Publishing Microdata to <odesi> Using Nesstar Publisher 4.X
(using DDI 2.x)

April 29, 2013

Alexandra Cooper, Queen’s University
Table of Contents

LIST OF REVISIONS ................................................................................................................................. 2
WHAT’S NEW IN NESSTAR PUBLISHER 4.0 .......................................................................................... 3
REFERENCE GUIDE FOR DIFFERENCES BETWEEN NESSTAR 3.5 AND NESSTAR 4 ...................................... 5
1) INTRODUCTION ................................................................................................................................. 6
2) IMPORTING FROM NESSTAR PUBLISHER 3.x .................................................................................. 7
3) OBTAINING THE DATASET .............................................................................................................. 8
4) PREPARING THE DATASET ............................................................................................................. 9
5) IMPORTING STUDY(s) OR DATASET(s) ........................................................................................... 12
6) MARKING UP THE DATASET - CREATING METADATA FROM SCRATCH ........................................ 14
   a) Document Description .................................................................................................................... 14
   b) Study Description ............................................................................................................................ 15
   c) Other Study Materials .................................................................................................................... 15
   d) Datasets ......................................................................................................................................... 15
   e) Variable Groups ............................................................................................................................ 20
   f) Other Materials .............................................................................................................................. 21
   g) External Resources ....................................................................................................................... 21
   h) Validate Metadata and Variables (Tools → Validate ...) ............................................................... 22
   i) Save .Nesstar file .......................................................................................................................... 22
   j) Exporting Metadata (.xml file) ..................................................................................................... 22

TIPS FOR MARKING UP FILES .............................................................................................................. 23
7) MARKING UP THE DATASET - USING METADATA FROM OTHER SOURCES .................................. 24
8) MANAGING SERVERS ...................................................................................................................... 28
9) PUBLISHING IN NESSTAR .............................................................................................................. 30
10) TEMPLATES IN NESSTAR .............................................................................................................. 32
11) RE-IMPORTING THE DATA FILE INTO NESSTAR ........................................................................ 37

APPENDIX A — TEMPLATE GUIDE - DESCRIPTION OF THE TAGS IN DDI 2.x .................................................. 39
APPENDIX B — WHEN SHOULD I ASSIGN WEIGHTS TO INDIVIDUAL VARIABLES? ............................. 45
APPENDIX C — CHANGING ALPHANUMERIC VARIABLES TO NUMERIC VARIABLES IN THE SPSS SYNTAX FILE .................................................................................................................. 46
APPENDIX D — ERROR WARNING WHEN IMPORTING DATA .................................................................. 47
LIST OF REVISIONS

July 10, 2012
- general editing (typos, spacing, re-arrange some pictures)
- added Appendix C – Changing Alphanumeric Variables to Numeric Variables
  (Revisions made by Queen’s)

July 24, 2012
- added 2) Importing from Nesstar 3.x
  (Revisions made by Queen’s)

July 30, 2012
- added 7)f) Tags to edited when XML is imported
  (Revisions made by Queen’s)

September 11, 2012
- added Appendix D – Error warning when importing data
  (Revision made by Queen’s)

April 29, 2013
- added note to 7 d) Edit Metadata as described in MARKING UP THE DATASET – CREATING METADATA FROM SCRATCH (Section 6)
  (Revision made by Carleton)
**WHAT'S NEW IN NESSTAR PUBLISHER 4.0**

New to Publisher is the concept of ‘Projects’. When survey data, cube (tabular) dataset or other resources is brought into Publisher it will be saved as a project. Projects can include a number of datasets, ‘Cube setups’ and resources.

The left hand side of Publisher will display the list of projects in Publisher. From here, you can open, close and edit files.

**Cube Builder** is no longer a separate program. It can be accessed by opening the Datasets folder in the file you want to use it with.

**Resource Publisher** is no longer a separate program. It can be accessed by opening the Dataset folder in the file you want to use it with.

Publishing at a later date:
- can set a later date and time for published file to appear in webview. The default is ‘immediately’.

**Email Notification:**
- when this option is turned on, user will have the option of being notified when changes have been made to chosen files. There is an email icon in webview which user can click.

**ID field:**
- doesn’t like underscore ( _ ) but file will still publish
- Nesstar replaces the underscore ( _ ) with an exclamation mark (!)
- when a file is downloaded, the file name (which is taken from the ID field) will have exclamation marks (!) instead of underscores ( _ ). The BPD now recommends that dashes (-) be used instead of underscores ( _ ).

**Datasets**
- encompasses tabs File Description, Variables, Variable Groups, Data Entry

**Import from dataset:**
- this option allows you to copy data from one dataset to another.
- see **MARKING UP THE DATASET – USING METADATA FROM OTHER SOURCES** (Section 7)

**Importing weights**
- when importing metadata using Import from DDI... or Import from Study..., whatever weights are declared in the source file are copied to the file you are importing to. This includes both weights declared for the file and for each individual variable.

**Copying variable groups:**
- see **MARKING UP THE DATASET – CREATING METADATA FROM SCRATCH** (Section 6)

**Select Column**
- this allows you to add the columns you want to the Variable view. You can also arrange the columns in the order you want.
- click → choose columns → OK
Is **Weight Variable** checkbox
- in the Variables section, you can add the **Is Weight Variable** column. This allows you to check each variable that is a weight variable in a file. See MARKING UP THE DATASET – CREATING MetadAtA FROM SCRATCH, Section 6)(iii)(2) Declare Weight Variable for entire file

Validating metadata and variables
- checks that all mandatory fields are filled in (if any are missing, file will not publish), variables are not missing categories or labels, or that variables do not have duplicate labels
REFERENCE GUIDE FOR DIFFERENCES BETWEEN NESSTAR 3.5 AND NESSTAR 4
This section is for those who have worked with Nesstar 3.5.

Add dataset to My Projects
- highlight ‘My Projects’
- File ➔ Import Study... OR File ➔ Import Multiple Studies...

Remove dataset from My Projects
- highlight file to be removed
- click

Create file with metadata only (for aggregate files)
- highlight ‘My Projects’
- File ➔ Add New Study ➔ add metadata

Open file in My Projects
- highlight study name
- double click on study name OR click

Close file in My Projects
- highlight study name
- File ➔ Close (or Close all) OR click

Add xml file
- Documentation ➔ Import ➔ Import from DDI

Resequence
- Open file in ‘My Projects’
- Open ‘Datasets’ folder ➔ click on ‘Variables’
- Variables ➔ Resequence

Change Template
- open Template Manager ➔ choose template ➔ Use

Close left hand navigation window
Particularly useful when working in the Variables window and adding question text
- click
INTRODUCTION

This guide is focused on marking up and publishing Statistics Canada microdata files for <odesi>. In some cases, marking up other files such as public opinion polls or surveys from non-Statistics Canada sources, will follow the same steps outlined in this guide; however, the data file will come from different sources. In addition, different DDI fields may be used for marking up public opinion polls or non-Statistics Canada files. Mark-up students should consult their supervisors for further information when marking up non-Statistics Canada files.

For aggregate data, such as B2020 or Excel files, see the External Resources Guide.

The Best Practices Document (BPD) provides additional information about the creation of metadata in Nesstar for <odesi>. If in doubt, please refer to the BPD for clarification and assistance (http://spotdocs.scholarsportal.info/display/odesi/best+practices+-+english).

