# SynBio2Easy – how to use … from the perspective of various scenarios

Installation and System Requirements

Currently, SynBio2Easy does not have a Graphical User Interface (GUI) and it is operated from the command line. Let’s begin with launching the program through the provided start-up scripts. If for any reason the start-up scripts do not work in your environment, there are instructions in Appendix C for opening a command prompt and starting the program manually.

## Starting a Command Prompt Window (MS Windows)

1. Click on the Windows ‘Search’ button and type in “PowerShell” to the search box, then click on the ‘Windows PowerShell’ icon.
2. A new window will be opened with a cursor where you can type in commands, we will call it “prompt” or “terminal” in this document.

## Starting a Terminal Window (MacOS)

1. Click on the small magnifying glass icon in your menu bar and type in “terminal.app” to the search box, then click on the ‘Terminal.app’ icon.
2. A new terminal window will be opened with a cursor where you can type in commands, we will call it “prompt” or “terminal” in this document.

## Checking if Java is installed

The tool needs Java installed in your system, so first check if you already have.

1. Open a prompt terminal
2. Type in:  
   `java -version`
3. If you see:  
   java version "11.0.3" or java version "13.x.x" or any other version higher than “11.x.x” than you are all set up
4. If an ‘unknown command’ error is generated, or you see an earlier version 10, or 1.8 then follow the instructions in Section 9 at the end of the document to install Java

## Installing SynBio2Easy

Installing the tool is as simple as downloading the appropriate package from the GitHub repository into your chosen location.

1. The GitHub repository for the SynBio2Easy is at <https://github.com/BioRDM/synbio2easy>. You can find the latest releases at <https://github.com/BioRDM/synbio2easy/releases>. Please download the most recent Java Archive (JAR) file from the latest release by clicking the ‘SynBio2Easy.jar’ link at the top of the page.
2. Next, please download the provided launcher script appropriate for your system: ‘SynBio2Easy.bat’ for Windows, or ‘SynBio2Easy.sh’ for Linux or MacOS
3. Create a directory for the tool, for example “C:\SynBio2Easy”
4. Copy the downloaded jar and script files into the directory

## Checking that everything works

1. In Windows Explorer or the equivalent file manager, navigate to the folder where you copied the files, e.g. “C:\SynBio2Easy”
2. Double-click on the start-up script
3. A new terminal window will open and you should now see the SynBioHub-Toolkit welcome message displayed at the top of the console window
4. If this did not work as expected, please follow the commands in Appendix C for launching the program manually without the start-up script

Metadata Input File

Metadata that describe the biological designs to be generated or annotated in SBOL documents and deposited in SynBioHub are specified for SynBio2Easy in a metadata input file. These data are expected to be in tabular format in an MS Excel spreadsheet file, including any of the column names as described in the list below.

* **Display\_id**: This is the only **mandatory** column in the spreadsheet. It is used as the primary key to match with component definition display IDs in SynBioHub and in SBOL documents.
* **Name**: The optional name of the design, which will be displayed as free text in the record in SynBioHub
* **Version**: The optional version of the component definition, which can be either numeric (e.g. 1.0) or free text (e.g. 1.0-alpha)
* **Attachment\_filename**: The absolute or relative (from the current working directory) path to a file to be attached to the design in SynBioHub
* **Summary**: The short description that will be written in the design’s description property of the SBOL component definition.
* **Description**: The text that will appear as the record description in SynBioHub (mutable description)
* **Notes**: The notes that will appear on the record in SynBioHub
* **Author**: The authors that will be listed in the design in the SBOL document
* **Key**: An entirely optional column that can be used to store a unique identifier string for a design, which can then be interpolated by SynBio2Easy when it is referenced in other columns, such as in the ‘display\_id’ and ‘name’ cells’ values

The cells in the spreadsheet columns support simple templating using keyword strings such as “{key}”, “{display\_id}” and “{name}”: the SynBio2Easy interpolation engine can then construct the target string value with the relevant values from those cells in the same row. In addition, standard Excel formulas are supported, for example ‘concatenate’ which provides powerful ways to combine values from other cells.

# Scenario 1 – Generating a New Design from a Metadata Definition File

In this scenario, the user intends to generate files describing their library of designs in SBOL document format. A template SBOL document is provided, along with the MS Excel spreadsheet containing metadata for each component design, and the program generates a new SBOL document based on the template for each component’s display\_id.

