

DTS Excel Wizard User Guide

Version 4.7

Introduction

The DTS Excel Wizard is an application for adding DTS terminology information to Excel files. Using existing cell data, such as Concept names or Concept codes, the Excel Wizard places associated Concept information into other cells. New columns can also be added. The Excel Wizard provides an easy-use alternative to direct API programming for populating spreadsheet information. The modification script (program) can optionally be saved in a file for reuse. The Excel Wizard is available both as a plug-in to the DTS Editor, and as a standalone application.

Installation

Extract the files in `ExcelWizard-4.7.zip` into your `DTSInstall` directory. Be sure the `Use folder names` box is checked. This will place all the Excel Wizard files into the appropriate folders:

Folder

`DTSInstall\bin\excelwizard`
`DTSInstall\lib\modules`

Files

`ExcelWizard.bat`
`excelwizard.jar`
`apelwizard.jar`
`xmldigester.jar`
`pluginutils.jar`
`poi-3.6-20091214.jar`
`poi-ooxml-3.6-20091214.jar`
`poi-ooxml-schemas-3.6-20091214.jar`
`xmlbeans-2.3.0.jar`
`excelwizarduserguide.pdf`
`excelwizarduserguide.htm`

`DTSInstall\docs`

`DTSInstall\docs\help`

Running the Excel Wizard

The ExcelWizard can be run in two ways:

1. As a plug-in to the DTS Editor by selecting the `New Excel Wizard` option in the Tools menu. See ***Using the DTS Excel Wizard Plug-In*** section below for details.
2. As a stand-alone application by executing `ExcelWizard.bat` in the `DTSInstall\bin\excelwizard` folder. See ***Using the DTS Excel Wizard Application*** section for details.

Preparing the Excel File

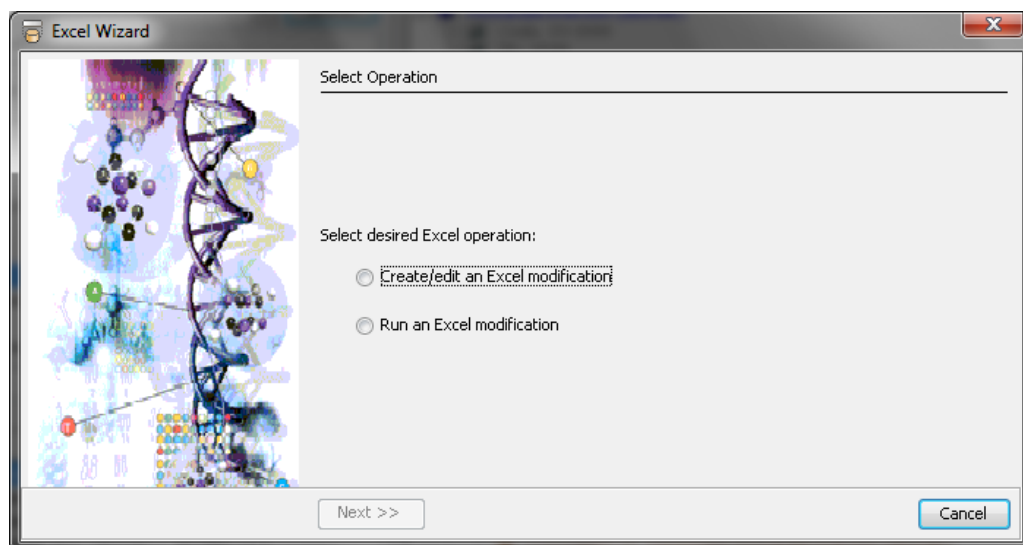
The Excel spreadsheet file to be processed must be structured so that the cell values in one column uniquely identify a DTS Concept. The values can be Concept Names, Concept Codes, Concept Ids, or the value of a Concept Property that uniquely determines a Concept. The Concept referenced in each row is called the *Key Concept* for the row and is the Concept from which the Attribute values to be added are retrieved.

Using the DTS Excel Wizard Plug-In

After selecting the New Excel Wizard DTS Editor menu option, you will be lead through a set of wizard pages. Click **Back** to go back to the previous page, **Next** to go to the following page. You may click on **Cancel** at any time to exit the wizard. Some options may not be available until certain selections are made in the page. The following sections describe the individual Excel Wizard pages.

Select Operation

Two options are available on the initial page: **Create/edit an Excel modification**, and **Run an Excel modification**. Select the first option to create or edit a modification specification. If this option is selected, a modification using the specification can be run without restarting the wizard. The second option displays the wizard pages to immediately perform a modification. If this option is selected, a modification specification must have already been created and saved to a file.



