



D4.3 Annex I

Data Curation Tool

User Guide

DISCLAIMER: Any dissemination of results reflects only the author's view and the European Commission is not responsible for any use that may be made of the information it contains.

COPYRIGHT MESSAGE: © FAIR4Health Consortium, 2019

This report contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both. Reproduction is authorised provided the source is acknowledged.

Document Information

| GRANT AGREEMENT NUMBER 824666 | | ACRONYM: FAIR4Health | |
|-------------------------------|---|----------------------|-----------|
| Full title | Improving Health Research in EU through FAIR Data | | |
| Horizon 2020 Call | SwafS-04-2018: Encouraging the re-use of research data generated by publically funded research projects | | |
| Type of action | Research and Innovation Action | | |
| Start Date | 1 st December 2018 | Duration | 36 months |
| Website | www.fair4health.eu | | |
| Project Officer | Pepa Krasteva | | |
| Project Coordinator | Carlos Luis Parra Calderón, Andalusian Health Service | | |
| Report | Data Curation Tool User Guide | | |
| Related task | Task 4.1 – Development of the FAIR4Health agents and the FAIR4Health platform | | |
| Release date | November 2020 | | |
| Dissemination Level | Public | | |
| Responsible Author | Alper Teoman (P15 – SRDC) | | |
| e-mail | alper@srdc.com.tr | | |
| Collaborators | A. Anil Sinaci (P15 – SRDC), Mert Gencturk (P15 – SRDC) | | |
| Keywords | Data curation tool, FAIRification, user guide | | |

Table of Contents

| | |
|--|----|
| 1. Introduction | 3 |
| 2. Data Curation Tool Functional Features and User Manual..... | 3 |
| 2.1. Curation Steps Manual | 4 |
| 2.1.1. Verify FHIR Repository and Terminology Service | 4 |
| 2.1.1.1. Add FHIR Repository URL | 4 |
| 2.1.1.2. Add Terminology Service..... | 5 |
| 2.1.2. Add Data Source..... | 6 |
| 2.1.2.1. Browse Data Source (.xls, .xlsx or .csv)..... | 6 |
| 2.1.2.2. Import from Saved Mapping | 7 |
| 2.1.3. Map Metadata | 8 |
| 2.1.3.1. Display of FHIR Resource and Profile Elements..... | 9 |
| 2.1.3.2. Display of Source Data Columns (Contents) | 10 |
| 2.1.3.3. Applicable Filters and Reloading the Data Source..... | 11 |
| 2.1.3.4. Mapping of FHIR Elements and Source Data Elements..... | 12 |
| 2.1.3.5. Assign a Default Value to a FHIR Element..... | 14 |
| 2.1.3.6. Add Concept Map (Selecting Source and Target Systems) | 15 |
| 2.1.3.7. Edit & Remove the Mapping..... | 17 |
| 2.1.3.8. Save & Export the Mapping..... | 18 |
| 2.1.4. Validate Created FHIR Resources..... | 20 |
| 2.1.4.1. Create Resources and Validate..... | 20 |
| 2.1.4.2. View Validation Details | 21 |
| 2.1.4.3. Export Created Resources | 22 |
| 2.1.5. Transform | 23 |
| 2.1.5.1. Put Resources into the FHIR Repository | 23 |
| 2.2. Tool Settings..... | 24 |
| 2.2.1. Change Language (Localization) | 24 |
| 2.2.2. Toggle Sidebar & Full Screen | 25 |
| 2.2.3. Debugging | 26 |
| 2.2.4. Log Locations..... | 27 |
| 2.2.5. Help..... | 27 |

List of Figures

| | |
|---|----|
| Figure 1. Home Screen..... | 3 |
| Figure 2. FHIR Repository Verification..... | 5 |
| Figure 3. Adding Terminology Service..... | 5 |
| Figure 4. Completed Verification Step..... | 6 |
| Figure 5. Browse File Options..... | 7 |
| Figure 6. Browse File - Select Files | 7 |
| Figure 7. Import from a Saved Mapping | 8 |
| Figure 8. Map Metadata FHIR and Data Source Parts | 9 |
| Figure 9. Select FHIR Resource..... | 9 |
| Figure 10. Select FHIR Profile | 10 |
| Figure 11. Select Data Source File | 10 |
| Figure 12. Select Data Source Sheet / Table | 11 |
| Figure 13. Search and Filter the Elements & Reloading the Data Source | 12 |
| Figure 14. Patient Resource Gender Field Mapping | 13 |
| Figure 15. Display of Complex Data Types | 13 |
| Figure 16. Patient Resource Country Field Mapping (Complex Data Type) | 14 |
| Figure 17. Assign Default Value - 1..... | 14 |
| Figure 18. Assign Default Value - 2..... | 15 |
| Figure 19. Add Concept Mapping to an Element | 16 |
| Figure 20. Select a Source and Target System for the Translation..... | 16 |
| Figure 21. Concept Map Error - No terminology service is connected..... | 17 |
| Figure 22. Actions on the Created Mapping | 18 |
| Figure 23. Mapping Edit Mode | 18 |
| Figure 24. Save & Export the Mapping..... | 19 |
| Figure 25. Save Mapping in the Project Local Directory | 19 |
| Figure 26. Export Mapping..... | 20 |
| Figure 27. Validation..... | 21 |
| Figure 28. Validation Result Details - 1 | 22 |
| Figure 29. Validation Result Details - 2..... | 22 |
| Figure 30. Transform Screen..... | 23 |
| Figure 31. Transform Result..... | 24 |
| Figure 32. Tool Language Settings | 25 |
| Figure 33. Tool Window View Options | 26 |
| Figure 34. Open Sidebar..... | 26 |
| Figure 35. Open Developer Tools..... | 27 |
| Figure 36. Help - GitHub Repository..... | 28 |

1. Introduction

Data Curation Tool, described in this manual, is a standalone, desktop application which can run on any operating system (Windows, macOS, Linux). The tool is developed by the FAIR4Health project (<https://www.fair4health.eu/>). It is an open source project and you can access it from <https://github.com/fair4health/data-curation-tool>. You can follow the setup instructions in the [repository](#) and create an executable.

2. Data Curation Tool Functional Features and User Manual

This manual consists of two sections: Data curation guide to describe the steps of curating existing data into an HL7 FHIR Repository and tool-wide settings such as language and window settings. Before you start using the application, you can browse the section summarizing the curation steps on the screen that opens when you run the application (Figure 1).

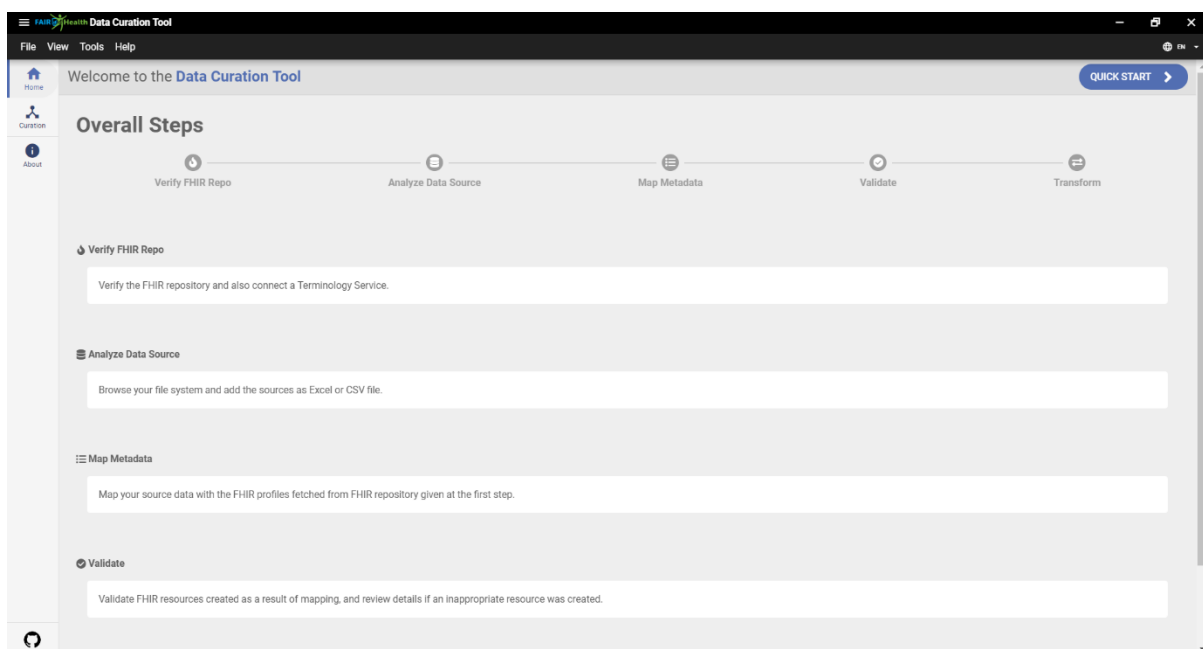


Figure 1. Home Screen

2.1. Curation Steps Manual

In this section, the steps of data curation using the tool is described with the help of screenshots. The data curation process consists of 5 steps:

- ❖ Verify FHIR Repository
- ❖ Analyze Data Source
- ❖ Map Metadata
- ❖ Validate
- ❖ Confirm & Transform

2.1.1. Verify FHIR Repository and Terminology Service

2.1.1.1. Add FHIR Repository URL

This tool transforms your existing data (such as from an Excel file) into FHIR resources and puts them into an HL7 FHIR Repository. Hence, the very first step is to verify that the tool can access a FHIR Repository. The tool asks from the user to provide a URL, as shown in Figure 2.