NOTE: This guide assumes that you have SPSS installed on your computer as well as Nesstar Publisher. If you use a statistical package other than SPSS, such as SAS, Stata, or Excel, some of the procedures in Preparing the Dataset (Section 4) may differ. Nesstar Publisher can import data in the following formats:

- Nesstar (*.Nesstar), new format for version 4.0
- NSDstat (*.NSDstat)
- DDI Document (*.xml)
- SPSS System (*.sav)
- SPSS Portable (*.por)
- SPSS Syntax (*.sps)
- Stata 7 and Stata 8 (*.dta)
- Statistica (*.sta)
- NSDstat (*.nsf), the old NSDstat format used by NSDstat Pro
- dBase (*.dbf)
- DIF (*.dif)
- Delimited Text (*.txt, *.csv, *.sdv, *.cdv, *.prn)
- PC-Axis (*.px)
- Excel (*.xls)
- Hierarchy Definition File (*.NSDstatHDef)

File size limitations: The maximum size of file that can be imported is approximately 10 Gigabytes, with a limitation within a file to 260 million cases. However, using files of this size will affect response times.

Parts of the manual are taken from Publishing to <odesi> Using Nesstar (using DDI 2.x), written by Alexandra Cooper, Jeff Moon (Queen’s University), Jane Fry (Carleton University), Amber Leahey, Leanne Hindmarch (Scholars Portal).
2) **IMPORTING FROM NESSTAR PUBLISHER 3.x**

If you have a Nesstar 3.x file (.NSDStat) it is very easy to import it into Nesstar 4.0.

a) **Importing .NSDStat file**

   - Some files can take a while to import – be patient
   - (i) Open Nesstar Publisher 4.0
   - (ii) Click on My Projects
   - (iii) (if this is not highlighted, then the Import Study option will not be available)
   - (iv) Click on File → Import Study...
   - (v) Select the .NSDStat from the drop down box → Double click on file name
   - (vi) File will open in the right hand window
   - (vii) Save the file → File → Save
   - (viii) If necessary, make any edits and save new edits
   - (ix) Publish

**Note:**
Most .NSDStat files will likely have an ID# with underscores ( _ ). These need to be changed to dashes ( - ) or they will not be read properly by Nesstar 4.
3) **Obtaining the Dataset**

You need to have a microdata file to work with in order to import it into Nesstar Publisher, add metadata, and publish the result to <odesi>’s Nesstar Server (after which the file will appear in <odesi>’s Nesstar Webview at http://odesi.scholarsportal.info). In most cases, this step will be performed by the DLI contact for your institution, since they are the ones with the passwords to access the Statistics Canada files. Mark-up students should consult their supervisors to retrieve the files for them from the appropriate source.

**a) Marking up a new file, not currently in <odesi>:**

(i) Check if Statistics Canada has already marked up the file
- Statistics Canada Nesstar: http://www62.statcan.ca/webview/
- If needed file is identified, mark-up students ask supervisor to download
- Download the **data** in SPSS format (.sav)
- Download the **documentation** in XML format (.xml)
- Unzip both files
- Import the dataset into Nesstar → see **Importing Study(s) or Dataset(s)** (Section 5)
- Import documentation to the file and make appropriate updates to the metadata → see **Marking up the Dataset – Using Metadata from Other Sources** (Section 7)

(ii) If Statistics Canada has not already marked up the file, download the file from Statistics Canada’s FTP site
- Mark-up students ask supervisor to download
- **Data** will usually be available in the form of an ASCII text file (.txt or .dat) and a syntax file (.sps).
- **Documentation** in .xml format will not yet have been created.
- **Documentation** in the form of a codebook, user guide and/or a questionnaire should be available (.pdf format usually).

**b) Correcting a file that is already in <odesi>:**

(i) If you were the person who originally marked up this file, you will want to work from your saved .Nesstar file (produced from going through the process described in this guide when you first marked up the file).

(ii) If you did not originally mark-up this file, you will need to download the file from <odesi>’s Nesstar (http://odesi.scholarsportal.info).
- Download the **data** in SPSS format (.sav)
- Download the **documentation** in XML format (.xml)
- Unzip both files
- Import the dataset into Nesstar → see **Importing Study(s) or Dataset(s)** (Section 5)
- Import documentation to the file and make appropriate changes to the metadata → see **Marking up the Dataset – Using Metadata from Other Sources** (Section 7)
- Be sure to be using the appropriate template → see **Templates in Nesstar** (Section 10)

**Note:** Sometimes, the necessary correction will already have been made by Statistics Canada in the intervening time since the file was first published to <odesi>. You may wish to check if Statistics Canada has already marked up the corrected file. If yes, students should ask their supervisor to download the data and documentation.
4) **PREPARING THE DATASET**

Before marking up the file (entering the metadata into Nesstar Publisher), always check to ensure that the file is ready for public use. You will need to check for the following:

- there is an id variable;
- all the variable labels are correct and unique (avoid lengthy labels/repeats that read as a literal question);
- all the value labels are correct;
- all the missing values have been declared;
- all the recoded variables are correct; and
- make sure the data are not weighted.

This guide assumes you will perform these checks using SPSS. However, sometimes quick fixes can be made in Nesstar. Generally, you will have either an SPSS system file (.sav format – obtained from <odesi> or Statistics Canada Nesstar), or you will have a combination of an ASCII text file (.txt or .dat) and a syntax file (.sps – obtained from Statistics Canada FTP site). The following will describe how to prepare the dataset in each case.

a) **Working with an “.sps” syntax file (creating an SPSS system file .sav)**

Open the syntax file (.sps file) by double-clicking on it. The .sps file is the program file which converts the raw data file (ASCII text file) to an SPSS system file (.sav) which is then imported into Nesstar.

(i) **Modify the SPSS command code**

See below for instructions on editing the syntax file.

Tips on editing SPSS command code:

- all commands end with a period
- all commands start in the first column; subsequent lines in the same command are indented by 1-2 columns
- all lines of text are in quotation marks
- LRECL (Logical Record Length) can be determined by looking at the last position of the last variable on the raw data file. Or, the last number at the end of the DATA LIST FILE

(ii) **Run the file**

SPSS will run the syntax file’s code against the raw data file referred to

(1) Save the file
(2) From the Menu bar, use the menu command: Run ▶ All to run the code. When it is done, “SPSS processor is ready” will appear at bottom of screen.

(iii) **Check results.**

(1) Check that the SPSS System File (.sav) was saved to the disk location you specified at the bottom of your SPSS program (.sav file will have been added to working directory / folder)
(2) From within the data editor screen (the one that looks like a spreadsheet), use the menu command: Analyze ▶ Descriptive Statistics ▶ Frequencies
(3) Choose an appropriate variable (e.g. Province, or some other categorized variable)
(4) Run the analysis and compare your results (e.g. number of cases) with the number of cases indicated in the “Readme” file or Data Dictionary for the survey
(5) The .sav file can now be imported into Nesstar Publisher
**TITLE** "GENERAL SOCIAL SURVEY CYCLE 1, 1985"
**SUBTITLE** "HEALTH AND SOCIAL SUPPORT"
**SET LENGTH**=NONE **WIDTH**=80.
**FILE HANDLE** GSS/NAME=C:\data\gss0185.dta*LRECL=559.

### DATA LIST

- **FILE** = GSS/
- **SEQNO** 1 - 5
- **GENNLTH** 6 - 6
- **BPCHECK** 7 - 7

```
       ICIMACT1  22 - 22
       DAYSMISS  23 - 24  (A)
       CUTDOW   25 - 25
       CUTALLDDY 26 - 27
       ICIMACT2  28 - 28
       DAYSNTAB  29 - 30  (A)
       SEEDOC   31 - 31
```

- **WEIGHTI** 91 - 93
- **WEIGHTM** 94 - 97  (1)
- **OVRUNDRE** 98 - 98
- **EXERCISE** 99 - 99
- **TYPEXER1** 100 - 100

```
       GCHILD  556 - 556
       RELATIV  557 - 557
       EXFARE  558 - 559.
```

**VARIABLE LABELS**

- **SEQNO**  "SEQUENCE NUMBER"
- **GENNLTH**  "DESCRIBE YOUR STATE OF HEALTH"
- **BPCHECK**  "TIME SINCE LAST BLOOD PRESSURE CHECK"
- **GCHILD**  "HOUSEHOLD INCLUDES RESPOND GR-CHILDREN"
- **RELATIV**  "HOUSEHOLD INCLUDES RESPONDENTS RELATIVES"
- **EXFARE**  "EXTENDED FAMILY SUMMARY VARIABLE".