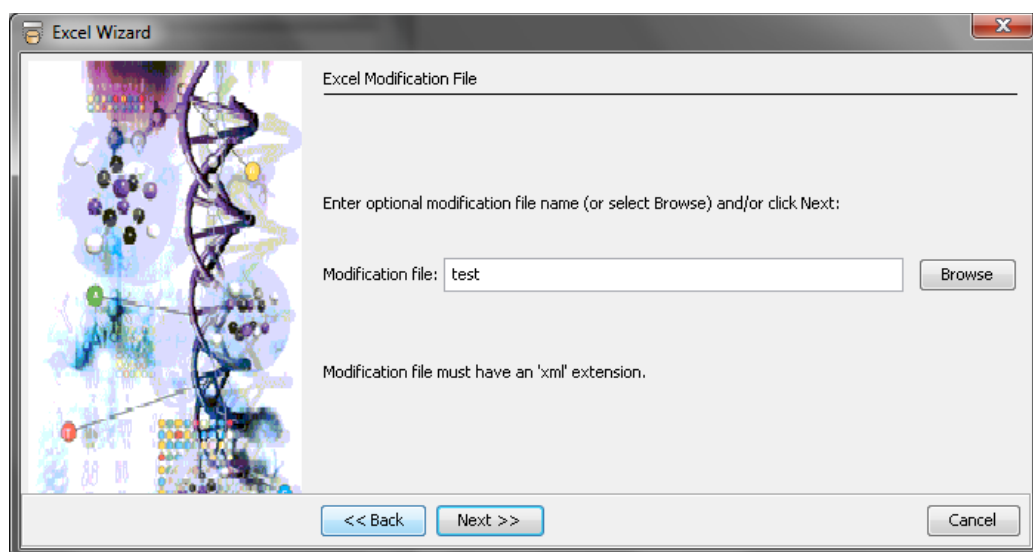
After selecting an option, click **Next**. If **Create/edit an Excel modification** was selected, see the **Specify a Modification File** section for further instructions. Otherwise go to the **Select Excel File** section.

Specify a Modification File

A modification specification is a list of the changes that are to be made to an Excel file. It is the “instructions” for updating the file, e.g. “put the Code in Source value associated with each Key Concept into cell E”. The modification specification can be saved to a file for later editing or execution. This is an optional step. Modification specification files are XML files and can theoretically be created in any standard editor, but it is recommended that these files be created and edited (or at least validated) through the Excel Wizard to avoid program errors.

If saving is desired, enter the path and name of a new or existing modification file in the text field (shown below), or click the **Browse** button to search the file hierarchy. The wizard will supply a default extension of “.xml”.

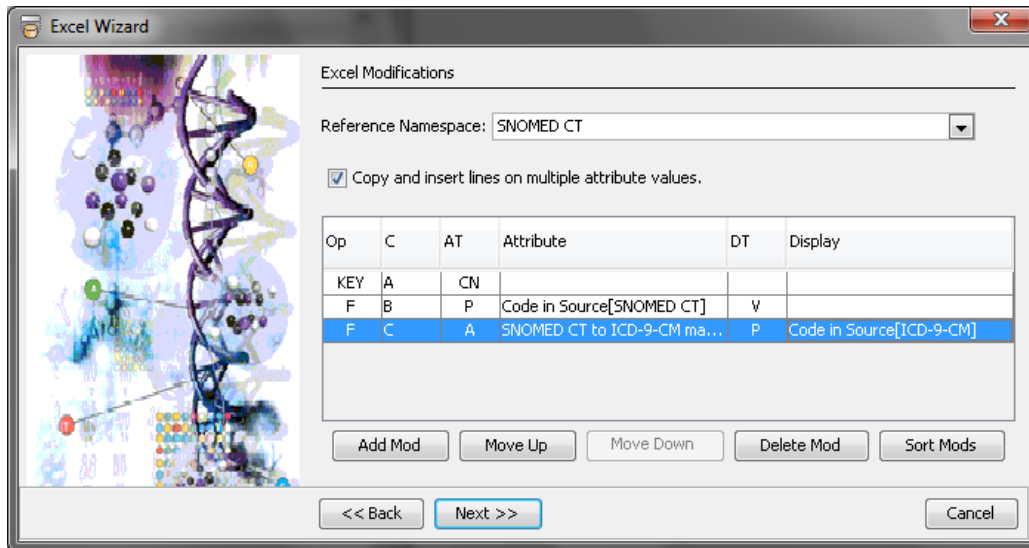
Click **Next** to continue.