We utilize <https://onfhir.io> as the FHIR Repository of the FAIRification workflow. Before you continue with the curation, it is necessary that the tool can connect to a valid FHIR repository, hence you verify it. You can download and run the onFHIR repository.

1 enter the FHIR repository **base URL** and click the verify button.

As an optional setting, you can provide a terminology service if your source data includes terminology systems (coding schemes) that need be mapped into the terminology systems accepted by the HL7 FHIR Repository that you are connected. With button **2**, you can connect a new terminology service.

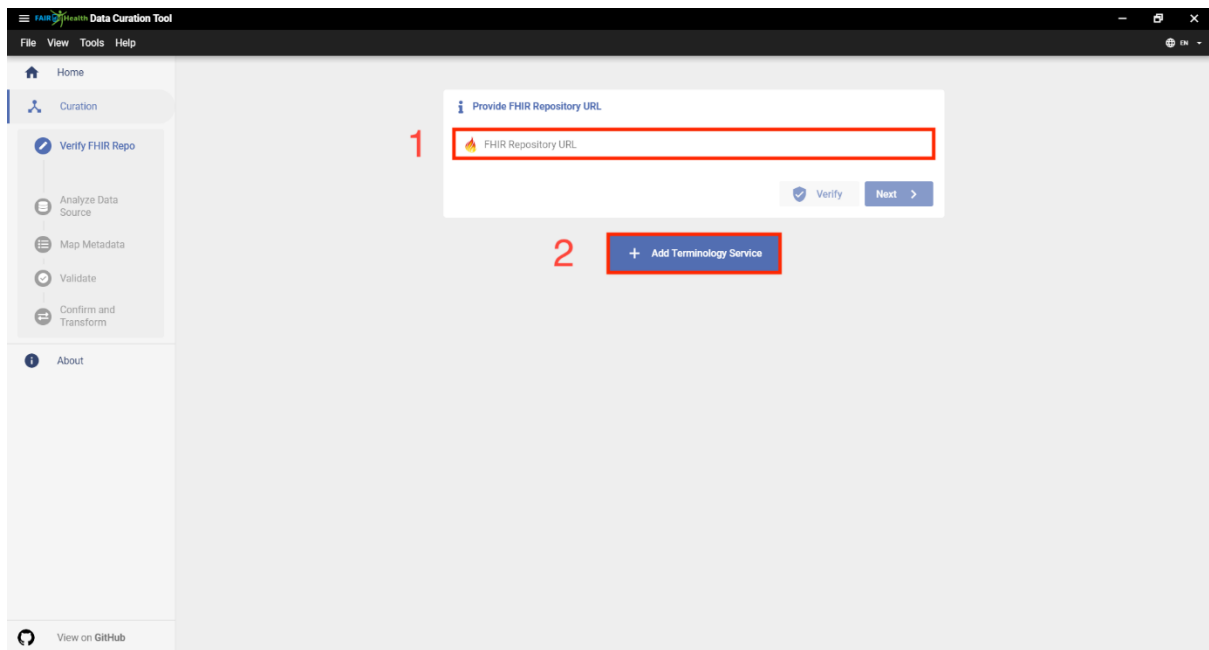


Figure 2. FHIR Repository Verification

2.1.1.2. Add Terminology Service

Enter the terminology service **base URL** in the framed section shown in the Figure 3 and click the verify button. In the Figure 4, you can see the status of both URLs verified.