**Note:**
- Make most changes in this section.
- Clean up variables if needed:
  - check spelling
  - spell out words in full (i.e. cigs = cigarettes)

**‘nickname’ of file**

same as DATA LIST

**FILE name**

full path to where raw data is located

**Logical Record Length (LRECL)**

this number is the LRECL — it is the last position of the raw data file

Notes:

- "(A)" — indicates that this variable is alphanumeric; its value label will be letters not numbers. If the data are numbers, this has to be changed. See Appendix C for instructions on how to change alphanumeric variables to numeric variables.

- "(1)" — indicates the number of decimal places to be read in; in this case, this variable should be interpreted as having 1 decimal place.

- A period (" . ") signifies the end of this command.
b) **Working with an SPSS system file (.sav) file.**
If you are starting directly with a .sav file, you do not need to go through the process that was described in part a) above.

(i) Open the .sav file in SPSS.
(ii) Go through the checklist listed at the beginning of this section.
(iii) Make any needed changes.
(iv) Save the file.
(v) The .sav file can now be imported into Nesstar Publisher.
5) **Importing Study(s) or Dataset(s)**

a) **Importing One Study or Dataset**
   Some files can take a while to import – be patient
   (i) Open Nesstar Publisher
   (ii) Ensure that you’re using an appropriate template (see the **TEMPLATES IN NESSTAR** (Section 10) for information on how to create Templates if there is not already one for you to use).
   
   1. Click \( T \) \( \rightarrow \) Template Manager window opens.
   2. Highlight the appropriate template \( \rightarrow \) click on Use
   (iii) Click on **My Projects**
   (if this is not highlighted, then the **Import Study** option will not be available)
   (iv) Click on **File** \( \rightarrow \) Import Study...
   (v) Select the file type (.sav) from the drop down box \( \rightarrow \) Double click on file name

   **ERROR NOTE:** when importing some files, there may be a warning: *“The measure definitions for variables appear to be wrong, do you want them corrected?”* 
   Click on **Yes** and the file will be imported. For information on this warning, see **Appendix D**.

   When data is imported, you will see this screen. To add/edit metadata, click on the section names on the left. To see the data, click on the ‘+’ beside ‘Datasets’, then on ‘Variables’.

   ![Nesstar Publisher Interface](image)

   **NOTE:** If you have pre-existing metadata (e.g. you are editing a file that has already been marked up and you have downloaded the documentation in XML format from either Statistics Canada or <odesi>’s Nesstar), see **MARKING UP THE DATASET – USING METADATA FROM OTHER SOURCES** (Section 7).

b) **Importing Multiple Datasets as One Study**
   This allows you to import multiple datasets as one study. To import more than one dataset as multiple studies, use Importing Multiple Datasets or Studies.
   (i) Open Nesstar Publisher
   (ii) Ensure that you’re using an appropriate template (see the **TEMPLATES IN NESSTAR** (Section 10) for information on how to create Templates if there is not already one for you to use).
   
   1. Click \( T \) \( \rightarrow \) Template Manager window opens.
   2. Highlight the appropriate template \( \rightarrow \) click on Use
   (iii) Click on **My Projects**
   (if this is not highlighted, then the Import Study option will not be available)
(iv) Import dataset – some files can take a while to import – be patient

(1) Click on File → Import Study...
   Note: do not click on Import Multiple Studies... this imports multiple datasets as individual studies, not as one study.

(2) Select the file type (.sav) from the drop down box
(3) Select the files to import using Shift/Ctrl keys → Click on Open
When data is imported, you will see this screen. To add/edit metadata, click on the section names on the left. To see the data, click on the ‘+’ beside ‘Datasets’, then on ‘Variables’.

When multiple datasets are first imported, by default the project name will be the name of one of the files. This can be changed by using the ‘Save As’ function.

c) Importing Multiple Studies or Datasets
This allows you to import multiple datasets as individual studies. To import multiple datasets as one study, use Importing Multiple Datasets as One Study.

(i) Open Nesstar Publisher
   Ensure that you’re using an appropriate template (see the TEMPLATES IN NESSTAR (Section 10) for information on how to create Templates if there is not already one for you to use).

   (1) Click T → Template Manager window opens.
   (2) Highlight the appropriate template → click on Use

(ii) Click on My Projects
   (if this is not highlighted, then the Import Study option will not be available)
(iii) Import dataset – some files can take a while to import – be patient

   (1) Click on File → Import Multiple Studies...
   (2) Select the file type (.sav) from the drop down box
   (3) Select the files to import using Shift/Ctrl keys → Click on Open
6) **MARKING UP THE DATASET - CREATING METADATA FROM SCRATCH**

Below is the list of sections in which you will be entering information. Each section contains a number of fields that must be filled out. See **APPENDIX A** for a list of all the fields and the formats for filling the fields for Document Description and Study Description. For more detailed information and examples, please refer to the Best Practices Document (http://spotdocs.scholarsportal.info/display/odesi/best+practices+-+english). After entering the metadata, the last two steps in marking up the dataset are saving the .Nesstar file and exporting the .xml file.

**NOTE:** If you have existing metadata in XML format, you can import it rather than starting from scratch – in this case, skip to **MARKING UP THE DATASET - USING METADATA FROM OTHER SOURCES** (Section 7).

Click on Document Description or Series Statement to see the information in the sections. To edit, click on the ‘+’ and then the field.

**a)** Document Description

Describes the metadata, that is, the document which you are presently putting together.

Click on ‘+’ beside Document Description to open section. Click on field name to edit.

![Metadata Editing Interface]

[Image description: A screenshot of the Nesstar Publisher v4.0.2 software interface showing the editing of a Document Description section. Fields such as Title, Authors, and Notes are displayed with example text entries.]

Alexandra Cooper (Queen’s University)   September 11, 2012
b) Study Description
Describes the study
NOTE: for Beyond 2020, Excel files, and any other aggregate data only:
• 2.4.1 Media – describes the type of aggregate data. These will display as a direct download of the data/page from the <odesi> search interface. Use one of the following tags – B2020, Excel, data.
• 2.4.1.1 Location of Data Collection – points to the URL of the aggregate data page (.html) or the data files on the <odesi> ftp server; this is the only time this field would be used
• Enter the link to the aggregate data file on the <odesi> ftp server:
  http://odesi.scholarsportal.info/documentation/{surveyfoldername}/{filename.html}

Click on ‘+’ beside Series Statement to open section. Click on field name to edit.

c) Other Study Materials
Describes other related studies or publications for this data file. To enter a study or publication:
(i) Highlight the section the survey is to go under
(ii) Click ‘+’
(iii) Enter title of study and any other relevant information

d) Datasets
The Dataset section shows all the datasets imported into the file. Click on the ‘+’ beside the dataset you want to edit.

(i) File Description
Describes file content and version information for the data
(1) File Content
Brief information about what the data covers
(2) Data Check
Information on any checks performed on the data
(3) Version Statement
   (i) Version Text
   Version number (if there is no version number, insert the date you are working on it)
   (ii) Version Date
   Date of revision
   (iii) Version Responsibility
   Who is responsible for the revision (eg, Statistics Canada)
   (iv) Version Notes
   Reason for revision (eg, errors in processing; re-weighting)

(ii) Key Resources & Relations
This section is used to specify the relationships between survey datasets. These studies may contain several separate datasets that are related in some way. Some files may be hierarchically related, i.e. they contain data at different levels, example, Household level data and Individual level data, or they may just all belong to the same study with no relationship between them.