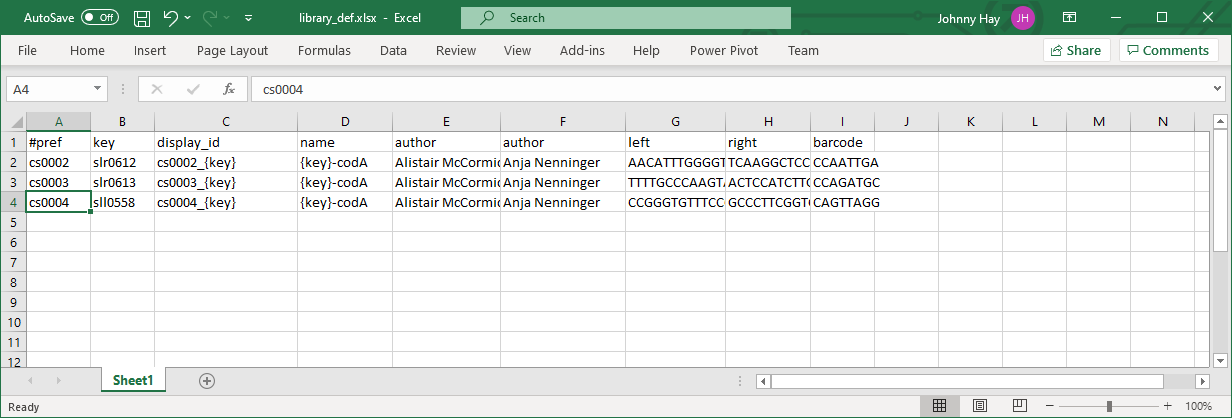
We will start the tool, specify where the input files are located, provide a filename prefix, version and target output directory.

## Preparation

It is assumed that the tool has been installed in ‘C:\SynBio2Easy’.

Copy the template component design SBOL document file named “design\_template.xml” into ‘C:\SynBio2Easy’ along with the Excel spreadsheet named “design\_metadata.xlsx” containing the columns of metadata. Your Excel file should contain one worksheet comprised of two or more columns: ‘display\_id’ and one or more of the optional metadata columns. The figure below shows an excerpt from a worksheet in an example input spreadsheet.

TODO: Insert screenshot of Excel metadata input file here



## Instructions for Generate

1. Double-click on the start-up script “C:\SynBio2Easy\SynBio2Easy-CLI.bat”
   1. Alternatively, use the instructions in Appendix C to start the program manually
2. Choose ‘generate’
3. Type in the name of the template file, “design\_template.xml” then <ENTER> when asked to specify the template file
4. Type in the name of the Excel file, “design\_metadata.xlsx” then <ENTER> when asked to specify the metadata input file
5. Press <ENTER> to specify the default ‘library’ filename prefix
6. Press <ENTER> to specify the default ‘1.0’ version string
7. Press <ENTER> to specify the default ‘library’ output directory

## Output

* The program will report the generation progress as each document is created
* Finally, the program will provide output to notify you that the SBOL documents have been created
* The program creates files under the ‘library’ output directory. The entire collection is split into multiple files, as SynBioHub could not import too many designs in one request.

Scenario 2 – Flatten Design Components and Sequence Hierarchy

In this scenario, the user intends to flatten the hierarchy of components and sequence annotations contained in an SBOL document. An input SBOL document file is provided, along with the output file path to contain the flattened SBOL component hierarchy.

We will start the tool, specify where the input file is located, provide an output filename, specify whether to flatten all root components found in the input SBOL document or just one specific design, and an optional suffix string to be appended to each flattened component ID.

## Preparation

It is assumed that the tool has been installed in ‘C:\SynBio2Easy’.

If you followed the instructions in Scenario 1 for generating a new design from a template and metadata file, the SBOL document files are expected to be located in the ‘library’ sub-directory, with a ‘library’ filename prefix.

## Instructions for Flatten

1. Double-click on the start-up script “C:\SynBio2Easy\SynBio2Easy-CLI.bat”
   1. Alternatively, use the instructions in Appendix C to start the program manually
2. Choose ‘flatten’
3. Type in the name of the input SBOL file, “library\library.1.xml” then <ENTER> when asked to specify the file containing the designs to be flattened
4. Type in the name of the output SBOL file, “library\library\_flattened.1.xml”
5. Press <ENTER> to accept the default behaviour to flatten all root components in the SBOL document
6. Type in a suitable suffix for each flattened component, e.g. “\_flat”

## Output

* The program will output how many designs were successfully flattened

Scenario 3 – Enrich SBOL Designs with Additional Metadata Annotations

In this scenario, the user intends to enrich their SBOL designs with extra descriptive metadata added into an existing file containing an SBOL document. An input SBOL document file is provided, along with an MS Excel file containing the ‘display\_id’ of each design to be enriched column and any other desired columns as described in the Metadata Input section.