Define the Modification

An Excel modification is defined by specifying (i) the Namespace in which the Key Concept will be found, (ii) information on the location (the column in the spreadsheet) and representation of the Key Concept identifier, and (iii) information on the desired location and representation of the DTS Attributes to be added to the spreadsheet. These pieces of data constitute the modification specification which will drive the updating of the spreadsheet and, if a modification specification file was selected, be saved in the file for later use.

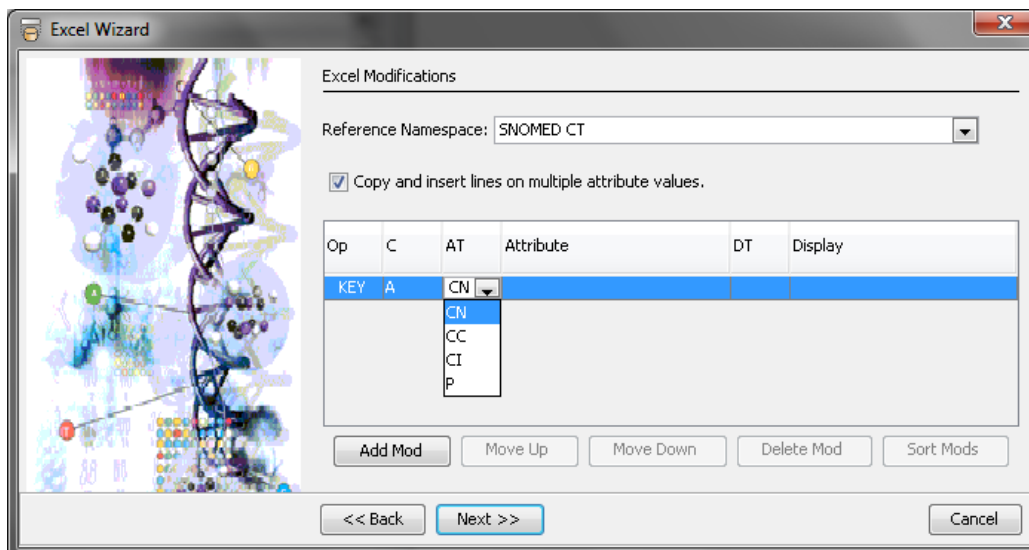
The Key Concept Namespace is selected from the Combo Box at the top of the **Excel Modifications** page (see screen shot below), while the Key Concept and DTS Attribute parameters are entered in the table at the center of the page. The **Copy and insert lines on multiple attribute values** checkbox should be checked if Excel rows should be copied and inserted when multiple instances of requested Attributes are found. The number of rows to be added and the cell population logic is a function of the number of Attribute instances and their relative dependencies. The logic is similar to that used in TQL. If this box is not checked, duplicate rows will not be created and only the first instance of an Attribute will be placed in a cell.



After completing the top of the wizard page, the Key Concept parameters should be selected. These parameters are in the first table row whose initial cell is **KEY**. For the Key Concept, the **C** (Column), **AT** (Attribute Type) and, optionally, **Attribute**, columns are used. Double click on the **C** cell in the **KEY** row to select the Excel column where the Key Concept identifier will be found. (This is the column in the original (unmodified) spreadsheet.) Then double click on the **AT** cell to select the Attribute Type (see below). The available Key Concept Attribute Types (representations) are:

- CN – the value in the spreadsheet cell is the Concept Name of the Key Concept
- CC – the value in the spreadsheet cell is the Concept Code of the Key Concept
- CI – the value in the spreadsheet cell is the Concept ID of the Key Concept
- P – the value in the spreadsheet cell is the value of a (unique) DTS Property of the Key Concept

If **P** is selected for the Attribute Type, then double click the **Attribute** cell and select the appropriate DTS Property Type (from the Key Concept Namespace) in the drop-down list.



This completes the Key Concept specification.

Table rows below the **KEY** row describe the DTS Attributes to be added to the spreadsheet. To add a new Attribute row, press the **Add Mod** button. This will add a new row below the selected row, or at the end of the table if no row is currently selected. (To select a row, click once on any cell in the row.) The new row's **Op** cell will be set to **F** (Fill) and the row's **Column** cell will be set to the first available column (a column not specified in the modification table) after the column of the previous row. These values are only defaults and can be changed as described below.

Attribute row cell descriptions are given below. Double click on any cell to select its value from the displayed dropdown. Cells should be completed in a left to right order since some cells are only editable for specific values of previous parameters.

