
- Not Repeatable
- Attributes: ID, xml:lang, source

Description: Information regarding responsibility for uses of the data collection.

Example:

<disclaimer>**The original collector of the data, CRIC and Carleton University bear no responsibility for uses of this collection, or the interpretations or inferences based upon such uses**.</disclaimer>

<disclaimer>**The original collector of the data, Statistics Canada, bears no responsibility for uses of this collection, or the interpretations or inferences based upon such uses.**</disclaimer>

2.5 <othrStdyMat>   Other Study Description Materials

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: Other materials relating to the study description. This section describes other materials that are related to the study description that are primarily descriptions of the content and use of the study, such as appendices, sampling information, weighting details, methodological and technical details, publications based upon the study content, related studies or collections of studies, etc. This section may point to other materials related to the description of the study through use of the generic citation element (A.6), which is available for each element in this section. <othrStdyMat> maps to Dublin Core Relation element.

Note that Section 5.0, Other Study-Related Materials, should be used for materials used in the production of the study or useful in the analysis of the study. The materials in Section 5.0 may be entered as PCDATA (ASCII text) directly into the document (through use of the txt element). That section may also serve as a "container" for other electronic materials by providing a brief description of the study-related materials accompanied by the "type" and "level" attributes further defining the materials. Other Study-Related Materials in Section 5.0 may include: questionnaires, coding notes, SPSS/SAS/Stata setup files (and others), user manuals, continuity guides, sample computer software programs, glossaries of terms,

interviewer/project instructions, maps, database schema, data dictionaries, show cards, coding information, interview schedules, missing values information, frequency files, variable maps, etc.

2.5.1 <relMat>   Related Materials

- Optional
- Repeatable
- Attributes: ID, xml:lang, source, callno, label, media, type

Description: 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. This element can contain either PCDATA or a citation or both, and there can be multiple occurrences of both the citation and PCDATA within a single element. 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).

Example(s):

<relMat> **Full details on the research design and procedures, sampling methodology, content areas, and questionnaire design, as well as percentage distributions by respondent's sex, race, region, college plans, and drug use, appear in the annual ISR volumes MONITORING THE FUTURE: QUESTIONNAIRE RESPONSES FROM THE NATION'S HIGH SCHOOL SENIORS**.</relMat>

<relMat>**Current Population Survey, March 1999: Technical Documentation includes an abstract, pertinent information about the file, a glossary, code lists, and a data dictionary. One copy accompanies each file order. When ordered separately, it is available from Marketing Services Office, Customer Service Center, Bureau of the Census, Washington, D.C. 20233**. </relMat>

<relMat>**A more precise explanation regarding the CPS sample design is provided in Technical Paper 40, The Current Population Survey: Design and Methodology. Chapter 5 of this paper provides documentation on the weighting procedures for the CPS both with and without supplement questions**.</relMat>

2.5.2 <relStdy>   Related Studies

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: 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.

Example(s):

<relStdy>**ICPSR distributes a companion study to this collection titled FEMALE LABOR FORCE PARTICIPATION AND MARITAL INSTABILITY, 1980: [UNITED STATES] (ICPSR 9199).**</relStdy>

2.5.3 <relPubl>   Related Publications

- Optional
- Repeatable
- Attributes: ID, xml:lang, source

Description: Bibliographic and access information about articles and reports based on the data in this collection. Can take the form of bibliographic citations.

Example(s):

<relPubl>**Economic Behavior Program Staff. SURVEYS OF CONSUMER FINANCES. Annual volumes 1960 through 1970. Ann Arbor, MI: Institute for Social Research**.</relPubl>

<relPubl>**Data from the March Current Population Survey are published most frequently in the Current Population Reports P- 20 and P- 60 series. These reports are available from the Superintendent of Documents, U. S. Government Printing Office, Washington, DC 20402. They also are available on the INTERNET at http:// www. census. gov. Forthcoming reports will be cited in Census and You, the Monthly Product Announcement (MPA), and the Bureau of the Census Catalog and Guide**.</relPubl>