(iii) Variables
   (1) Resequence variables
   This is very important to do and must be done for each dataset. It recalculates the start and end positions for each of the variables. If it is not done, then the user will get an error when trying to export to spss, sas or ASCII.
   (i) Open Datasets → click on Variables
   (ii) Variables → Resequence
   (iii) Resequence dialog box – “Do you want to recalculate the widths? Recalculating the widths are recommended if you plan to export as Fixed Format Text, SAS or SPSS syntax.” → Yes

(2) Declare Weight Variable for entire file
State which variable is the weight variable (repeat for each weight variable). This applies the weight to the entire file; to add weight to individual variables, see Add weight variable to individual variables below.
(i) Choose weight variable (user guide will list weight variable(s))

(ii) In lower right-hand corner, under Variable Information, check Is Weight Variable (do this for each weight variable)

OR

Turn on Is Weight Variable checkbox

(iii) Click on Select Columns icon

(iv) Customize the variable list dialog box opens

(v) Click Is Weight Variable → OK

(3) Add weight variable to individual variables

(i) Click on Weights tab in the middle of the screen

(ii) Highlight variables that weight is to be added to

(iii) Click on ‘+’ beside Documentation box; the Select Variable dialog box opens
(iv) Choose weight variable and click **OK**
- if there is more than one weight variable, the proper weight variable must be applied to the appropriate variable. The User Guide will usually have this information.
- if there is more than one weight variable, and each one can be applied to any of the variables (for example, an individual weight, a household weight), then **do not** do this step.
- for more information on when to assign weights to individual variables, see **Appendix B**

(v) Weight variable is now listed under the **Weights** tab

(vi) Click on **Statistics** tab, choose variable that weight has been applied to. The **Preview** window will now show the variable with both weight and non-weighted values.
(4) **Check variables** (most of this should have been done in SPSS):
   (i) Variable and value labels correct
   (ii) Missing values declared (if not, dataset will have to be corrected in SPSS and re-imported)

(5) **Question Text information**
   (i) On left hand side of screen, click on **Variables** under **Datasets**
   (ii) Go to Documentation tab in the middle of the screen
   (iii) Fill in the Pre-Question Text, Literal Question, Post-Question Text, and Interviewer Instructions, fields. Cut and paste from Questionnaire or User Guide, when possible.
   (iv) Fill in the Universe and Notes fields - this section applies to each individual variable. Cut and paste from codebook or user guide, when possible.

(iv) **Data Entry**
    This section shows the data. Do not change anything here.

(v) **Cell Notes**

(vi) **Cubes Setup**
    See guide for Cubes.
e) **Variable Groups**  
Variable groups cluster variables with similar characteristics.  
Some basic groups:  
- *Administration* – variables that concern the administering of the survey and not the actual respondent or any of their responses, e.g. Caseid  
- *Weight* – weight variable(s)  
- *Demographics* – age, sex, education, etc.

**NOTE:** Variables can be in more than one group; i.e. “Languages other than Fr/Eng taught” would be in the Languages Group and the Multiculturalism Group.

(i) **Create Variable Groups**  
1. Click on Description tab  
2. On left-hand side, click on Variable Groups, then on ‘+’  
3. Type group name in Label: field (right side of screen) and any additional information about the variable group  
4. Next variable → click on Variable Groups, then on ‘+’  
   - if you click on ‘+’ before Variable Groups, you will create a sub-folder of the current group; to remove folder, click on ‘−’

(ii) **Add Variables to Groups**  
1. Click on Variables tab  
2. Highlight the Variable Group that you want to add the variables to  
3. Click on ‘+’ on far right side of screen  
4. Select Variables window opens  
5. Choose variables to add, click OK  
   - a green dot appears beside a variable when it has been added to a group  
6. To remove variable from group → highlight variable → click on ‘−’  
7. To move variables up/down → highlight variable → click on ↓ or ↑  
8. When done, check the list to make sure all the variables in the dataset have been put in a group. If they are, they will all have a green dot beside them.

(iii) **Copying Variable Groups**  
Variable groups can be copied within a file. There are two options: *Copy* and *Copy Shared*.  
**Copy:**  
1. Highlight the group(s) to copy  
2. Right click and choose Copy  
3. Where copied groups are to be placed, right click and choose Paste

**Copy Shared:**  
With Copy Shared, all actions of the original group are replicated in the copied group (renaming the group, adding, and deleting variables).  
Copy Shared groups are denoted by a blue variable group icon  
1. Highlight the group(s) to copy
(2) Right click and choose **Copy**
(3) Where copied groups are to be placed, right click and choose **Paste**

**f) Other Materials**
Provides links to all documentation that accompanies the data file for this study: codebooks, user guides, questionnaires, SPSS and SAS program files, etc.

There are two steps to make documents available in Webview:
- FTP the documents to the `<odesi>` server
- add the link in Publisher

**i) FTP Documents to the `<odesi>` Server**
(1) open FTP client to the `<odesi>` docs folder
(2) go to the survey folder (or add survey folder)
(3) upload the documents in folder
(4) close FTP client

**ii) Add the link in Publisher**
(1) Click on ‘+’
(2) Enter title of study and location on the `<odesi>` ftp server
  - Location of ftp site:
    http://odesi.scholarsportal.info/documentation/{surveyID}/{filename}
    `{surveyID}` is the file name of the dataset (minus the file extension)
    **Note:** make sure there are no special characters in the file name

**g) External Resources**
Replaces the ‘Resource’ Publisher that was part of Publisher 3.5.
See External Resources Guide.
h) Validate Metadata and Variables (Tools → Validate …)

(i) Validate metadata
This option checks that all mandatory metadata fields are completed. If any empty fields are found they will be marked in red. Fields are marked mandatory in the Template Editor.

(ii) Validate External Resources
This options checks that all mandatory fields for ‘External Resources’ are completed. If any empty fields are found they will be marked in red. Fields are marked mandatory in the Template Editor.

(iii) Validate Variables
This option enables you to validate the variables within a dataset and to locate those that may have missing information, e.g. missing category labels. The following options are available:

- **Nominal variables without categories**
  Checks for missing category labels within nominal variables.

- **Variables with missing label**
  Checks for any variables with missing labels.

- **Variables with same label**
  Use to quickly locate variables that have the same label.

- **Variables with missing mandatory fields**
  Checks for missing mandatory information.

- **Scale variables with non-missing categories**
  Locates scale variables that contain categories that are not defined as ‘missing’. For example, some ‘Age’ variables record a respondent’s exact age, but may include a category ‘99’ labelled ‘Not answered’.

To use:

(i) Select variables to be validated by highlighting variable
(ii) Tools → Validate Variables
(iii) Select the type of validation from the drop down list
(iv) Click OK

A message will be displayed saying how many variables were found that matched the validation criteria, and these will be highlighted in the Variables window.

i) Save .Nesstar file

(i) File → Save As...
(ii) Name file same as .sav and save in appropriate directory → Save

j) Exporting Metadata (.xml file)

Export the file to create the .xml file. Make sure that you have saved the file before exporting the metadata

(i) Click on Documentation → Export → Export DDI...
(ii) Save DDI file as .xml in proper directory
TIPS FOR MARKING UP FILES

- For *Document Description* and *Study Description* the correct title is necessary, including the year and study # (if applicable).
- When working on the left-hand side of the page, you can highlight a field to see more detail, which is then shown on the right-hand side of the screen.
- When you are filling in the information on the right-hand side and you want to add more entries, click on the ‘+’ sign on the right-hand side. This can be done as often as necessary. It will ensure that the entry is on a different line when it is published. If there is no ‘+’ sign, then it is not possible to add more entries in that particular field.
- Likewise, you can click on the ‘−’ sign to take away an entry you don’t want.
- If you want to re-order your entries, put the cursor over the entry you want moved and click once so there is a box around it, and then click on the up or down arrow on the right-hand side.
- If there is a Globe icon on the right-hand side, clicking on it will enable you to put in a URI (external link) that is relevant to that particular field.
- Correct spelling and grammar are of the utmost importance.
- It may seem like you are repeating yourself when you are entering metadata in different sections and that is because you are! Remember that the different fields you are entering information into will appear in different parts of the finished document and must be complete on their own.