Op	The default value for the Op (Operator) cell is F for Fill. This means fill the existing spreadsheet column (specified in C) with the designated Attribute value. The alternate value is IF for Insert and Fill. This means that a new column (given by parameter C) will be added to the spreadsheet and filled with the Attribute value. See the Using Insert and Fill box below.
C	As in the Key Concept row, select the column where the Attribute value should be placed. Further information on the interpretation of the column parameter can be found in the Using Insert and Fill box below.
AT	Select the Attribute Type, i.e., the type of Attribute to be placed in the cell. This can be any of the following Type (representation) codes: CN – the value placed in the spreadsheet cell is the Concept Name of the Key Concept CC – the value placed in the spreadsheet cell is the Concept Code of the Key Concept CI – the value placed in the spreadsheet cell is the Concept ID of the Key Concept CS – the value placed in the spreadsheet cell is the Concept Status of the Key Concept PN – the value placed in the spreadsheet cell is the Preferred Name of the Key Concept. An empty value is used if there is no Preferred Name. RN – the value placed in the spreadsheet cell is the Resolved Name of the Key Concept, i.e., the Preferred Name of the Key Concept or the Concept Name if there is no Preferred Name. S – the value placed in the spreadsheet cell is the value of a Synonym of the Key Concept P – the reference value is the value of a Property of the Key Concept; see the DT and Display cells for further information SP – the reference value is the Concept value of a Superconcept; see the DT and Display cells for further information (only available if reference Namespace is Ontylog or Ontylog Extension) SB – the reference value is the Concept value of a Subconcept; see the DT and Display cells for further information (only available if reference Namespace is Ontylog or Ontylog Extension) DC – the reference value is the Concept value of a Defining Concept; see the DT and Display cells for further information (only available if reference Namespace is Ontylog or Ontylog Extension) DR – the reference value is the Concept value of a Defining Role; see the DT and Display cells for further information (only available if reference Namespace is Ontylog or Ontylog Extension)

- R – the reference value is the Concept value of a Role of the Key Concept; see the DT and Display cells for further information (only available if reference Namespace is Ontylog or Ontylog Extension)
- IR - the reference value is the Concept value of an Inverse Role of the Key Concept; see the DT and Display cells for further information (only available if reference Namespace is Ontylog or Ontylog Extension)
- A – the reference value is the Concept value of a Concept Association of the Key Concept; see the DT and Display cells for further information
- IA - the reference value is the Concept value of an Inverse Concept Association of the Key Concept; see the DT and Display cells for further information

Attribute If the Attribute Type was specified as S, P, DR, R, IR, A or IA, select the name of the DTS Synonym Type, Property Type, (Inverse) Role, or (Inverse) Association Type whose value should be used. Note that Types from all Namespaces are available to be selected.

DT The Display Type. If the Attribute Type was specified as P, select from the following two options:

- V – the value placed in the spreadsheet cell is the value of the Property whose Type is the Attribute Property Type
- PQ - the value placed in the spreadsheet cell is the value of the Property Qualifier (whose Type is given in the Display cell) on the Property whose Type is the Attribute Property Type.

If the Attribute Type was specified as DC, DR, R or IR, select from the following options:

- CN – the value placed in the spreadsheet cell is the Concept Name of the target of the Attribute Type
- CC – the value placed in the spreadsheet cell is the Concept Code of the target of the Attribute Type
- CI – the value placed in the spreadsheet cell is the Concept ID of the target of the Attribute Type
- CS - the value placed in the spreadsheet cell is the Concept Status of the target of the Attribute Type
- PN – the value placed in the spreadsheet cell is the Preferred Name of the target of the Attribute Type
- RN – the value placed in the spreadsheet cell is the Resolved Name of the target of the Attribute Type
- P – the value placed in the cell is the value of the Property (whose Type is given in the Display cell) on the target of the Attribute Type

If the Attribute Type was specified as A or IA, select from the following options:

- CN – the value placed in the spreadsheet cell is the Concept Name of the target of the Attribute (Inverse) Association Type

- CC – the value placed in the spreadsheet cell is the Concept Code of the target of the Attribute (Inverse) Association Type
- CI – the value placed in the spreadsheet cell is the Concept ID of the target of the Attribute (Inverse) Association Type
- CS - the value placed in the spreadsheet cell is the Concept Status of the target of the Attribute Type
- PN – the value placed in the spreadsheet cell is the Preferred Name of the target of the Attribute (Inverse) Association Type
- RN – the value placed in the spreadsheet cell is the Resolved Name of the target of the Attribute (Inverse) Association Type
- P – the value placed in the cell is the value of the Property (whose Type is given in the Display cell) on the target of the Association whose Type is the Attribute (Inverse) Association Type
- AQ – the value placed in the spreadsheet cell is the value of the Association Qualifier (whose Type is given in the Display cell) on the Association whose Type is the Attribute (Inverse) Association Type

Display If PQ, P, or AQ was selected for the Display Type (DT), select the display Attribute Type from the list of Property Qualifier Types, Property Types, or Association Qualifier Types.

