

Import Wizard

Import Wizard	1
Formatting items.....	1
Formatting dates	1
Formatting latitude/longitude coordinates	2
Order of records in an import file	2
Adding traits	2
Setting default items	2
Import options	2
Delimited Text.....	3
Excel Worksheet.....	3
Errors found during import.....	3
Importing references.....	4

Import Wizard

The Import Wizard is used to import existing, electronic information into a BioLink database. The Wizard is started by selecting the **Tools→Import→Taxa and Material records** menu item. Currently the information being imported can be stored in a text files and Excel spreadsheets. The Wizard presents a series of steps or question during the import. These are detailed below.

The first step in importing is to select a data source type. A list of the available formats is presented. Select the required type and either double click or press the **Next** button. This will open a window where information about the import file can be specified. See *Import Options* below for details concerning these windows. Once the import file has been selected and its options set press the **OK** button. This will close the Import Options window and return to the Import Wizard.

The Import Wizard will now contain two lists of data items, one from the file being imported and one for those available within BioLink. The next step is to map or link the items being imported with those in BioLink. Pressing the **Auto** button (found between the two lists) will attempt to make these mappings based on the names of the import items. Items that cannot be guessed at will be left blank. To link these items, locate the corresponding item in the BioLink list and press the < button. This will add the name of the BioLink item in the Destination column next to the import item it corresponds with. Continue this process for all items to be imported (items which are unmapped will be ignored during the next step). To remove a mapping use the > button or to remove all mappings press the >> button. When complete, press the **Next** button.

At this stage the mappings can be saved to a template. This template can be used in the future if data with the same structure needs to be imported. Press the **Save Template** button and specify a name for the template file. Pressing the **Start Import** button will begin the actual import. A progress log will be displayed during the import process, showing the status of the import. When the import is complete this log file can be saved if required. Once complete, press the **Close** button to exit the Import Wizard.

Formatting items

Most items can be formatted as required and BioLink will interpret them correctly during import. However, special care should be take with dates and latitude/longitude coordinates.

Formatting dates

For best results, dates should take the format YYYYMMDD (or YYYYMM or YYYY for incomplete dates) as this is the internal format used by BioLink. Dates in other formats, such as 4 May 1987, May 1987, 4/5/87 or 4-

5-87, can be used but the results of these may not be as expected. BioLink will attempt to interpret these and convert them to proper dates, but this can be a difficult task and may have an unacceptably high error rate.

Formatting latitude/longitude coordinates

The preferred format for latitude/longitude coordinates is decimal degrees with northern latitudes positive, southern latitudes negative, western longitudes negative and eastern longitudes positive (this is the internal format used by BioLink). Other formats can be used, such as degrees-minutes-seconds-direction, but BioLink will need to interpret these and convert them to decimal degrees during the import. This conversion step may introduce errors.

Order of records in an import file

The order of records in the import file can have a significant impact on the performance of the Import Wizard. This is because BioLink uses an internal cache to store recently imported data. Optimal use of this cache means less searching of the database, a slow operation compared to reading memory.

BioLink stores all taxon names in memory so taxa can appear in any order in the import file. However, BioLink only stores the most recent material record in memory. If the next record being imported shares the same site and site visit as the previous record, BioLink can use this site and site visit without having to find them (again) in the database. By sorting the import file by sites followed by site visits, BioLink can reduce the number of times it has to search the database. This can reduce import times by as much as one third.

Adding traits

Often there will be imported fields that do not correspond with existing BioLink items. In these cases the imported items should be stored as traits. Traits can be attached to Sites, Site Visits, Material and Taxa. To create a trait, map the imported field to the Other item under the appropriate group. This will automatically create a trait for the item being imported.

Setting default items

In some cases it is desirable to preset a value for all records being imported. There are two ways to do this: set a Default Value or add a Fixed Field. A default value is set for a field in the import file. When set, the default value will override values found in the import file. A fixed field is a new field added directly in the Import Wizard. This new field has a default value and is mapped to a BioLink item in the same manner as fields in the import file.

To add a Default Value, select a field, click the right mouse button, select the **Set Default Value** command, enter the desired default and press the **OK** button. To clear this value click the right mouse button and select the **Clear Default Value** command.

To add a Fixed Field, click the right mouse button, select the **Add New Fixed Field** command, enter the default value for the new field and press the **OK** button. This will add the new field to the end of the list of fields being imported. The field can then be mapped to an existing BioLink field. To change the default value of a fixed field click the right mouse button and select the **Edit Fixed Value** command. To delete a fixed field click the right mouse button and select the **Delete Fixed Value** command.

Import options

BioLink supports a number of formats when importing data from sources outside BioLink. The following formats are currently supported.

Delimited Text: A delimited text file.

Excel Worksheet: A Microsoft Excel worksheet.

A number of options are available for each of these formats. Details of the available formats are listed below. With all import formats, pressing the **Refresh Preview** button will show how the Import Wizard has interpreted the import file. If this is incorrect the import options or import file may need to be modified.