🌟 Always remember to SAVE the file and export the .xml file regularly!
7) **MARKING UP THE DATASET - USING METADATA FROM OTHER SOURCES**

As described in **OBTAINING THE DATASET** (Section 3), sometimes you are working with a file for which you already have metadata in XML format. For example, you’re correcting a file that was already published in <odesi>, or you’re working on a file that has already been published on the Statistics Canada Nesstar site. In this case, you will import the metadata into Nesstar, edit it as necessary, then save it and export the new .xml file.

a) **Import your dataset** as described in **IMPORTING STUDY(S) OR DATASET(S)** (Section 5)

b) **Import your metadata**

   (i) **Documentation → Import → Import from DDI... (.xml file)** OR **Documentation → Import → Import from Study... (.Nesstar file)**
   These two options are essentially the same. One uses the .xml file the other, the .Nesstar file

   (ii) **Import Documentation** dialog box
       (1) Check the sections you want to import
       (2) Click OK
       (3) Find the .xml file
       (4) Click on **Open** and the metadata will be added to your file

In Publisher 4.0, there are some new import options. Below is a list of options that are different from Publisher 3.5.

**Variable Information** – imports the following:
- variable label
- location information (start column, end column, width, record number)
- decimals for numeric variables
- measure definition for numeric variables
- missing definitions for numeric and fixed string variables
- min and max for numeric and data variables are reset
- statistics options are reset for numeric variables
- universe and notes for variables

**Variable Documentation** – imports all the documentation for a variable, except for question text and variable labels

**Question Text** – imports the literal, pre- and post-questions and interviewers instructions; universe and notes are imported with Variable Information

**Weights** – if weights are declared in the source file, then they will be copied when imported (for both the file and individual variables)

**NOTE:** when importing, the file imported will override any previous default fields from the template.
c) **Resequence Variables**

This is very important to do and must be done for each dataset. It recalculates the start and end positions for each of the variables. If it is not done, then the user will get an error when trying to export to spss, sas or ASCII.

(i) **Open Datasets → click on Variables**

(ii) **Variables → Resequence**

(iii) **Resequence** dialog box – “Do you want to recalculate the widths? Recalculating the widths are recommended if you plan to export as Fixed Format Text, SAS or SPSS syntax.” → **Yes**

The StartCol and EndCol will have an asterisk until the columns are resequenced.

d) **Edit Metadata** as described in **MARKING UP THE DATASET – CREATING METADATA FROM SCRATCH** (Section 6)

(i) **Document Description**

The following fields **must** be filled in/changed when editing metadata that has already been done or that you are using from another source.

The metadata you are creating will be a subsequent version of the metadata you have imported or you are already using. **Do not remove any of the previous information!**

If you are editing a file which was previously done by your institution and you already have the XML file, then you do not have to import anything. You will be working with the XML file which you already have.

Check the **Best Practices Document** for examples of what goes in each field (http://spotdocs.scholarsportal.info/display/odesi/best+practices++english).

- **Authoring Entity** (1.1.2.1)
  Add your name to the top of the list of people who have worked on this file.

- **Producer** (1.1.3.1)
  Add your institution’s name.

- **Copyright** (1.1.3.2)
  Add current date.

- **Production Date** (1.1.3.3)
  Add current date and name of institution.

- **Production Place** (1.1.3.4)
  Add city your institution is in.

- **Funding** (1.1.3.6)
  If the file is from outside ODESI, add funding information (e.g. Ontario Council of University Libraries (OCUL)).
- **Distributors** (1.1.4.1)
  Add your institution’s name.
- **Contact Persons** (1.1.4.2)
  Add your institution’s email.
- **Version Text** (1.1.6.1)
  List the date.
- **Version Responsibility** (1.1.6.2)
  Add name of your institution.
- **Version Notes** (1.1.6.3)
  Indicate the date of the version and what you have edited e.g. “Weighting was added to variables.”, “Variable groups were created”, etc.
- **Bibliographic Citation** (1.1.7)
  Replace with your institution and date of publishing. This is the only instance where you will replace information.
- **Holdings** (1.1.8)
  Optional

(ii) **Series Statement**
- **Restrictions** (2.4.2.3)
  Change restriction statement to ODESI’s restriction:
  “The data is restricted to use by current students, faculty and staff of <a href="http://www.ocul.on.ca/" target="new">Ontario Council of University Libraries (OCUL)</a> Member Libraries for academic research and teaching only.”

(iii) **Other Materials**
- Change file path for guides to point to the ODESI server

e) **Import from Dataset**
This option allows you to import dataset information (file description, variable information and documentation, categories, question text, keys and weight) from other files. **Import from Dataset** can only be used in the Dataset section. This option is especially useful when working with a study that has multiple datasets. For example, once question text has been entered for one dataset, it can be easily copied to another.

(i) **Highlight dataset**

(ii) **Documentation → Import → Import from Dataset…**

(iii) **Choose file to import from**

(iv) **Choose dataset to import from and sections to import → OK**
If the file you choose to import from has multiple datasets attached to it, you will have the option of which dataset to use. The ‘Match %’ shows the percentage of variables that match the variables of the file you are importing to.
f) Tags that must be edited when XML file is imported
The following tags must be checked and edited when an XML file from another source is used. Depending on the survey and the source of the survey, there may be other tags.
Section 1:
1.1.2.1 Authoring Entity
1.1.3.1 Producer
- your name and institution should be listed at the top of the list for Producer and Authoring Entity
1.1.3.2 Copyright
1.1.3.3 Production Date
- date you are working on the file for both Copyright and Production Date
1.1.3.4 Place of Production
1.1.3.6 Funding
- make sure it has Ontario Council of University Libraries (OCUL) listed
1.1.4.1 Distributors
1.1.4.2 Contact Persons
- your institution should be listed at the top of the list for Distributors and Contact Persons
1.1.6.1 Version
1.1.6.2 Version Responsibility Statement
1.1.6.3 Notes and Comments
1.1.7 Bibliographic Citation
Section 2:
2.1.6 Version Statement
- this section is only changed if changes are made to the data itself
- if only the metadata is being updated, then this is not used

g) Save .Nesstar file
(i) File → Save As...
(ii) Name the file the same as .sav file and then save it in the appropriate directory → Save

h) Exporting Metadata (.xml) - Export the metadata to create the .xml file.
(i) File → Save Dataset
(ii) Click on Documentation → Export → Export DDI
(iii) Save DDI file as a .xml file in proper directory
(iv) Click on Close
8) **Managing Servers**

This section shows you how to: add and remove servers; add folders; move folders and files; and delete folders and files.

a) **Adding a Nesstar Server**

   **NOTE:** This only has to be done the first time you publish to a particular server
   (i) **Publishing → Add server...** dialog box opens
   (ii) Enter URL of the server with the port number
   (iii) Enter user name and password
   (iv) Click OK

b) **Removing a Nesstar Server**

   (i) **Publishing → Server Name → Remove Server**

c) **Managing Folders**

   (i) **Publishing → Server Name → Manage server...**
   (ii) **Add folder**
      (1) Highlight the folder where the new folder is going into
      (2) Click on '+'
      (3) Type name of new folder and hit Enter
      (4) Click on folder icon to open **Catalog Properties**
          Here you can enter information about the folder
          • **Comment:** field – this field appears in WebView when you click on the folder
      (5) Click on **Close**

   (iii) **Delete folder**
      (1) Highlight folder to be deleted
      (2) Click on '−'
      (3) Click on **Close**
d) Managing Files within Folders

NOTE: It moves slowly!