To delete an Attribute row from the specification, select the Attribute row as described above and press the **Delete Mod** button.

Using Insert and Fill

The ability to insert columns complicates the interpretation of the Attribute column (C) parameter. It is first necessary to understand how the Excel Wizard handles insertions. Before any rows are processed, the wizard inserts new columns (as specified in each I&F row's column parameters) into the spreadsheet in the order they appear in the specification. A new column is inserted at the position given by the column parameter. Each insertion pushes subsequent columns out one position. Then, each row is processed by filling the cell, given by the row's column parameter, with the appropriate Attribute value. Thus, all Attribute column parameters are interpreted as the column position in the final (perhaps extended) spreadsheet.

This procedure means that the interpretation of each I&F column parameter is relative to the history of previous insertions, i.e., the order of the rows is relevant. It also complicates recognition of conflicts. The sequence I&F: 5, I&F: 3, F: 6 results in a column conflict, since after the new column 3 is added, the originally added column 5 is the same as the final column 6. While these conflicts are detected and reported by the wizard, they are more easily seen if the Attribute rows are kept in increasing column order. The **Sort Mods** button is available for this reason.

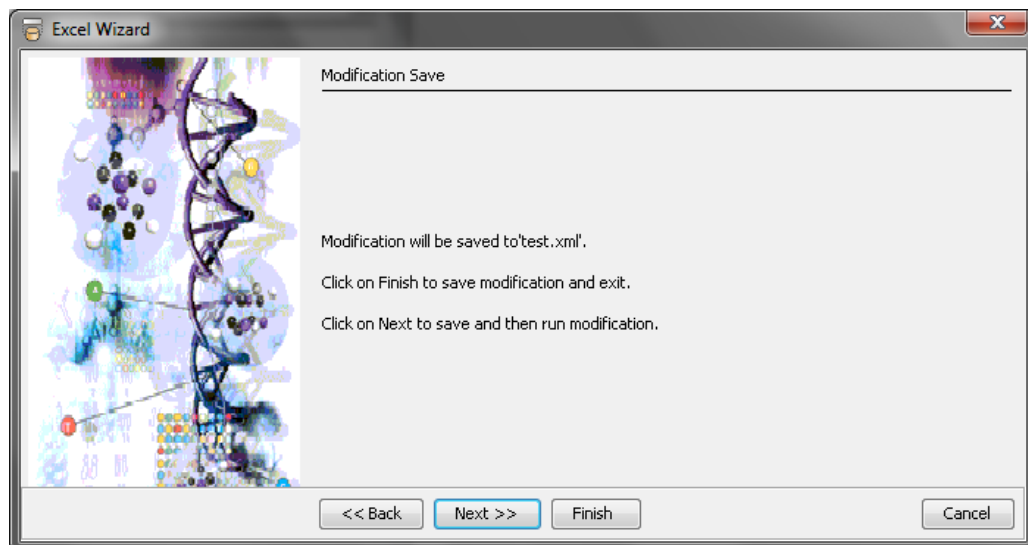
The **Move Up**, **Move Down**, and **Sort Mods** buttons can be used to rearrange the rows. **Move Up** and **Move Down** move the selected row up and down one position, respectively. **Sort Mods** sorts the **Attribute** rows in increasing column order. See the **Using Insert and Fill** box above for more information on row sorting.

Press **Next** when the modification specification is completed. Before proceeding, the modification is checked for errors, e.g., multiple rows referring to the same column or overwriting of the **Key Concept** column, and a message is shown if any are found.

If a modification file was specified, see the **Confirm Modification Save** section. Otherwise, go to **Select Excel File**.