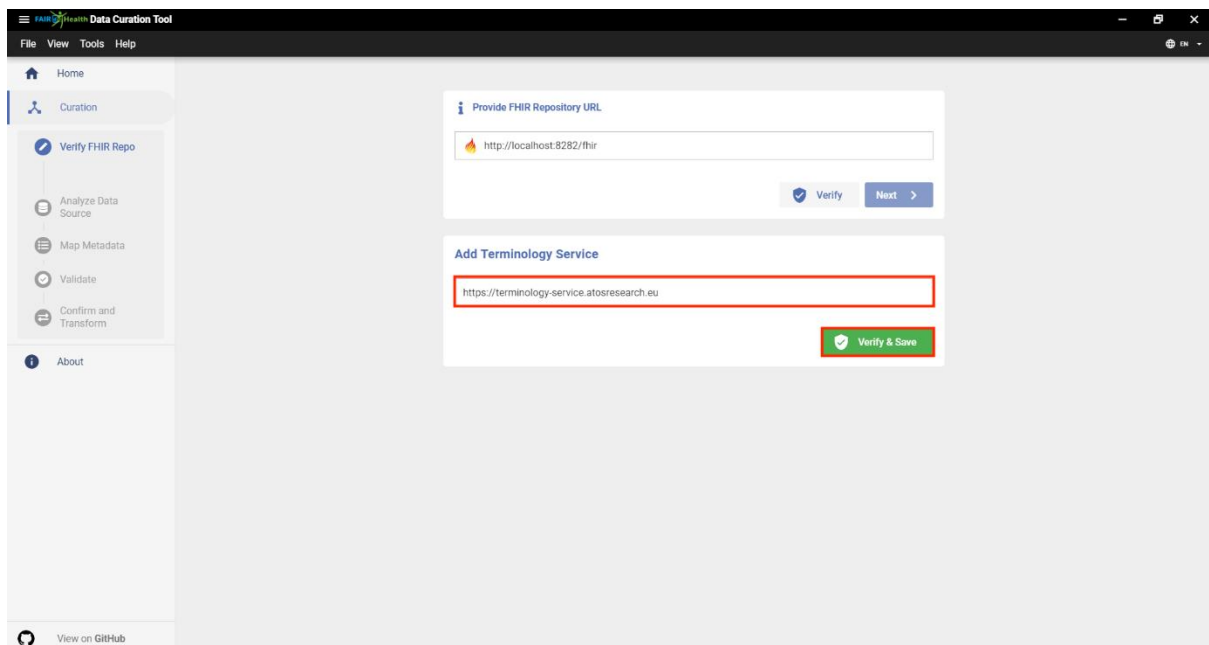


Figure 3. Adding Terminology Service

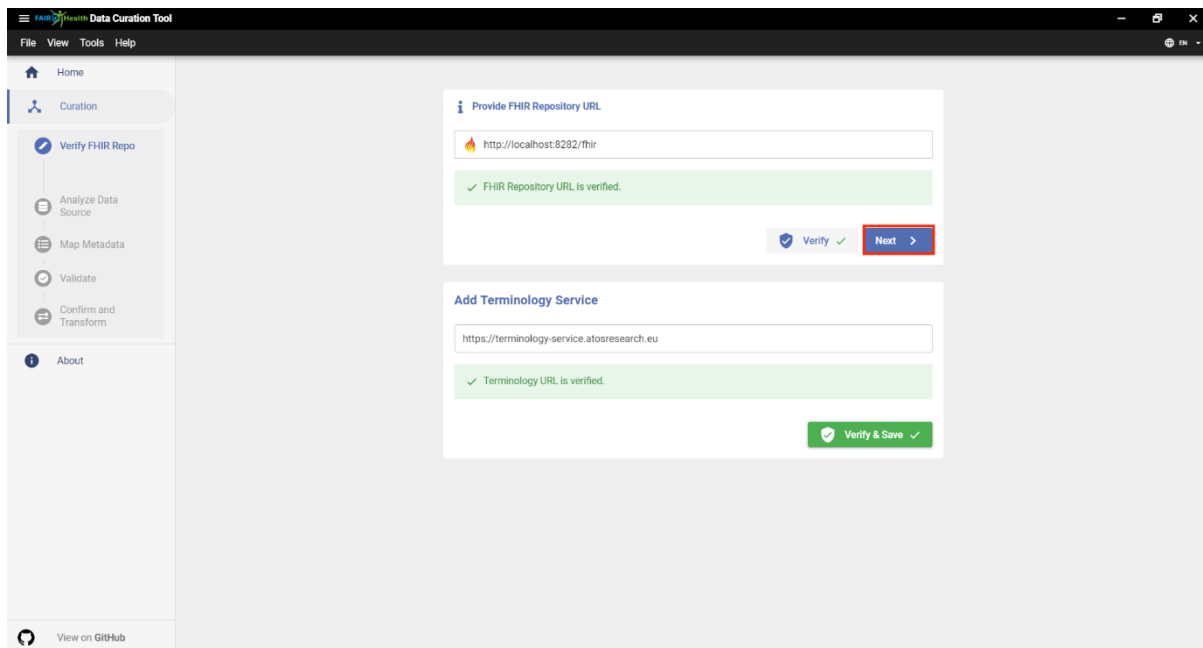


Figure 4. Completed Verification Step

2.1.2. Add Data Source

2.1.2.1. Browse Data Source (.xls, .xlsx or .csv)

With drag and drop method (Figure 5), you can drop the excel and csv data files into the box in **1**. Or you can select and add files by clicking the **2** Browse button. The Browse button will open a window in your file manager and will ask you to select files in .xls .xlsx and .csv extensions.

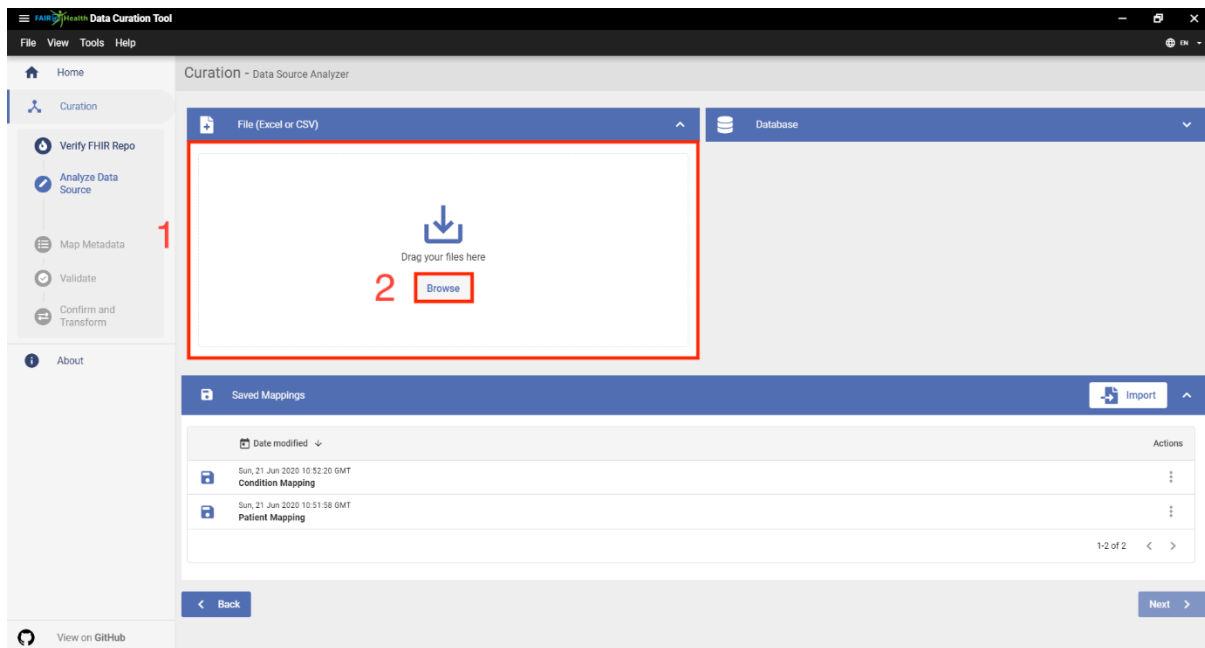


Figure 5. Browse File Options

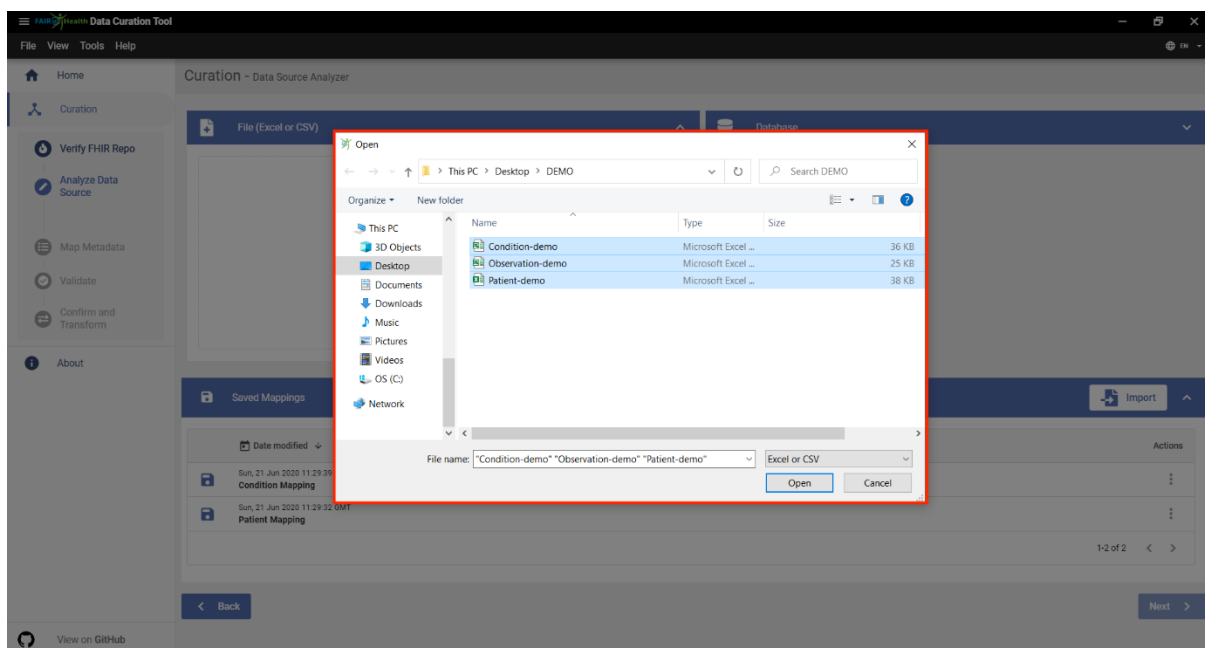


Figure 6. Browse File - Select Files

2.1.2.2. Import from Saved Mapping

If you want to import a mapping you have already made, you can do this from the Saved Mappings section at this step (Figure 7). From the button in **1**, you can add the mappings you have exported. Or you can choose from the mappings you have saved in the tool in section **2**. Saving and exporting a mapping is mentioned under the section **2.1.3.8**.