Example of Complete DDI compliant codebook for Section 2.0 Study Description

```
<stdyDscr>
- <citation>
- <titlStmt>
  <titl>Survey of Household Spending,  2002 [Canada]</titl>
  <subTitl />
  <altTitl>SHS 2002</altTitl>
  <parTitl>Enquêtes sur les dépenses des ménages, 2002 [Canada]</parTitl>
  <IDNo agency="Statcan">62M0004-XCB 2002</IDNo>
     </titlStmt>
- <rspStmt>
  <AuthEnty>Income Statistics Division, Statistics Canada</AuthEnty>
  </othID>
     </rspStmt>
- <prodStmt>
  <producer>Statistics Canada</producer>
  <copyright>Copyright© Statistics Canada, Data Liberation Initiative, 2003</copyright>
  <prodDate date="2003-11-31">November 31, 2003</prodDate>
  <prodPlace>Ottawa, Ontario</prodPlace>
  <software version="2.1 Beta 9" date="2003-11-14">Nesstar Publisher</software>
  <fundAg>Centre for Research and Information on Canada (CRIC), [Ottawa, Canada]</fundAg>
  <grantNo>Award for the Prettiest Research, Grant #28745B</grantNo>
     </prodStmt>
- <distStmt>
  <distrbtr abbr="DLI" affiliation="Statistics Canada" URI="http://www.statcan.ca/english/Dli/dli.htm">Data Liberation Initiative</distrbtr>
     </distStmt>
- <serStmt>
  <serName>Survey of Household Spending</serName>
  <serInfo>Conducted since 1997, the SHS provides information about the spending habits, dwelling characteristics and household equipment of
     Canadian households.  Annual from 1997-1999, it has become a biennial survey as of 2000.</serInfo>
     </serStmt>
- <verStmt>
  <version>Version 1.0</version>
  <verResp />
  <notes />
```

```
    </verStmt>
    <biblCit format="APA">Statistics Canada. 2003.  Survey of Household Spending  [public-use microdata file].  Ottawa, Ontario: Statistics Canada.
        Income Statistics Division [producer and distributor].
    </biblCit>
    <holdings />
    </citation>
- <stdyInfo>
- <subject>
  <keyword>Accommodation costs</keyword>
  <keyword>Air conditioners</keyword>
  <keyword>Alcoholic beverages</keyword>
  <keyword>Audiovisual materials</keyword>
  <keyword>Bathrooms</keyword>
  <keyword>Cable television</keyword>
  <keyword>Cellular telephones</keyword>
  <keyword>Child care needs</keyword>
  <keyword>Clothing</keyword>
  <keyword>Communications</keyword>
  <keyword>Compact disc players</keyword>
  <keyword>Computers</keyword>
  <keyword>Dishwashers</keyword>
  <keyword>Dwelling characteristics</keyword>
  <keyword>Dwellings</keyword>
  <keyword>Education</keyword>
  <keyword>Electrical appliances</keyword>
  <keyword>Expenditures</keyword>
  <keyword>Food</keyword>
  <keyword>Freezers</keyword>
  <keyword>Fuel costs</keyword>
  <keyword>Furniture and fixtures</keyword>
  <keyword>Head of household</keyword>
  <keyword>Health care</keyword>
  <keyword>Heating systems</keyword>
  <keyword>Household expenditures</keyword>
  <keyword>Insurance payments</keyword>
  <keyword>Internet</keyword>
  <keyword>Internet use</keyword>
```

```
<keyword>Motor vehicles</keyword>
<keyword>Pensions</keyword>
<keyword>Period of construction</keyword>
<keyword>Property tenure</keyword>
<keyword>Reading materials</keyword>
<keyword>Recreation</keyword>
<keyword>Refrigerators</keyword>
<keyword>Social surveys</keyword>
<keyword>Surveys</keyword>
<keyword>Taxes;</keyword>
<keyword>Telephones</keyword>
<keyword>Televisions</keyword>
<keyword>Tobacco products</keyword>
<keyword>Transportation services</keyword>
<keyword>Travel</keyword>
<keyword>Video recorders</keyword>
<keyword>Washing machines and dryers</keyword>
<topcClas>Mass Political Behaviour</topcClas>
   </subject>
<abstract>This public-use microdata file presents data from the 2001 Survey of Household Spending (SHS) conducted in January through March
   2002. Information about the spending habits, dwelling characteristics and household equipment of Canadian households during 2001 was
   obtained by asking people in the ten provinces and the three territories to recall their expenditures for the previous calendar year (spending
   habits) or as of December 31 (dwelling characteristics and household equipment).</abstract>
<abstract>Conducted since 1997, the Survey of Household Spending integrates most of the content found in the Family Expenditure Survey and