Confirm Modification Save

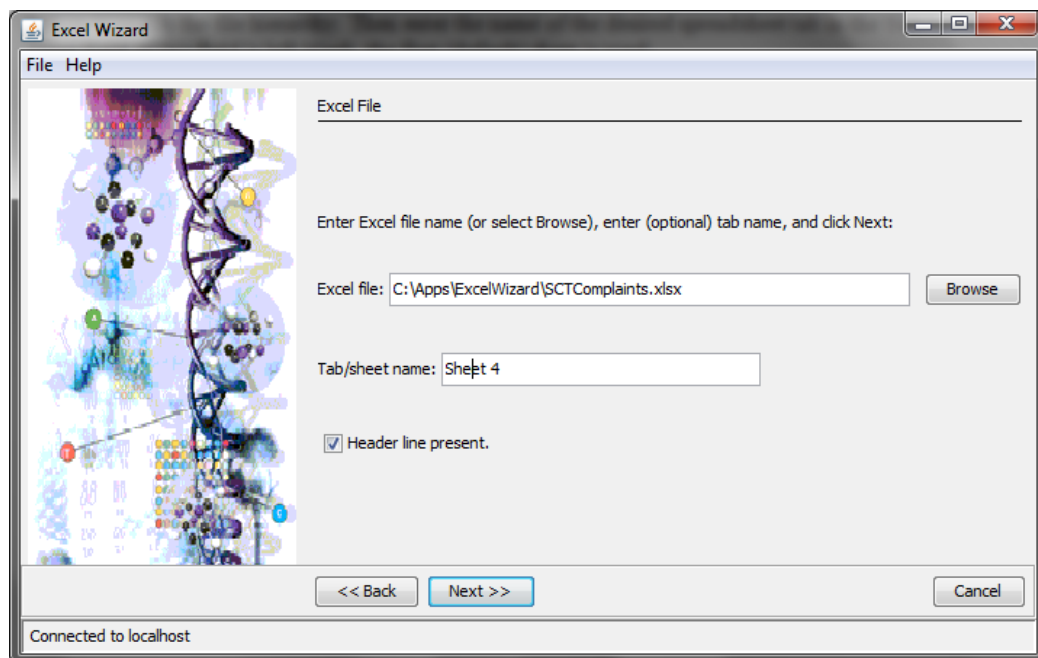
Click **Finish** to save the modification and exit the wizard. Otherwise click **Next**.



Select Excel File

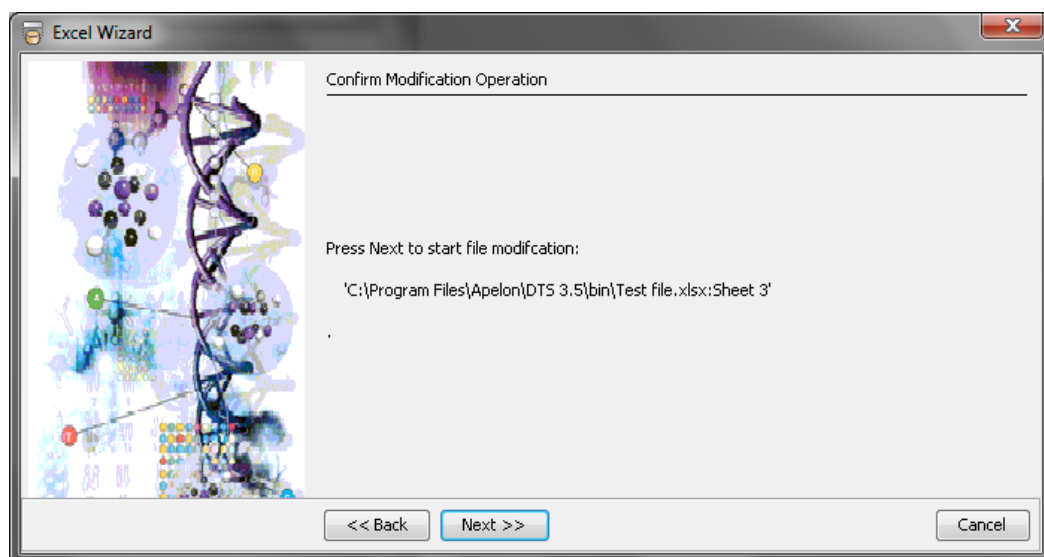
Enter the path and name of the target Excel file (.xls or .xlsx) in the first text field, or click the **Browse** button to search the file hierarchy. Then enter the name of the desired spreadsheet tab in the **Tab/sheet name** field. If this field is left blank, the first (default) sheet is used. Check the **Header line present** box if the sheet has a header line. If checked, column labels will be added for the modifications.

Click **Next** to continue.



Confirm Modification

Click **Next** to perform the modification on the named Excel file.