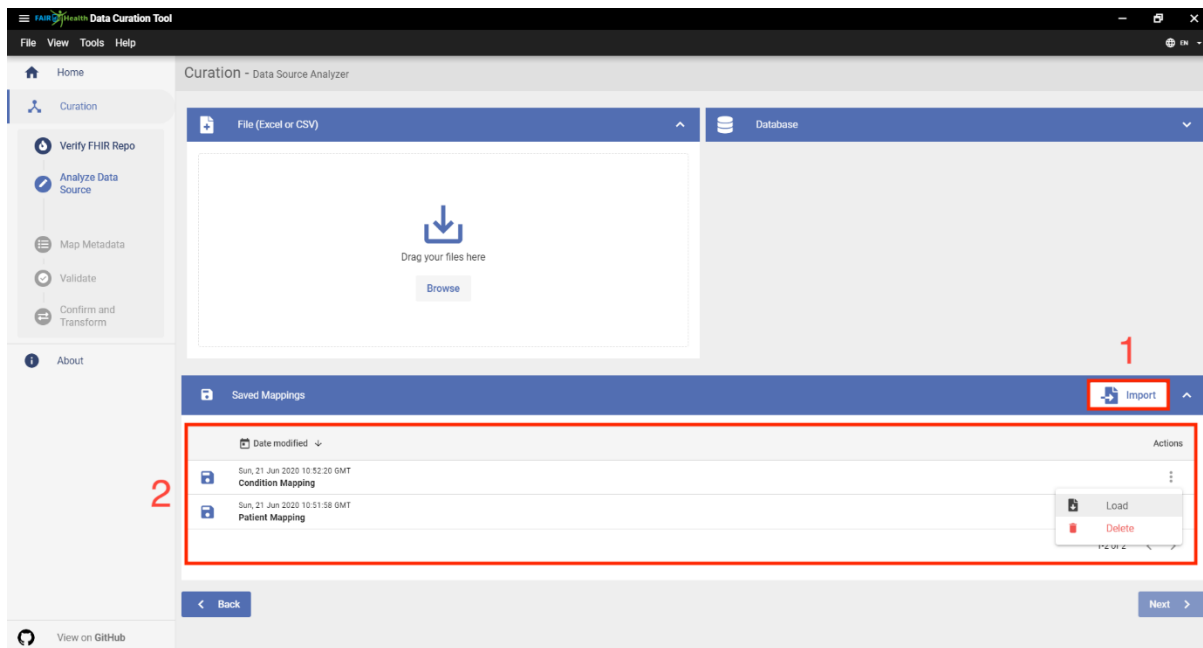


Figure 7. Import from a Saved Mapping

2.1.3. Map Metadata

FHIR Repository has been connected, source files have been added, now it is time to map metadata (map your source files to the FHIR resources of FAIR4Health common-data-model). At this step (Figure 8), the screen will be split in two. On the left **1**, you will see the FHIR resources, profiles, and their elements in the FHIR Repo you have connected. On the right **2**, you will see the table/sheet components inside your data source.

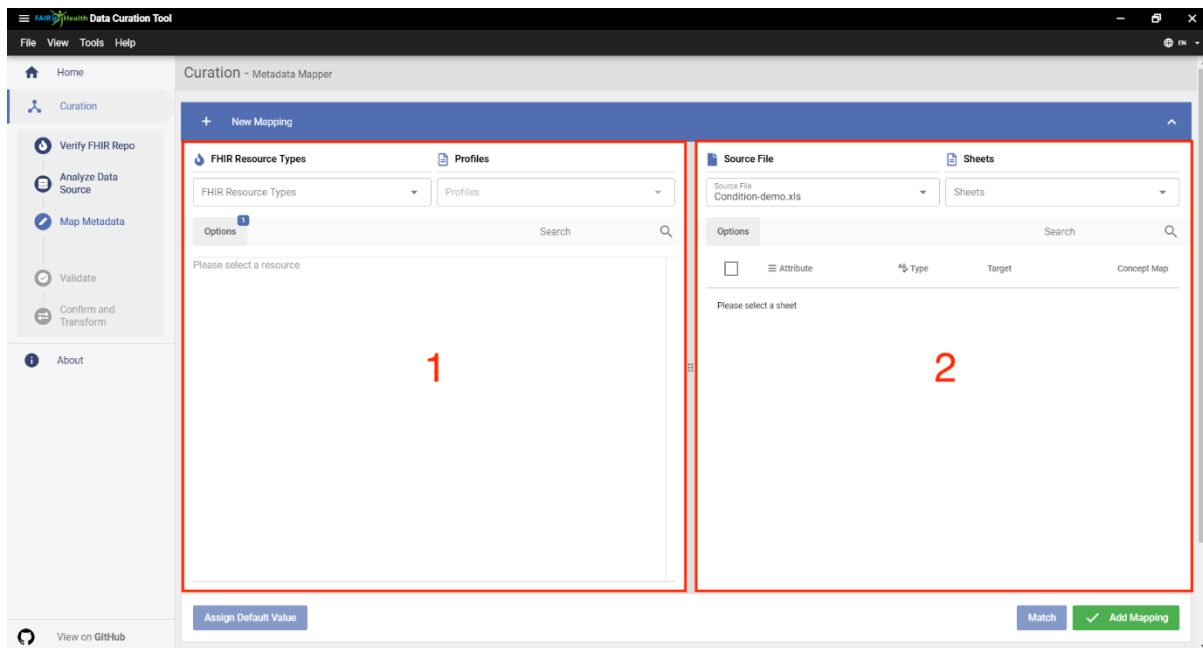


Figure 8. Map Metadata FHIR and Data Source Parts

2.1.3.1. Display of FHIR Resource and Profile Elements

The dropdown lists in figures Figure 9 and Figure 10 shows the resources types fetched from the FHIR repo (according to its CapabilityStatement) and the profiles defined under these resource types. After selecting a profile, you can see the elements of that resource types that you need to map data to.