(i) Publishing → Server Name → Manage server...

(ii) Move files between folders

- Highlight file to move → drag and drop to proper folder

(iii) Reorder files

- Highlight file to move → Click on up or down arrow

(iv) Deleting Files

Only delete files within Nesstar Publisher; when a file is published, there are a number of files created but hidden from view. If you don’t delete the file in Nesstar Publisher, the associated files are not deleted.

1) Publishing → Server Name → Manage server
2) Highlight file to be deleted
3) Click on ‘−’
4) Click on Yes if you want to delete file
5) Click on Close
9) **Publishing in Nesstar**

a) **Publish Data and Metadata**

(i) Open file in Publisher

(ii) **Publishing → Server Name → Study → Publish data and metadata...**

(iii) **Select catalogs to publish to** dialog box → find folder that study belongs in → put checkmark in box beside folder name

   (1) If a folder does not exist for a file, create a new one using [Managing Files within Folders](MANAGING SERVERS, Section 8)

   (2) Check to make sure that no other folder is selected; Publisher will sometimes ‘remember’ previously selected folders

(iv) If the file is not to be made public immediately:

   (1) Under **Time of availability:** → check **At the given time and date:**

   (2) enter time and date

(v) Click on **Publish**

![Select catalogs to publish to dialog box](image)

(vi) Window opens – **Working, Please Wait...** (some files can take a while to publish)

(vii) When file is published, **Publishing Report** will open

(viii) Click on **Open in Web client...** to view and perform quality checks on final published product
b) **Publish Metadata alone**
   (i) Publishing → Server Name → Study → Publish metadata only...
   (ii) *Select catalogs to publish to* dialog box → find folder that study belongs in → put checkmark in box beside folder name
       (1) If a folder does not exist for a file, create a new one using *Managing Files within Folders* (MANAGING SERVERS, Section 8).
       (2) Check to make sure that no other folder is selected; Publisher will sometimes ‘remember’ previously selected folders
   (iii) Click on **Publish**
   (iv) Window opens – *Working, Please Wait...* (some files can take a while to publish)
   (v) When file is published, **Publishing Report** will open → click on **Close**

c) **Republish**
   Only use this for problems that can be fixed in Publisher; if problem is fixed in SPSS, you must re-import the file. See **RE-IMPORTING THE DATA FILE INTO NESSTAR** (Section 11)
   (i) Publishing → Server Name → Study → Republish
     • study will publish to folder(s) originally published to
   OR
     Publishing → **Republish on all servers**
     • this will publish the file to any of the servers it has already been published to
   (ii) Window opens - *Working please wait...* (some files can take a while to publish)
   (iii) When file is published, **Publishing Report** will open → click on **Close**
10) **Templates in NESSTAR**

_Why use a Template:_ If you are importing a series of surveys into Nesstar, creating and using a template means that you don't have to type in the same information again and again and again and again and again. It means that the default fields will already have information in them when you use the correct template for your survey.

If you make a mistake when creating a Template, you can edit it (see Section C on **Editing Templates**)

a) **Creating templates** (see Appendix A for a list of the tags)

(i) Click on **Template Manager** window opens.

(ii) Click on New on right-hand side

(iii) Enter information about the template in the **Description** tag

(iv) Click on **Content** tag

(v) Left-hand window shows a series of folders

(1) Create groups to organize DDI fields:

(i) Highlight one tab (i.e. Document Description) → click on ‘+’

(ii) Create group under Document Description – type name for group (i.e. Title Statement – use something that describes DDI fields in group)

(2) Add items to group:

(i) Highlight new group

(ii) Under **Available Items:** open folder → list of fields will open

(iii) Highlight field → click on « (item will be added to template)

(3) When you have all the items for Document Description in the template, go to the next tab and repeat steps for Study Description, File Description and Variable Description.

(4) **Other Study Description Materials / Other Materials / Cell Notes / Cube Setups**
In these four sections, make sure that to check the boxes “Visible” boxes. If these boxes are not checked, then these sections will not be visible in the template and you will not be able to enter any information in the fields. You cannot add/remove any of the fields from these sections.

(vi) Add defaults to items
(1) Click on item in left-hand of screen (i.e. 1.1.1.1 Title)
(2) Item Description displays information about selected item:
   (i) Custom Label – DDI field name; can be edited
   (ii) Mandatory – denote if field is mandatory
   (iii) Fixed – users can view but not change field if this box is checked
   (iv) Description – information about DDI field; can be edited; appears in Publisher
   (v) Defaults – enter default text/values for field
   (vi) Controlled Vocabulary – creates list of terms for a field
      • enter text → click on ‘+’ → enter next text
      • use arrows to reorganize list
      • click on ‘−’ to remove item from list

(vii) When you have all the items you need for the template, click on OK, then save template

b) Saving templates
   NOTE: In order for the template to be properly saved, you must export it.
   In the Template Manager dialog box:
   (i) Click on Export
   (ii) Choose directory to save template in
   (iii) Click on Save *when saving your template, make sure to give it a meaningful name so you will be able to find it again

c) Editing templates
   In the Template Manager dialog box:
   (i) Click on Edit
   (ii) Make changes
   (iii) Click on OK
   (iv) Save template – Export → choose directory → Save
   NOTE: If you don’t do this step, changes will not be saved
d) Importing templates
   In the Template Manager dialog box:
   (i) Click on Import
   (ii) Choose template to import
   (iii) Click on Open
   (iv) Click on Use

e) External Links
   External links can be added to some DDI fields to provide links to pages
   In the Template Manager dialog box:
   (i) Highlight item on left side of screen
   (ii) Click on Defaults on right side of screen
   (iii) Click on globe at side of screen
   (iv) Enter URI and Title
   (v) Click OK

f) Using an existing template to create a new template
   ☆ Before creating the new template, make sure the original template is saved.
   As you go through the new template, check all fields and defaults, including external links, since these will all be copied from the original template.
   (i) Click ➔ Template Manager window opens.
   (ii) Highlight the template you want to use from the list of templates ➔ click on Edit
   (iii) Click on the Description tab ➔ give template a new name.
       Type in any new information you want saved in the appropriate fields ➔ click OK
       NOTE: When you are going through the new template, check every field as all the defaults that are in the template you copied from, will be in the new template.
   (iv) You will be brought back to the Template Manager ➔ click on Export
   (v) Choose directory to save template in ➔ click on Save
      The new template has now been saved
   (vi) You will be brought back to the Template Manager.
      To bring the template you copied back into the list ➔ click on Import.
      Find the template to import ➔ click on Open

g) Adding a default to a template
   Why: If you are working on a number of surveys from the same organization, or on a series of surveys, you can add defaults for the tags that are the same across the different surveys. This will save you a great deal of time.
   In the Template Manager dialog box:
   (i) Highlight the template you want to use from the list of templates
   (ii) Click on Edit in the right-hand box and a Template Editor window comes up. Click on Content tab to get list of entries to be edited
   (iii) Click on the field for which you would like to add the default.
   (iv) Click on the Default button in the right half of the window and add the text you would like to appear.
   (v) When you are finished, click OK on the right-hand side.
   (vi) Click on Export to save your changes. If you neglect this step, your changes will NOT be saved.

   Note: if a default is added to a template and the template is then imported into a file that has already been saved, the template will not override fields previously filled in.
h) **External Links**

External links can be added to some DDI fields to provide links to pages.

In the **Template Manager** dialog box:

(i) Highlight item on left side of screen

(ii) Click on **Defaults** on right side of screen

(iii) Click on globe 🌍 at side of screen

(iv) Enter URI and **Title**

(v) Click **OK**

---

**Image:**

![Template Manager](image.png)

---

i) **Adding extra lines to a template**

In the **Template Manager** dialog box:

(i) Highlight **User Template** on the left-hand side and click on **Edit** on the right-hand side and a template editor comes up.