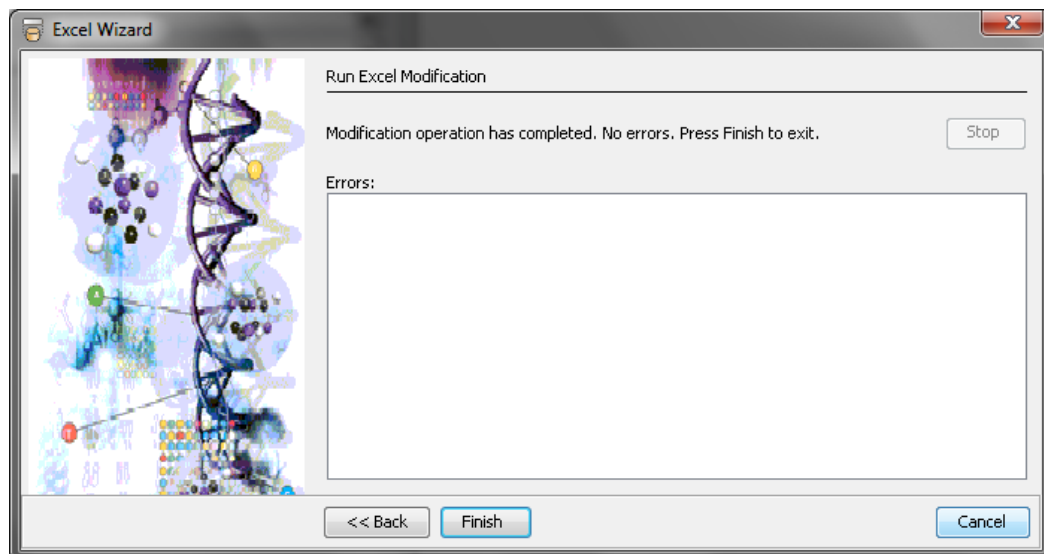


Run Excel Modification

The modification is run immediately. Press the `Stop` button if it is desired to prematurely terminate processing. Any run-time errors are reported in the error window. During execution, the Excel Wizard writes a log file, `ExcelWizard.log`, in the default directory. This file contains a record of all modification actions including errors detected.

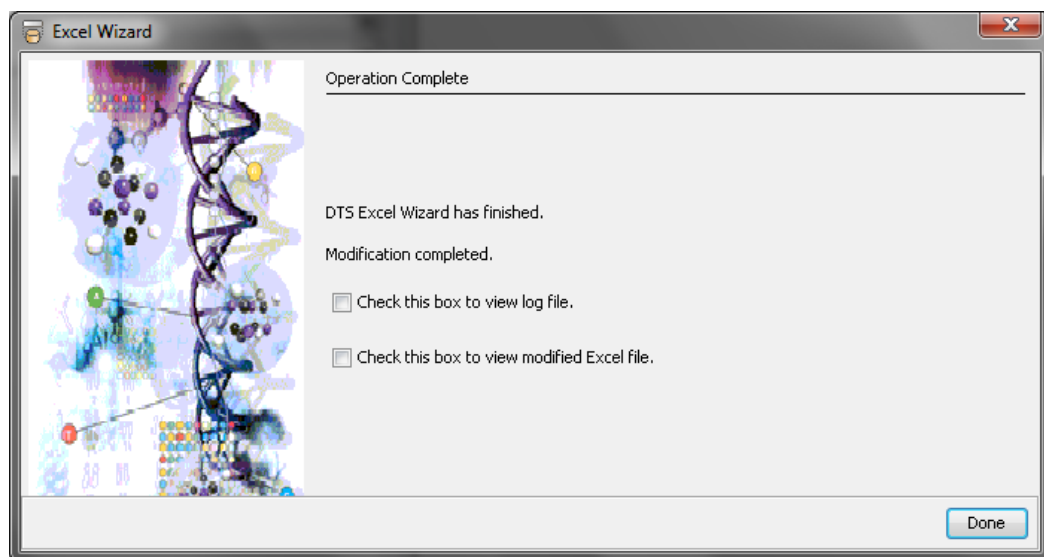
Note: The `ExcelWizard.log` log file is overwritten after each run. The log file should be renamed if it needs to be accessed after another modification.

Click `Finish` to complete the wizard.