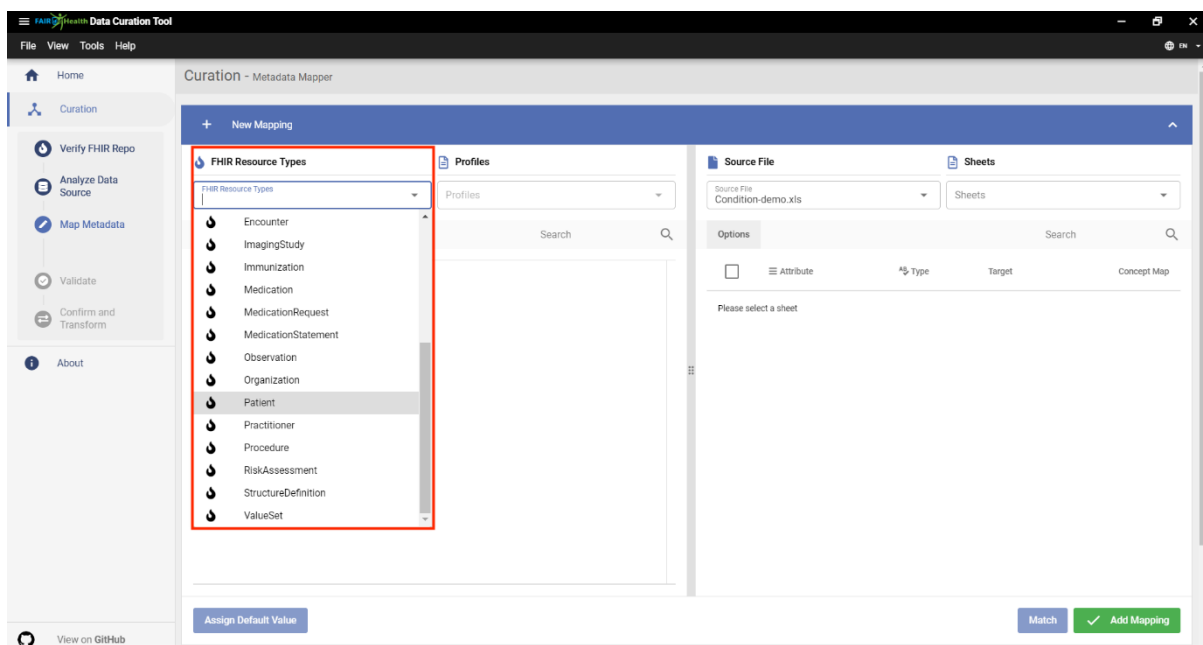


Figure 9. Select FHIR Resource

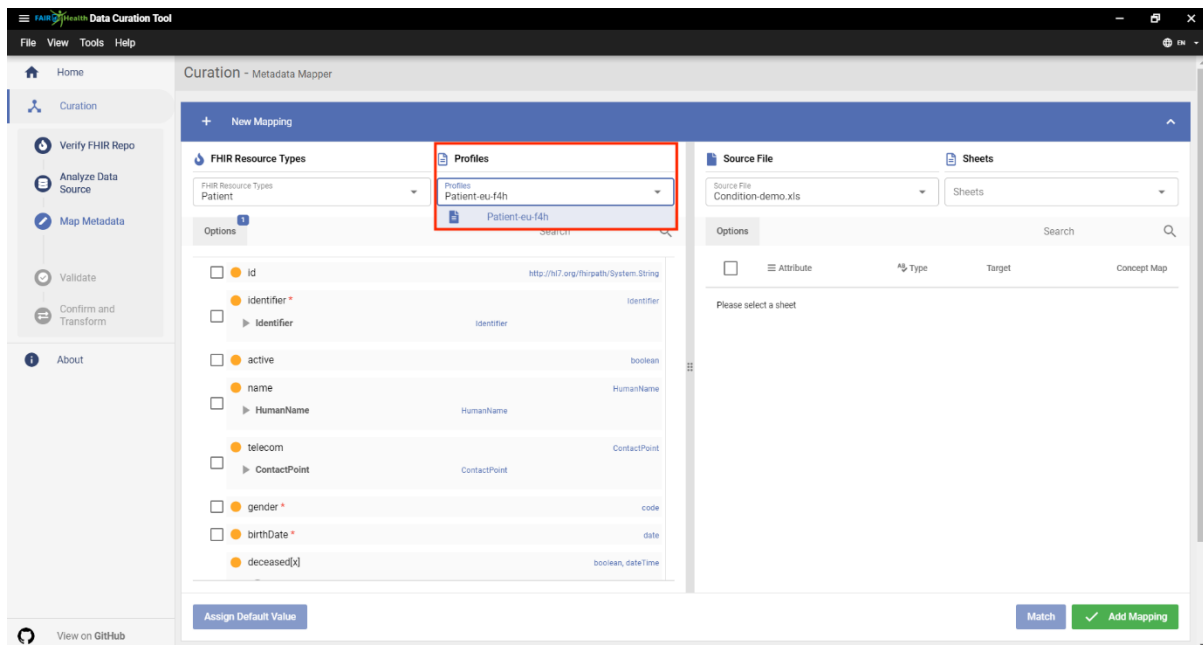


Figure 10. Select FHIR Profile

2.1.3.2. Display of Source Data Columns (Contents)

Similar to the FHIR resource type selection process, choose the source file and sheet/table from the dropdown lists (Figure 11 and Figure 12).

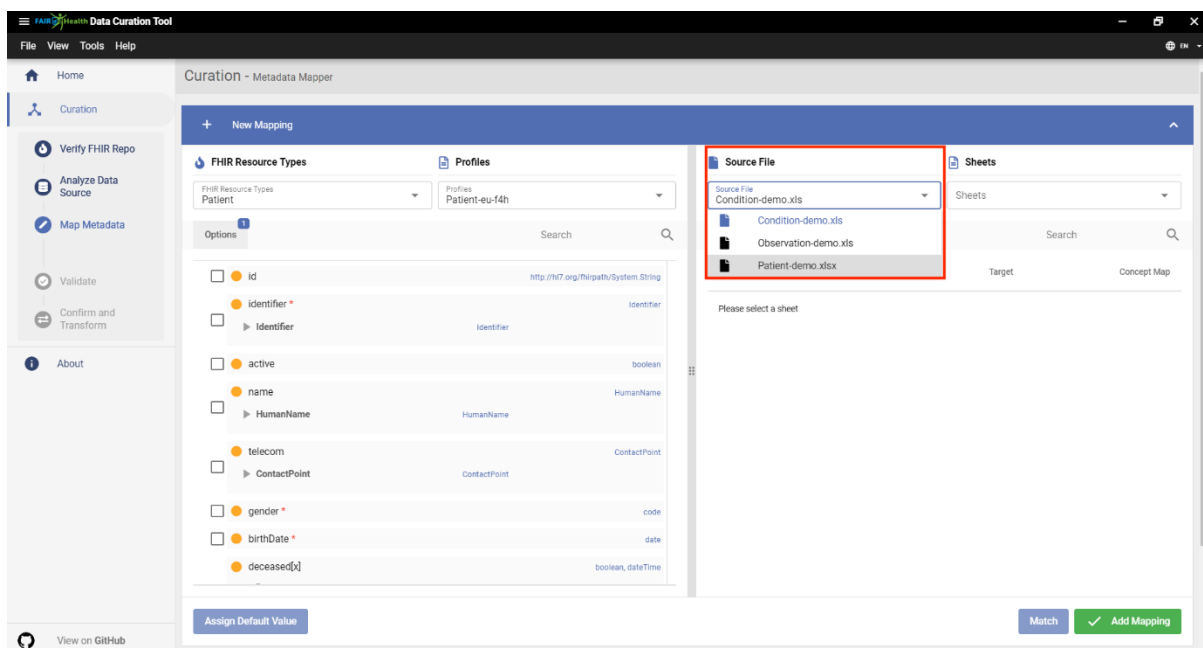


Figure 11. Select Data Source File

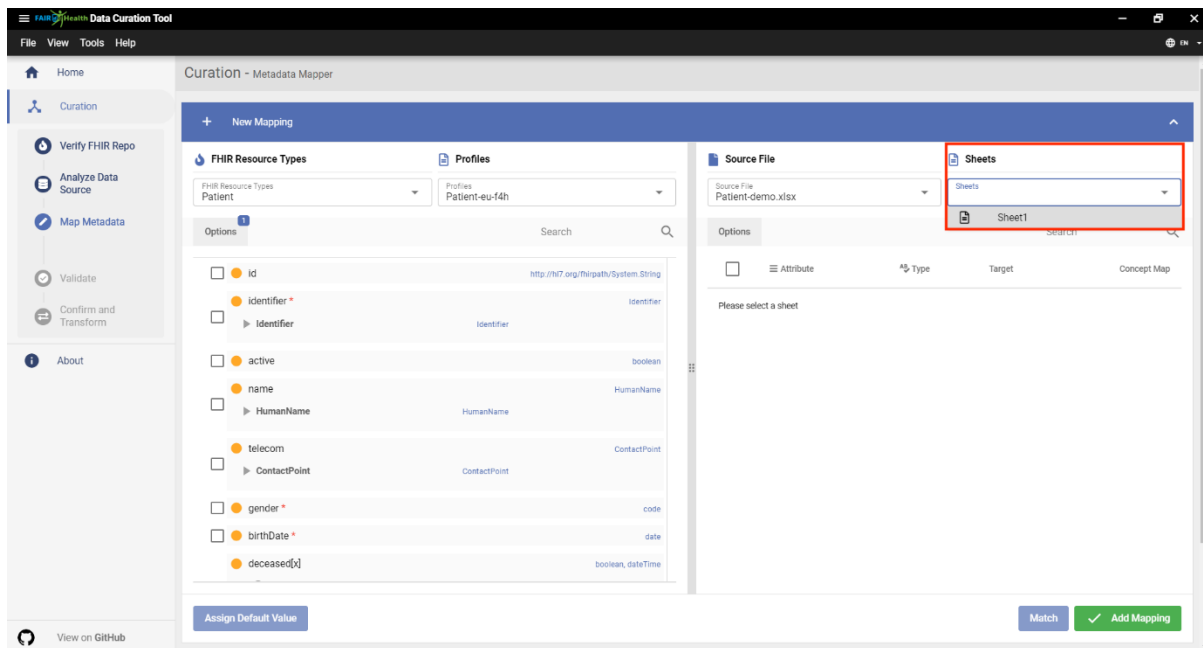


Figure 12. Select Data Source Sheet / Table

2.1.3.3. Applicable Filters and Reloading the Data Source

In Figure 13, filter options and search bars can be seen.

As shown in field **1**, two filters can be applied to the FHIR elements.

- ❖ **Show mandatory elements only:** Highlights the fields that must be provided within the selected FHIR profile.
- ❖ **Hide base resource elements:** Hides the elements coming from the base resource types in FHIR. These are Resource and DomainResource elements. You can find it from <https://www.hl7.org/fhir/resource.html> and <https://www.hl7.org/fhir/domainresource.html>

In fields **2**, you can search for the names of the elements.

With the number **3** Reload File button, you can reload your data from the file (it can be used if there is a change in the file) and clear the mappings made before saving.