(ii) Find the folder you want to add lines to on the left-hand side and click once on it. On the right-hand side you will see the different lines that are available for you to add.

(iii) Find the line you want to add, double-click on it, and it will be moved over to the left-hand side.

---

j) **Adding a Note Field**

This has to be done in the template.

In the **Template Manager** dialog box:

(i) Highlight the template you are currently using

(ii) Click on **Edit** in the right-hand box and a **Template Editor** window comes up

(iii) Add the note field to this template

(iv) Click on **OK**, **Template Editor** closes

(v) Click on **Use**, **Template Manager** closes. Note field will be added to Publisher.
k) **Getting rid of an invisible line of writing in a template**

This is the line of writing that you can't see in the template, (eg. the DLI License Link) but when the document is published, the line is there.

In the **Template Manager** dialog box:

(i) Highlight the template you are currently using
(ii) Click on **Edit** in the right-hand box and a **Template Editor** window comes up
(iii) Go to the area that you want to change
(iv) Click on **Default** on the right-hand side.
(v) Delete/change the info in this box. If the box is blank, click on the **globe** on the right-hand side of the screen and delete/change the URI and any other information that needs to be changed.
(vi) Click on **OK**, **Template Editor** closes
(vii) Click on **Use**, **Template Manager** closes.
11) **RE-IMPORTING THE DATA FILE INTO NESSTAR**

If there is a problem with the data file that cannot be fixed in Publisher, you will have to fix it in SPSS and re-import the datafile. Before you do this, ensure that you have exported the metadata as described above in **MARKING UP THE DATASET – CREATING METADATA FROM SCRATCH** (Section 6). Once you have finished fixing the data file, you must re-import the datafile into Publisher and then also import the metadata as follows.

**a) Import dataset** – some files can take a while to import – be patient

(i) Open Nesstar Publisher

(ii) Ensure that you’re using an appropriate template (see the **TEMPLATES IN NESSTAR** (Section 10) for information on how to create Templates if there is not already one for you to use).

1. Click $\Rightarrow$ Template Manager window opens.
2. Highlight the appropriate template $\Rightarrow$ click on Use

(iii) Click on My Projects

(if this is not highlighted, then the **Import Study** option will not be available)

(iv) Click on **File $\Rightarrow$ Import Study**...

(v) Select the file type (.sav) from the drop down box $\Rightarrow$ Double click on file name

**b) Resequence variables**

This is very important to do and must be done for each dataset. It recalculates the start and end positions for each of the variables. If it is not done, then the user will get an error when trying to export to spss, sas or ASCII.

(i) Open **Datasets $\Rightarrow$ click on Variables**

(ii) **Variables $\Rightarrow$ Resequence**

(iii) **Resequence** dialog box – “Do you want to recalculate the widths? Recalculating the widths are recommended if you plan to export as Fixed Format Text, SAS or SPSS syntax.” $\Rightarrow$ Yes

[c] 

**c) Declare Weight Variable for entire file**

States which variable is the weight variable (repeat for each weight variable). This applies the weight to the entire file; when re-importing into Nesstar, you do not need to add the weight to each individual variable. This information is in the .xml document that will be imported in the next step.

See section 6 **MARKING UP THE DATASET – CREATING METADATA FROM SCRATCH** subsection (iii) (2)

**d) Import metadata file**

See Section 7 **MARKING UP THE DATASET – METADATA FROM OTHER SOURCES** subsection b)
e) Tags that must be edited when XML file is imported
   For a list of tags that must be edited when using metadata from another source, see section 7, subsection f: **Tags that must be edited when XML file is imported**.

f) Save .Nesstar file
   (i) File → Save As...
   (ii) Name file same as .sav and save in appropriate directory → Save

g) Exporting Metadata (.xml file)
   Export the file to create the new .xml file. Make sure that you have saved the file before exporting the metadata
   (i) Click on Documentation → Export → Export DDI...
   (ii) Save DDI file as .xml in proper directory
APPENDIX A – TEMPLATE GUIDE - DESCRIPTION OF THE TAGS IN DDI 2.x

Note: This is a list of the possible tags that may be used in your template.

Note: for more information see the Best Practices Document.

Section 1: Document Description

Citation – Title Statement

1.1.1.1 Title – survey title, year [country, if not in survey title]: subtitle
   • same as 2.1.1.1
   • General Social Survey, 2005 [Canada]: Cycle 19, Time Use Survey, Main File
   • Canadian Tobacco Use Monitoring Survey, 2006: Cycle 1, Household File

1.1.1.2 Subtitle
   • same as 2.1.1.2
   • include everything after the colon in the title

1.1.1.3 Alternative Title – Acronym Year: Subtitle
   • same as 2.1.1.3
   • CTUMS 2005: Cycle 1, Household File
   • GSS19 2005: Cycle 19, Time Use Survey, Main File

1.1.1.4 Parallel Title (French title)
   • same as 2.1.1.4

1.1.1.5 ID Number – acronym-CatalogueNumber-language-year-subset
   • same as 2.1.1.5
   • ctums-82M0020-E-2004-c1-household-file
   • CatalogueNumber – use microdata file number (do not include the XCB or similar letters in catalogue number)

Citation – Responsibility Statement

1.1.2.1 Authoring Entity
   • name – Last name, First name
   • affiliation – Name of Institution. Name of Data Centre

Citation – Production Statement

1.1.3.1 Producer
1.1.3.2 Copyright
1.1.3.3 Production Date
1.1.3.4 Place of Production
   • Institution, City, Province, Country

Citation – Production Statement

1.1.3.6 Funding
   • used if there is external funding outside of normal operations

Citation – Distribution Statement

1.1.4.1 Distributors
1.1.4.2 Contact Persons

Citation – Series Statement

1.1.5.1 Series Name
1.1.5.2 Series Information

Citation – Version Statement

1.1.6.1 Version Text
   • date of version
   • most recent version first followed by older versions
1.1.6.1 Version Date
1.1.6.2 Version Responsibility
Section 2: Study Description

Citation

2.1.1 Title – survey title, year [country, if not in survey title]: subtitle
- same as 1.1.1.1
- General Social Survey, 2005 [Canada]: Cycle 19, Time Use Survey, Main File
- Canadian Tobacco Use Monitoring Survey, 2006: Cycle 1, Household File

2.1.2 Subtitle
- same as 1.1.1.2
- include everything after the colon in the title

2.1.3 Alternative Title – Acronym Year: Subtitle
- same as 1.1.1.3
- CTUMS 2005: Cycle 1, Household File
- GSS19 2005: Cycle 19, Time Use Survey, Main File

2.1.4 Parallel Title (French title)
- same as 1.1.1.4

2.1.5 ID Number – acronym-CatalogueNumber-language-year-subset
- same as 1.1.1.5
- ctums-82M0020-E-2004-c1-household-file
- CatalogueNumber – use microdata file number (do not include the XCB or similar letters in catalogue number)

Citation – Responsibility Statement

2.1.2.1 Authoring Entity
- name of institution as it appears in User Guide, usually on title page
- Statistics Canada. Income Statistics Division

2.1.2.2 Other Identifications
- other divisions mentioned as supporting or contributing to survey

Citation – Production Statement

2.1.3 Producer
2.1.3.2 Copyright
- year is release data of the data

2.1.3.3 Date of Production
- release date of the data
- if you are unsure of the date, use one of the following:
  - [1971 or 1972] - one year or the other
  - [1969?] - probable date
- [between 1906 and 1912] - used only for dates less than 20 years apart
- [ca. 1960] - approximate date
- [197-] - decade certain
- [197-?] - probable decade
- [18-] - century certain
- [18-?] - probable century

2.1.3.4 Place of Production

- *Institution, City, Province, Country*
- *Statistics Canada, Ottawa, Ontario, Canada*
- *Statistics Canada, Ottawa, Ontario, Canada: Special Surveys Division*