Delimited Text

This format imports data from a delimited text file with one record per row. Each field or data item needs to be separated from others by a unique character that is not found in any of the fields. A tab or pipe (‘|’) characters are generally used. The values should not be enclosed in quotes or double quotes. The first row should contain the names of the columns but this is not essential. If present, these names will be used by the Import Wizard during mapping to the BioLink database. If not present, the fields will be sequentially numbered.

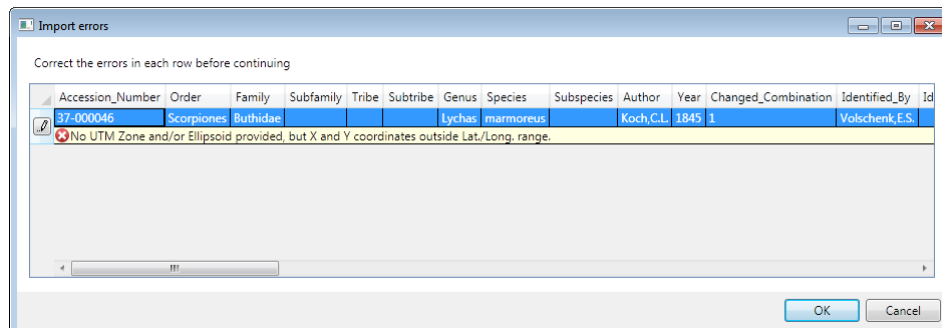
Excel Worksheet

This format imports data from a Microsoft Excel worksheet. The information to be imported must be on a single worksheet with each row representing a single record. The first row should contain the names of the columns but this is not essential. If present, these names will be used by the Import Wizard during mapping to the BioLink database. If not present, the fields will be sequentially numbered.

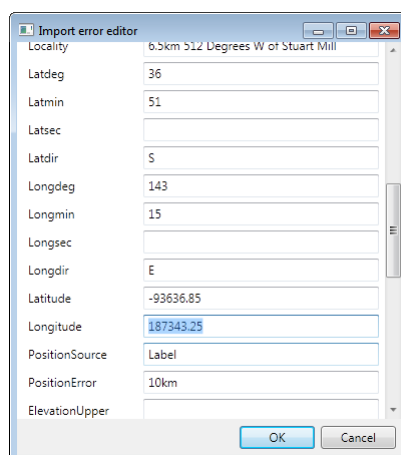
Errors found during import

If errors are encountered in the import file while importing, the offending row(s) are written to a SQLite database along with the cause of the error (where possible). The name and location of this database is given in the import log file. There is a facility within the Import Wizard that will allow you to examine errors, fix any problems and reimport the rejected rows.

Restart the Import Wizard and select the **Import Error Database** as the data source. Find the error database file using the Open File dialog box and press the **Open** button. Each rejected row will appear in a grid, and as each row is selected, the error that caused the row to be rejected will be displayed.



Using the error message as a guide, correct the row by either double clicking a single cell to change its value, or click on the pencil icon to edit the entire row in a separate window.



Once all rows have been corrected, click ok to continue through the Import Wizard as normal. Only the corrected rejected rows will be imported. Note that this process may create another error database if any rows remain in error.

Note that the taxonomic hierarchy is not maintained during importing. This allows easier data entry during importing and reduces the number of error rows produced. Because of this, imported taxonomic data should be checked carefully and problems resolved after import.

Importing references

The Reference Import Wizard is used to import reference information into a BioLink database. The Wizard is started by selecting the **Tools→Import→References** menu item. The information being imported can be in delimited text files or excel spreadsheets.

The first step in importing is to select an import file. Press the ellipsis button to the right of the file name field to open the File Open dialog. Once the import file has been selected press the **OK** button. This will close the Import Options window and return to the Import Wizard.

The Reference Import Wizard will now contain two lists of data items, one from the file being imported and one for those available within BioLink. The next step is to map or link the items being imported with those in BioLink. Pressing the **Auto** button (found between the two lists) will attempt to make these mappings based on the names of the import items. Items that cannot be guessed at will be left blank. To link these items, locate the corresponding item in the BioLink list and press the < button. This will add the name of the BioLink item in the Destination column next to the import item it corresponds with. Continue this process for all items to be imported (items which are unmapped will be ignored during the next step). To remove a mapping use the > button or to remove all mappings press the >> button. When complete, press the **Next** button.

At this stage the mappings can be saved to a template. This template can be used in the future if data with the same structure needs to be imported. Press the **Save Template** button and specify a name for the template file. Pressing the **Start Import** button will begin the actual import. A progress log will be displayed during the import process, showing the status of the import. When the import is complete this log file can be saved if required. Once complete, press the **Close** button to exit the Wizard.

It should be noted that import process, apart from the field mappings, is identical to the more general import wizard for Taxa and Material, including the error handling facility.