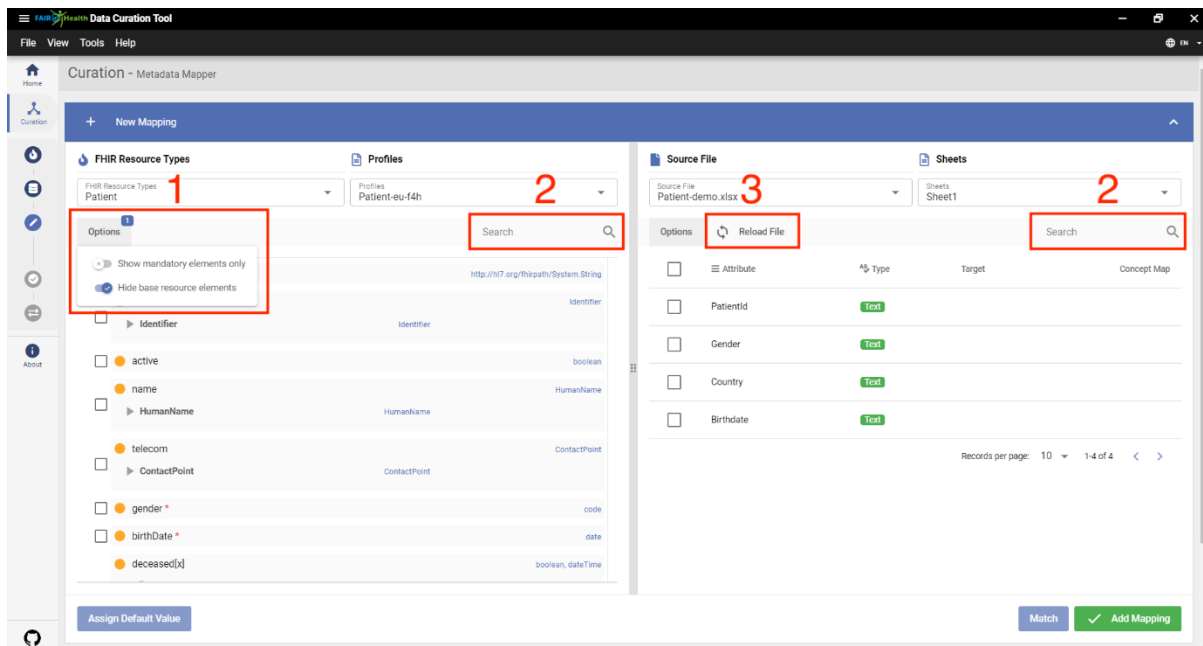


Figure 13. Search and Filter the Elements & Reloading the Data Source

2.1.3.4. Mapping of FHIR Elements and Source Data Elements

To map a field in the source data to an FHIR element, regardless of order, select the FHIR element and the source field to be mapped, then click the Match button. It means that you want to fill the FHIR element of that resource with the values coming from the source field of your file. As detailed in Figure 14, select the target FHIR element from field 1 and specify your source in field 2 and complete the matching. Once you match your source data fields with the FHIR elements you can finish the mapping and add it to your mapping configuration. There are different scenarios you may encounter during this matching process. These are:

1. If the FHIR resource element is a simple data type: text, id, code, boolean
2. If the FHIR resource element is a complex data type: CodeableConcept, HumanName, Address

The mapping in the first scenario is performed by selecting the field directly, as shown in Figure 14.

For the situation in the second scenario, the fields of the complex data type are expected to be expanded and selected from inside as shown in Figure 16. Please be sure to select an element inside the complex data when mapping the elements in the complex data structure. If you do not, the system will give a warning message.

As shown in Figure 15, you can expand your complex data type, select the field, and perform your mapping.