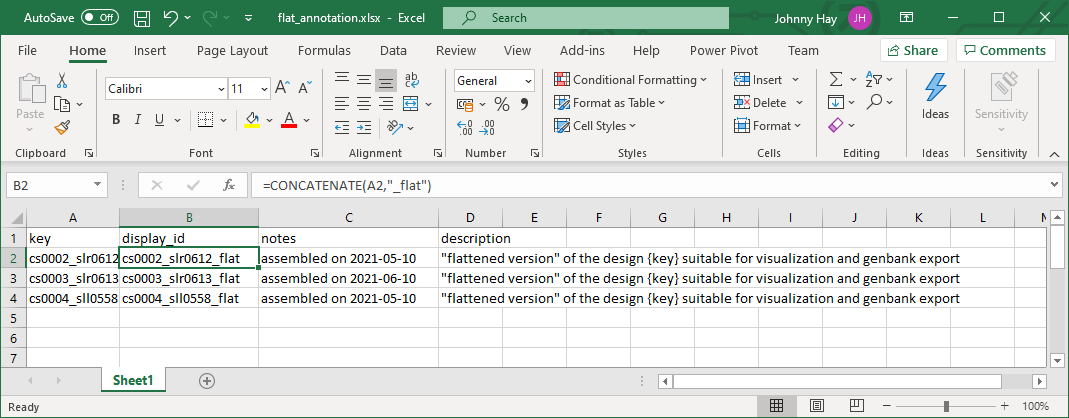
We will start the tool, specify where the input files are located, provide an output filename, specify whether to overwrite or append to existing metadata field values, and whether to halt the annotating procedure if either the input Excel file contains a component ID that is not found in the SBOL document, or if the Excel file contains blank values for metadata.

## Preparation

It is assumed that the tool has been installed in ‘C:\SynBio2Easy’.

If you followed the instructions in Scenarios 1 and 2 for generating a new design from a template and metadata file, then flattening them, the SBOL document files are expected to be located in the ‘library\’ sub-directory, with a ‘library\_flattened’ filename prefix. Your Excel file should contain one worksheet comprised of two or more columns: ‘display\_id’ and one or more of the optional metadata columns. The figure below shows an excerpt from a worksheet in an example input spreadsheet.

TODO: Insert screenshot of Excel metadata input file here



## Instructions for Annotate

1. Double-click on the start-up script “C:\SynBio2Easy\SynBio2Easy-CLI.bat”
   1. Alternatively, use the instructions in Appendix C to start the program manually
2. Choose ‘annotate’
3. Type in the name of the input Excel file, “library\annotations.xlsx” then <ENTER> when asked to specify the file with designs to be updated
4. Type in the name of the input SBOL file, “library\library\_flattened.1.xml”
5. Press <ENTER> to accept the default behaviour to append to any existing description or comment metadata fields if they already exist in the SBOL components
6. Press <ENTER> to accept the default behaviour to continue the annotating procedure without halting in case any missing component IDs are encountered
7. Press <ENTER> to accept the default behaviour to continue the annotating procedure without halting in case any missing metadata are encountered

## Output

* The program will report which designs were successfully annotated
* Any designs that were not found or which had missing metadata will also be reported

# Scenario 4 – Deposit SBOL Files into a New Collection in SynBioHub

In this scenario, the user intends to upload files describing their library of designs (for example, as generated in Scenario 1) to SynBioHub. All designs in a specified directory will be deposited into a new collection in SynBioHub.

We will start the tool, specify where the files are located, name the collection and provide necessary SynBioHub details.

## Preparation

It is assumed that the tool has been installed in ‘C:\SynBio2Easy’.

Create a folder (for example ‘library’) under the ‘C:\SynBio2Easy’ directory and copy all the designs files into the new folder. If you generated the library as in Scenario 1, the files are already present in ‘C:\SynBio2Easy\library’.

We assumed that all the data files have the standard ‘.xml’ extension, and that there are no other files in the ‘library’ directory.