Citation – Production Statement

2.1.3.6 Funding

- used if there is external funding outside of the normal operating expenses

Distributor Statement

2.1.4.1 Distributors
2.1.4.2 Contact Person
2.1.4.3 Depositor

- the name of the institution (or person) who provided this data collection to the archive storing it

2.1.4.4 Date of Deposit

Distribution Statement

2.1.4.5 Date of Distribution

Series Statement

2.1.5.1 Series Name
2.1.5.2 Series Information

Citation – Version

2.1.6.1 Version
2.1.6.1 Version Date
2.1.6.2 Version Responsibility
2.1.6.3 Version Notes

Bibliographic Citation

2.1.7 Bibliographic Citation

Note: This will be different than the Bibliographic Citation under Document Description

Scope – Subject Information

2.2.1.1 Keywords

- list alphabetical
- capitalize first word

2.2.1.2 Topic Classifications

- subjects

Abstract

2.2.2 Abstract

- take from user guide or Statistics Canada website

Scope – Summary Data Description

2.2.3.1 Time Periods

- time period to which data refers
- *year-month-day*

2.2.3.2 Dates of Collection

- date when the data was collected
- *year-month-day*
Scope – Summary Data Description
2.2.3.3 Country
• Canada

2.2.3.4 Geographic Coverage
• each level of geography covered by the data
2.2.3.5 Geographic Unit
• lowest level of geography covered by the data

2.2.3.6 Unit of Analysis
• Individual, Household, Family, Episode

2.2.3.7 Universe
• population covered by file
• Included: all Canadian citizens 15 years of age and older
  Excluded: residents of Indian Reserves; and residents of institutions

2.2.3.10 Kind of Data
• census data
• aggregate data
• survey data

Notes – Study Scope
2.2.4 Notes

Methodology – Data Collection
2.3.1.1 Time Method
• cross-sectional
• longitudinal
• time series

2.3.1.2 Data Collectors
• only use if you know specific division

2.3.1.3 Frequency of Data Collection
• Quinquennial – every 5 years
• Biennial – every 2 years
• Annual – every year
• Biannual – twice a year
• Monthly – every month
• Weekly – every week
• Occasional – conducted more than once, but not on a regular basis
• One-time – only conducted once

2.3.1.4 Sampling Procedure
• type of sample and sample design used to select respondents
• Stratified random sample
• Random digit dialing

Methodology – Data Collection
2.3.1.6 Mode of Data Collection
• method used to collect data
• computer assisted, random digit dialing, telephone interview

2.3.1.8 Data Sources
2.3.1.9 Characteristics of the Data Collection Situation
• noteworthy aspects of the data collection; i.e. number of respondents, length of interview

2.3.1.12 Weighting

Notes – Methodology and Processing
2.3.2 Notes

Data Appraisal

2.3.3.1 Response Rate
- percentage of respondents who provided information

2.3.3.2 Estimate of Sampling Error
- measure of how precisely one can estimate a population value from a given value

Data Access – Data Set Availability

2.4.1 Media

2.4.1.1 Location of Data Collection
- this field is used only for Beyond 2020, Excel files, and any other aggregate data
- Enter the link to the aggregate data file on the <odesi> ftp server
  (http://odesi.scholarsportal.info/documentation/{surveyID}/{filename})
  {surveyID} is the file name of the dataset (minus the file extension)

2.4.1.2 Original Archive where Collection Stored

2.4.1.3 Availability Status

2.4.1.4 Extent of Collection

2.4.1.5 Completeness of Collection Stored

Data Access – Data Set Availability

2.4.1.6 Number of Files

2.4.1.7 Notes

Data Use Statement

2.4.2.3 Restrictions

2.4.2.4 Access Authority (Contact Person)

2.4.2.5 Citation Requirement

2.4.2.7 Access Conditions

2.4.2.8 Disclaimer

Section 3: File Description

File Contents

3.1.2 File Contents
- abstract or description of the data file

Data Check

3.1.8 Extent of Processing Checks

Version Statement

3.1.12.1 Version Text

3.1.12.2 Version Responsibility

3.1.12.3 Version Notes

Section 4: Variable Description

Question

4.2.8.1 Pre-Question Text

4.2.8.2 Literal Question

4.2.8.3 Post Question

4.2.8.6 Interviewer Instructions

Universe

4.2.12 Universe

Question
4.2.16  Standard Categories
4.2.16  Standard Categories Date
4.2.19  Coding Instructions
Notes
4.2.24  Notes
APPENDIX B — WHEN SHOULD I ASSIGN WEIGHTS TO INDIVIDUAL VARIABLES?

Where there is a clear one-to-one match between a weight variable and a survey variable, assign the appropriate weight to the appropriate variable.

Where there are multiple (>1) weights that could be assigned to one survey variable, do not assign the weight. In the Canadian Election Surveys, for example, a question about political ethics could have both the provincial and national weight applied to it. We need to leave this decision to the user. Skipping this step only affects the display of each variable in the “Description” tab in Webview (the view that shows frequencies or marginals for a given variable). So, if weight is assigned in Nesstar, N and NW (sample count and weighted counts, respectively) appear in the “Description” tab. If weights are not applied, only the N is seen.

An example of a file with multiple weights where you would apply these weights to the variables is the Violence Against Women Survey. This survey has two weights – PERWGHT and INCWGHT. In the documentation it states that the PERWGHT weight variable is to be used for questions about the individual and the INCWGHT weight variable is to be used for questions about incidents. This is noted in the Weight field (2.3.1.12) and in the Notes field (4.2.24) for each of the variables. When you click on a variable and look at the marginal you will see the weight listed that has been used.
APPENDIX C – CHANGING ALPHANUMERIC VARIABLES TO NUMERIC VARIABLES IN THE SPSS SYNTAX FILE

In some instances, variables will be coded as alphanumeric when they are really numeric variables. In these cases, the variables need to be changed to numeric in the SPSS syntax file.

Alphanumeric variables are noted by the ‘(A)’ beside the variable in the Data List. To check if the variable is truly an alphanumeric variable, scroll down to the Value Labels and look at the values for the variables. If the value of a variable is listed in quotes (‘a’ “Excellent” or ‘1’ “6 months”) then the variable is coded as alphanumeric. If the value is a letter, then it is fine; if the value is a number, then it needs to be changed.

To change a variable from alphanumeric to numeric:

i) remove the ‘(A)’ from the Data List
ii) remove the quotes from around the value in the Value Labels list
iii) save file and run syntax file with data file to create new .sav file
iv) import the new .sav file into Publisher

The first two variables here are coded as alphanumeric. The third is numeric. By looking at the Values Labels, you can tell which variables are coded correctly.

Since the values for Q14MS are numbers, then it should be coded as numeric. The quotes need to be removed from the numbers here and the ‘(A)’ in the Data List will have to be removed.

Since the values for Q15FAM are letters, then the variable is coded correctly as alphanumeric.

This variable is coded correctly as numeric.
APPENDIX D — ERROR WARNING WHEN IMPORTING DATA

In some cases when you import data, Nesstar may give the following warning:

Click on Yes and the data will import properly.

The reason for this error is that Publisher has a set of rules that it follows for measure values. If the file being imported doesn’t match the rules, this message is displayed.

The following rules apply:

- String (alphanumeric) variables are set to nominal
- String and numeric variables with defined value labels are set to nominal
- Numeric variables without defined value labels but less than a specified number of unique values are set to ordinal
- Numeric variables without defined value labels but more than a specified number of unique values are set to scale

For SPSS `.sav` files Publisher uses the defined measure unless this is very different from what seems logical. In these instances the user will be asked during the import process, whether the measure definitions should be changed.

(Nesstar Publisher v4.0 User Guide - http://www.nesstar.com/help/4.0/publisher/extras/downloads.html)