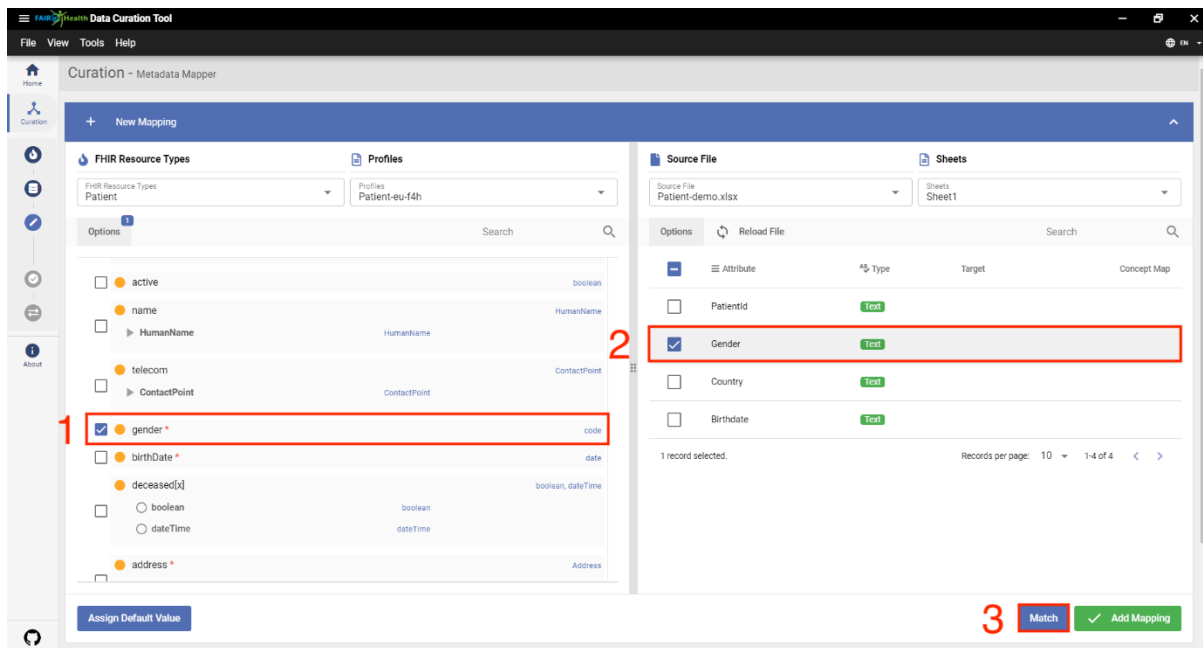


Figure 14. Patient Resource Gender Field Mapping

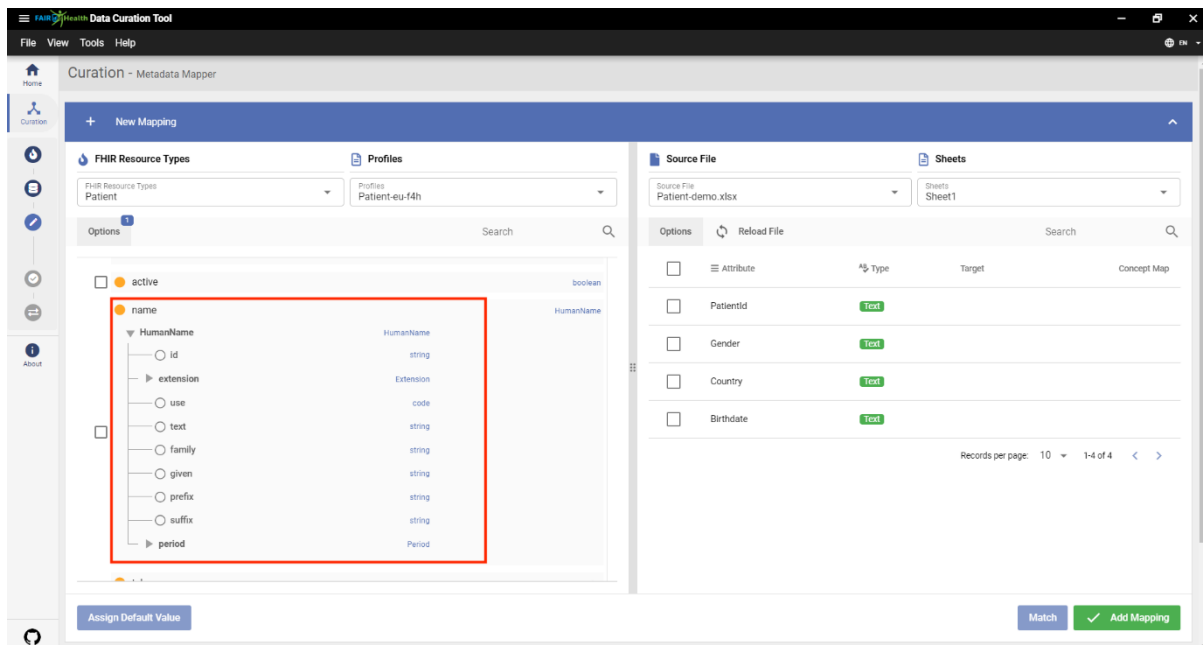


Figure 15. Display of Complex Data Types

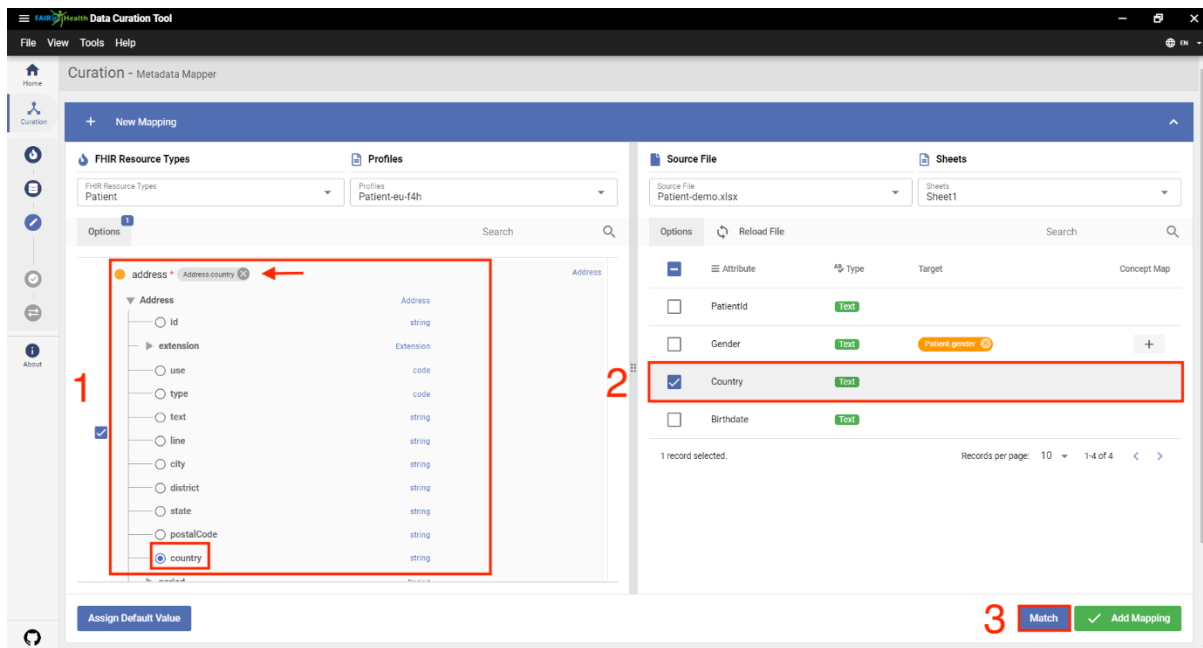


Figure 16. Patient Resource Country Field Mapping (Complex Data Type)

2.1.3.5. Assign a Default Value to a FHIR Element

If you want to add something that is not in your source data to an FHIR element, that is, if you want to assign a default value, follow the actions you see in Figure 17. First select the FHIR element you want to assign a value to (for example, see box **1**), then click the button number **2**. In the popup box in Figure 18, enter the desired value and click OK.

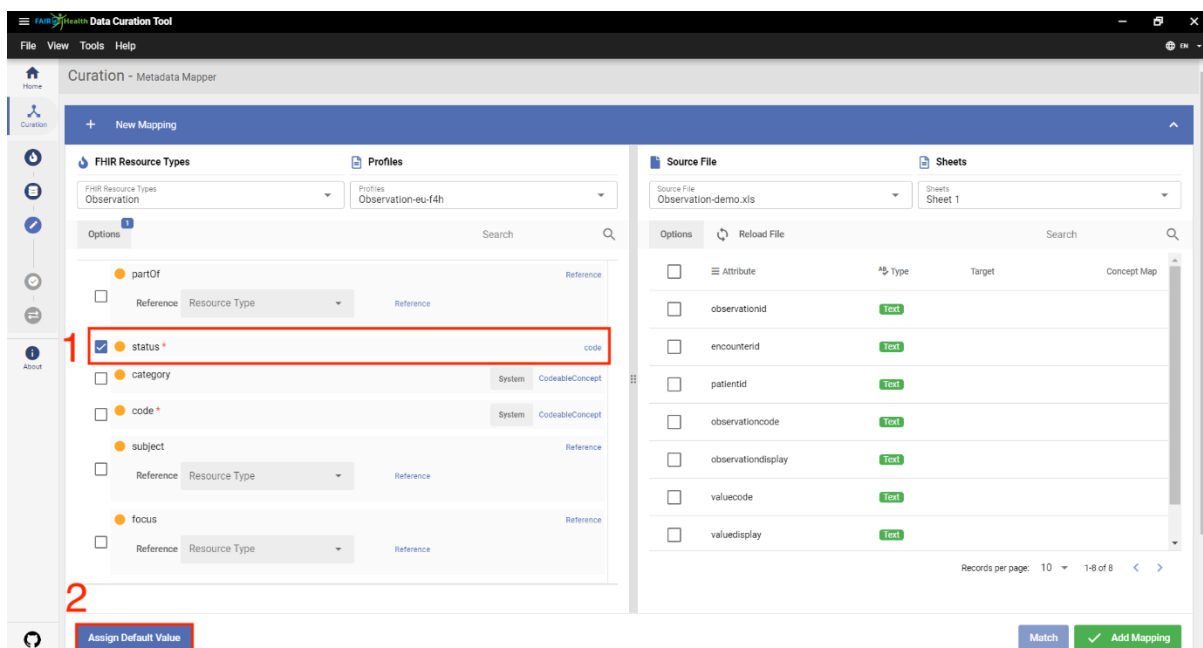


Figure 17. Assign Default Value – 1

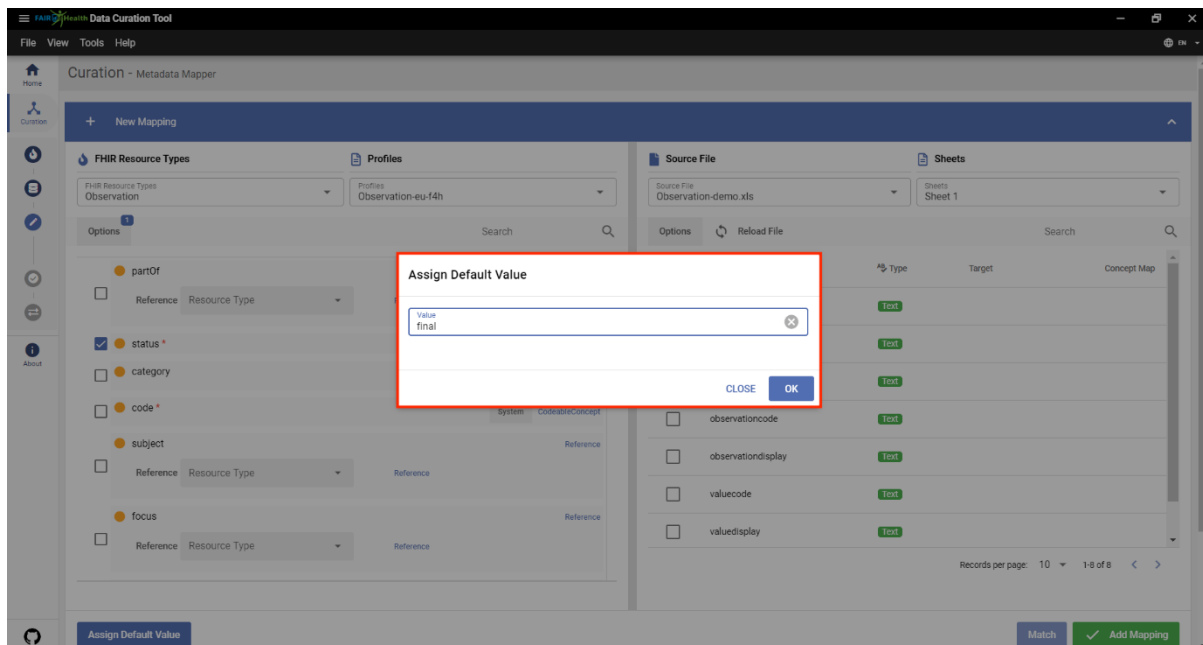


Figure 18. Assign Default Value - 2

2.1.3.6. Add Concept Map (Selecting Source and Target Systems)

If you want to translate your data, for example for the gender field, you are holding the genders as "m" and "f" and they should be kept as "male" and "female" in FHIR. You will choose source and target systems to translate those values into some other concepts, which translates m into male and f into female (or to some standard terminology systems such as SNOMED-CT). You can do this by clicking the plus button next to the relevant field on the source data side (Figure 19). Fill in the source and target fields, which is also illustrated in Figure 20.