## Instructions for Deposit

1. Double-click on the start-up script “C:\SynBio2Easy\SynBio2Easy-CLI.bat”
   1. Alternatively, use the instructions in Appendix C to start the program manually
2. Choose ‘deposit’
3. Type ‘library’ into the command prompt and then press <ENTER> when asked to specify the data directory (as our data directory is relative to where we are currently)
4. Press <ENTER> to select files of all types (since only the desired data .xml files are present in the directory anyway)
5. Enter ‘n’ if asked to create multiple collections
6. Enter ‘y’ when asked to create a new collection
7. Enter the name ‘synbio2easy’ for the collection
8. Press <ENTER> to specify the default ‘https://synbiohub.org’ target server
9. Press <ENTER> to accept the default ‘1.0’ version
10. Enter your SynBioHub username, usually your email address
11. Enter your SynBioHub password. **N.B.** the password input field is hidden so no text will appear as you type here

## Output

* The program will report the URL for the newly created collection
* The program will report which files were successfully deposited and those which failed
* Any files that failed to be deposited most likely failed because they contain invalid SBOL

Scenario 5 – Clean SBOL Files Downloaded from SynBioHub for Re-depositing

In this scenario, the user intends to clean SBOL document files that they have downloaded from SynBioHub, which are no longer valid for re-submission because of namespace adjustments. The cleaning process allows them to be deposited in SynBioHub again. An input SBOL document file to be sanitised is provided.

We will start the tool, specify where the input file is located, provide an output filename, specify any custom namespace URLs if desired, and instruct the tool whether to remove any references to existing SynBioHub collections in the output file.

## Preparation

It is assumed that the tool has been installed in ‘C:\SynBio2Easy’.

If you followed the instructions in the previous scenarios for generating a new design from a template and metadata file, then flattening, annotating and depositing them, you can download the SBOL files from the target SynBioHub instance. It is assumed the input SBOL file has been downloaded to the ‘library\downloads’ directory.

## Instructions for Clean

1. Double-click on the start-up script “C:\SynBio2Easy\SynBio2Easy-CLI.bat”
   1. Alternatively, use the instructions in Appendix C to start the program manually
2. Choose ‘clean’
3. Type in the name of the input SBOL file, “library\downloads\design.1.xml” then <ENTER> when asked to specify the document to be cleaned
4. Type in the name of the output SBOL file, “library\design\_clean.1.xml”, then <ENTER> when asked to specify the cleaned output file to generate
5. Press <ENTER> to accept the default namespace URL <http://biordm.sbs.ed.ac.uk>, or type in the desired custom namespace URL then press <ENTER>
6. Press <ENTER> to accept the default behaviour to remove any references to existing SynBioHub collections in the cleaned SBOL output document, or enter ‘N’ followed by <ENTER> to retain these references in the output

## Output

* If successful, the program will report the location of the cleaned SBOL document file

Scenario 6 – Prepare Template for Batch Update of Designs in an Existing SynBioHub Collection

In this scenario, the user intends to update a batch of multiple designs stored in an existing SynBioHub collection, all at once. This is done by generating a template Excel file that will contain the ‘display\_id’, ‘name’, ‘version’, and ‘uri’ values of the designs present in the collection, along with metadata columns that can be modified. The URL of an existing collection in the target SynBioHub instance is provided.

We will start the tool, provide an output filename for the generated Excel template, and specify the URL of the existing collection in the target SynBioHub instance.

## Preparation

It is assumed that the tool has been installed in ‘C:\SynBio2Easy’.

If you followed the instructions in the previous scenarios for generating and depositing a library of designs into SynBioHub, the collection URL should be something similar to ‘https://synbiohub.org/user/<user>/synbio2easy/synbio2easy\_collection/1.0’.