Complete Wizard

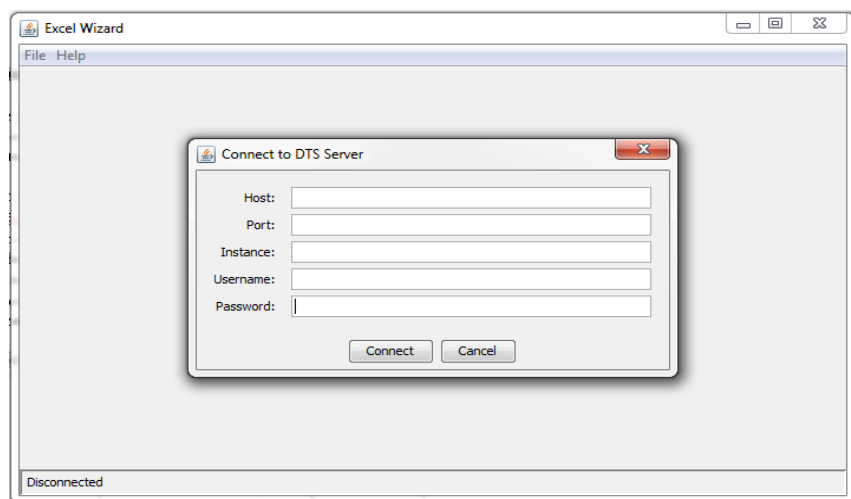
If a modification was run, the log file can be viewed by selecting the first checkbox, and the modified Excel file can be viewed by checking the second. Click `Done` to exit, and view any selected files.



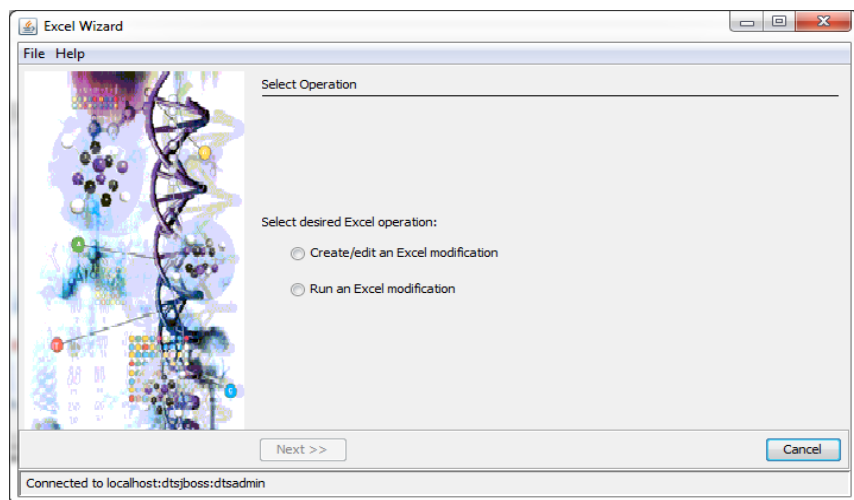
Using the DTS Excel Wizard Application

The ExcelWizard application (see screen shots below) provides a “standalone” presentation of the wizard pages. The application window consists of a central Wizard Area which displays the wizard pages, with a Menu Bar above and Status Line below.

Run the Excel Wizard application by executing the `ExcelWizard.bat` file found in `DTSInstall\bin\excelwizard`. When the application starts (see first screen shot below), it opens a Connect to DTS Server dialog. Fill in the associated parameters, and click Connect. If successful, the host name will be shown in the application’s Status Line, and the wizard’s Select Operation page appears in the Wizard Area (see following screen shot). Wizard page flow and operation is identical to that described in **Using the DTS Excel Wizard Plug-In**.



The Menu Bar gives access to additional application functions. The File menu includes the Connect, Disconnect, and Exit options. Connect opens the Connect to DTS Server dialog, while Disconnect breaks the connection. Note that disconnecting terminates the active wizard operation. The Help menu includes options to display the **Excel Wizard User Guide** (this document) and show an explanatory About dialog.



Revision History

Version 1.0	Initial release.
Version 1.1	Support for Defining Concepts, Defining Roles, and (Inverse) Roles.
Version 1.2	Evaluation of Inverse Associations corrected.
Version 1.3	Support for DTS V4 login by ExcelWizard.bat. Addition of the Concept Status (CS) Attribute Type for concepts in all Namespaces. Addition of the Superconcept (SP) and Subconcept (SB) Attribute Types for concepts in Ontylog and Ontylog Extension namespaces.
Version 1.4	Updates for DTS V4.1.
Version 1.5	Bug Fixes. Setting of default modification column on row insert.
Version 4.6	Version renumbered for consistency with DTS. Support for Extension Roles added. Improved attribute load performance.
Version 4.7	Updated poi (Excel) utilities to correct formatting errors when rewriting spreadsheets. The versions are now 3.9 (but same names).