   the Household Facilities and Equipment Survey. Many data from these two surveys are comparable to the Survey of Household Spending
   data. However, some differences related to methodology, to data quality and to definitions must be considered before comparing these data.
   See User Guide Section 1.1.4., for further information.</abstract>
<abstract>New for 2001: For 2001 only, extra questions were included in the survey so that data from the SHS could be used in the weighting of the
   Consumer Price Index. This change may affect historical comparisons for a few variables. See User Guide section 4.2.4., Comparability Over
   Time. There were also new questions added to the 2001 SHS. The expenditure questions covered yarn and thread, regular and other leasing
   fees for vehicles, and cable and satellite services. Questions were also added about the presence in the home of a satellite dish, DVD player or
   CD writer, and the type of Internet connection.</abstract>
<sumDscr>
 <timePrd date="2001-01-01" event="start">January 01, 2001</timePrd>
 <timePrd date="2001-12-31" event="end">December 31, 2001</timePrd>
 <collDate date="2002-01-01" event="start">January 01, 2002</collDate>
 <collDate date="2002-03-31" event="end">March 31, 2002</collDate>
```

```xml
<nation>Canada</nation>
<geogCover>Canada, Provinces and Territories</geogCover>
<geogUnit>Province, Territory</geogUnit>
<anlyUnit>Household</anlyUnit>
<universe level="study" clusion="I">  Included: The target population of the SHS is comprised of individuals living in private, Canadian households
    who are neither official representatives of foreign countries living in Canada and their families, nor residents of Indian reservations or Crown
    lands.</universe>
<universe level="study" clusion="E"> Excluded: The following groups were excluded from the survey; those living on Indian reserves and crown
    lands; official representatives of foreign countries living in Canada and their families; members of religious and other communal colonies;
    members of the Canadian Armed Forces living in Military Camps; people living in residences for senior citizens; people living full time in
    institutions: for example, inmates of penal institutions and chronic care patients living in hospitals and nursing homes. The survey covers
    about 98% of the population in the ten provinces. In the Territories, coverage was restricted to 88% in the Yukon, 92% in the Northwest
    Territories, and 89% in Nunavut. Information was not gathered from persons temporarily living away from their families (for example, students
    at university), because it would be gathered from their families if selected. In this way, double counting of such individuals was
    avoided.</universe>
<dataKind />
    </sumDscr>
<notes>Data refer to previous year, known as the Reference Year.</notes>
    </stdyInfo>
<method>
<dataColl>
<timeMeth>Cross-Sectional</timeMeth>
<dataCollector />
<frequenc>Annual 1997-1999 Biennial 2000 -</frequenc>
<sampProc>Stratified Sample</sampProc>
<collMode>Computer assisted telephone interviews (CATI)</collMode>
<source />
<collSitu />
<weight>WEIGHT=Weight at household level</weight>
    </dataColl>
  <notes />
  <anlyInfo />
  <respRate />
    </method>
<dataAccs>
<setAvail>
```

```xml
  <accsPlac />
<useStmt>
 <restrctn>The data may be used by current students, faculty and staff of the University of Guelph, the University of Waterloo and Wilfrid Laurier
     University for academic and personal research only.</restrctn>
 <contact />
 <citReq />
<conditions>
    Data Liberation Initiative
 <ExtLink URI="http://tdr.tug-libraries.on.ca/HELPS/dlilicence.html" title="DLI Licence Agreement" />
     </conditions>
     <disclaimer />
     </useStmt>
     </dataAccs>
   <othrStdyMat>
     <relMat />
     <relStdy> The Survey of Household Spending integrates most of the content found in the Family Expenditure Survey  (Cat. no.  62M0001XCB)
      and the Household Facilities and Equipment Survey (Cat. No.  64M0002XCB) . </relStdy>
   <relPubl />
   </stdyDscr>
```