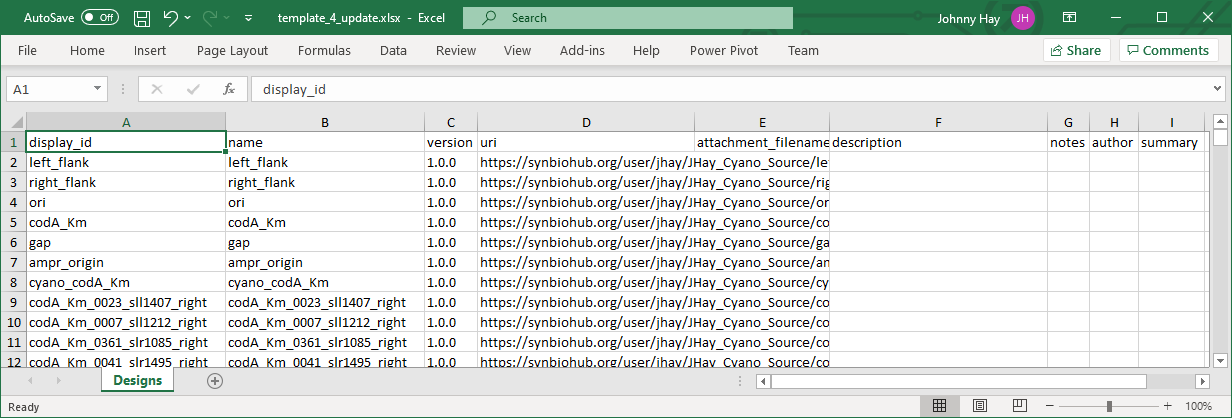
Instructions for Template4Update

1. Double-click on the start-up script “C:\SynBio2Easy\SynBio2Easy-CLI.bat”
   1. Alternatively, use the instructions in Appendix C to start the program manually
2. Choose ‘synbio2table’
3. Type in the name for the output Excel file, “library\template\_4\_update.xlsx” then <ENTER> when asked to specify the template file to be generated
4. Type in the URL of the existing collection in the target SynBioHub instance, then press <ENTER>
5. Type in your username, usually your email address, and press <ENTER>
6. Type in your password and press <ENTER>. **N.B.** the password input field is hidden so no text will appear as you type here

## Output

* The program will report which designs were generated in the template file
* The generated template Excel file will contain one worksheet comprised of nine columns: ‘display\_id’, ‘name, ‘version’, ‘uri’ and several optional metadata columns. The figure below shows an excerpt from a worksheet in an example template spreadsheet.

TODO: Insert screenshot of Excel template output file here



Scenario 7 – Update Existing Designs in SynBioHub Collection with File Attachments and Additional Metadata

In this scenario, the intention is to attach files containing supporting data, along with additional enriched metadata, to selected existing designs already deposited in SynBioHub.

We will use the Excel file which we generated in Scenario 6 to specify which files should be attached to which designs in SynBioHub; we will start the tool, point it to the excel file, instruct it where the designs are in SynBioHub (in which collection) and provide the connection details.

## Preparation

Create a folder named ‘metadata’ under ‘C:\SynBio2Easy’. If you generated the template for update as in Scenario 6, the metadata template spreadsheet will already be present in ‘C:\SynBio2Easy\metadata\template\_4\_update.xlsx’. Create a sub-folder named ‘attachments’ under ‘C:\SynBio2Easy\metadata’.

Copy the files you wish to attach to the designs in SynBioHub into ‘C:\SynBio2Easy\metadata\attachments’.

Edit the ‘template\_4\_update.xlsx’ file as you wish: it must contain the ‘display\_id’ column, but you can remove any of the other columns if they are not relevant to the metadata you wish to attach. The ‘attachment\_filename’ value is an absolute or relative path to a file for attachment to a design; in our case as we copied the files to the ‘attachments’ sub-directory, it will just be the relative path, e.g. “.\metadata\attachments\sll0199\_left.gb”. In, the ‘notes’ column you should explain what the file is, for example “sll0199\_left.gb is a PCR verified sequence.” HINT: you can use a formula to create the note from ‘attachment\_filename’ and some standard text, for example, assuming the attachment row is ‘B’, using a function such as ‘CONCATENATE($B$2, “ is a PCR file”)’.

Figure 1 below shows an example Excel work sheet. It is assumed that your Excel file is named ‘attachments.xlsx’ and is located in the same ‘sequences’ sub-directory.

TODO: Insert screenshot of Excel metadata input file here

## Instructions for Update

1. Double-click on the start-up script “C:\SynBio2Easy\SynBio2Easy-CLI.bat”
   1. Alternatively, use the instructions in Appendix C to start the program manually
2. Choose ‘update’
3. Enter the name of the MS Excel file containing the entity display IDs, file attachments and metadata you wish to upload: ‘metadata/attachments/template\_4\_update.xslx’
4. Enter the URL for the existing collection containing the designs you wish to update. Go to SynBioHub, open your collection, and copy the address from the web browser, for example ‘https://synbiohub.org/user/<USERNAME>/synbio2easy/synbio2easy\_collection/1.0’
5. Enter your SynBioHub username, usually your email address
6. Enter your SynBioHub password. **N.B.** the password input field is hidden so no text will appear as you type here

## Output

* The program will report which designs it found in the specified collection with display IDs matching those in the ‘display\_id’ column of the spreadsheet
* The program will also report which display IDs it could not find in the specified collection

TODO: Insert screenshot of updated design in SynBioHub UI here