If you have not added Terminology service in the first step, be aware that you will see a warning message as shown in Figure 21.

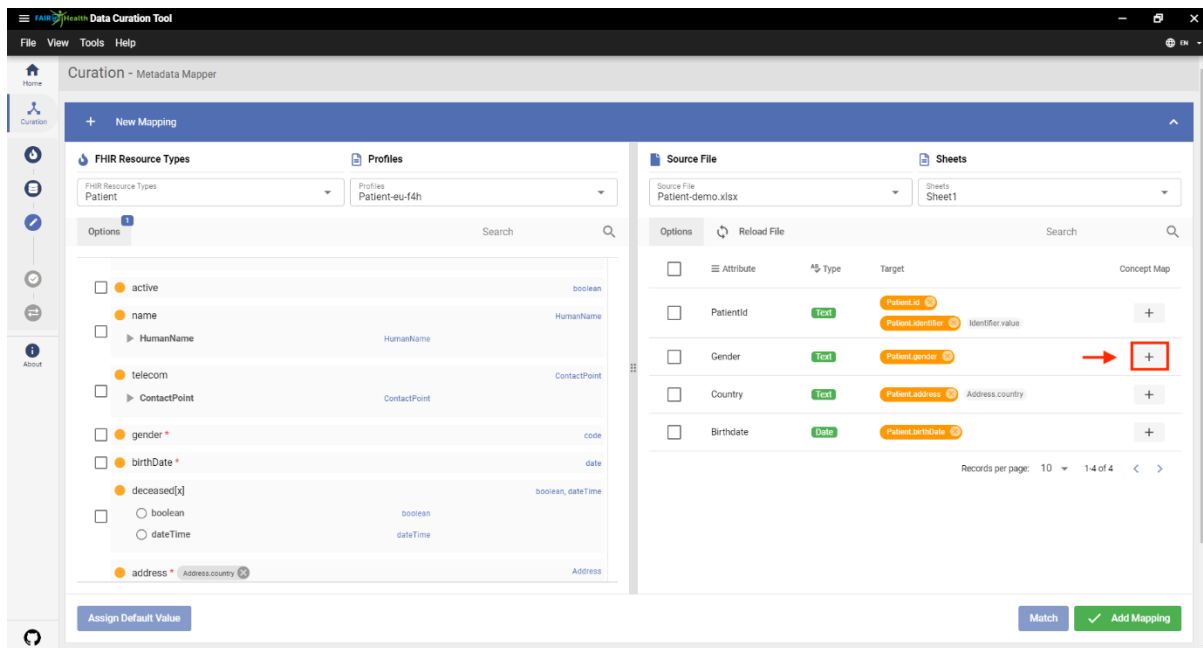


Figure 19. Add Concept Mapping to an Element

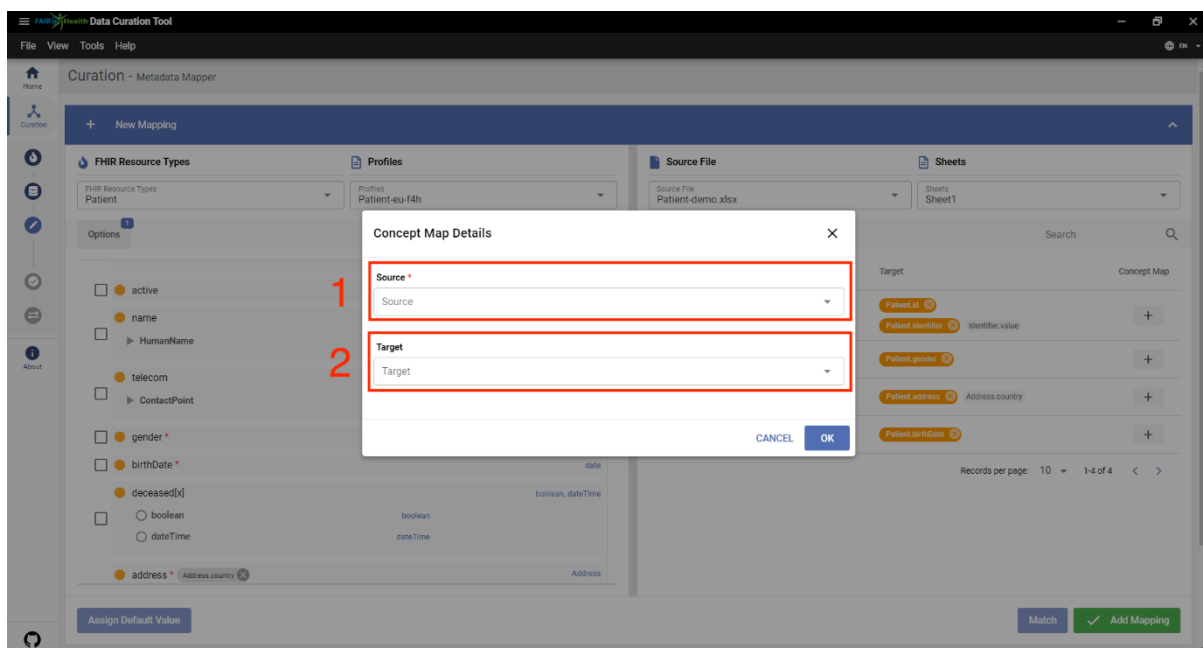


Figure 20. Select a Source and Target System for the Translation

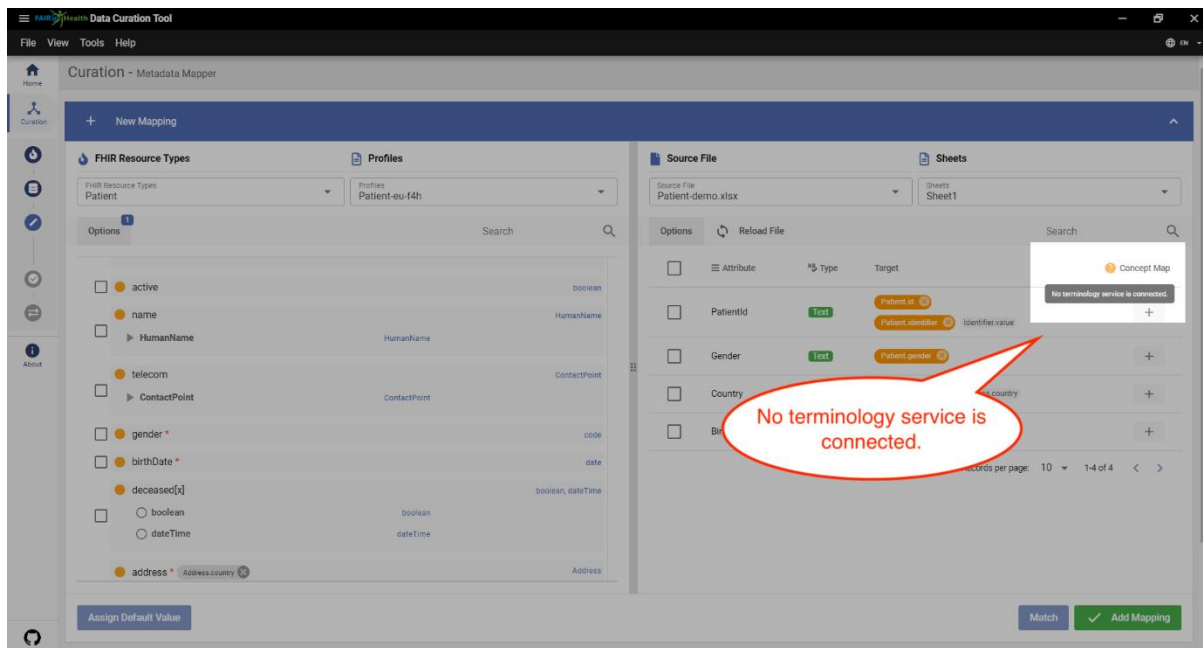


Figure 21. Concept Map Error - No terminology service is connected

2.1.3.7. Edit & Remove the Mapping

After you finish a mapping and click the Add Mapping button, scroll down and you will see your mapping in this box as depicted in Figure 22. Here you can see the main components of the mappings you've made (which profile and to which fields you created matchings Figure 22).

With the button **1**, you can see this mapping in more detail.

With the button **2**, you can edit your match.

And with the button **3**, you can delete this mapping.

Sample Edit mode of a mapping is illustrated in Figure 23. In this section, you can apply any changes to mapping.

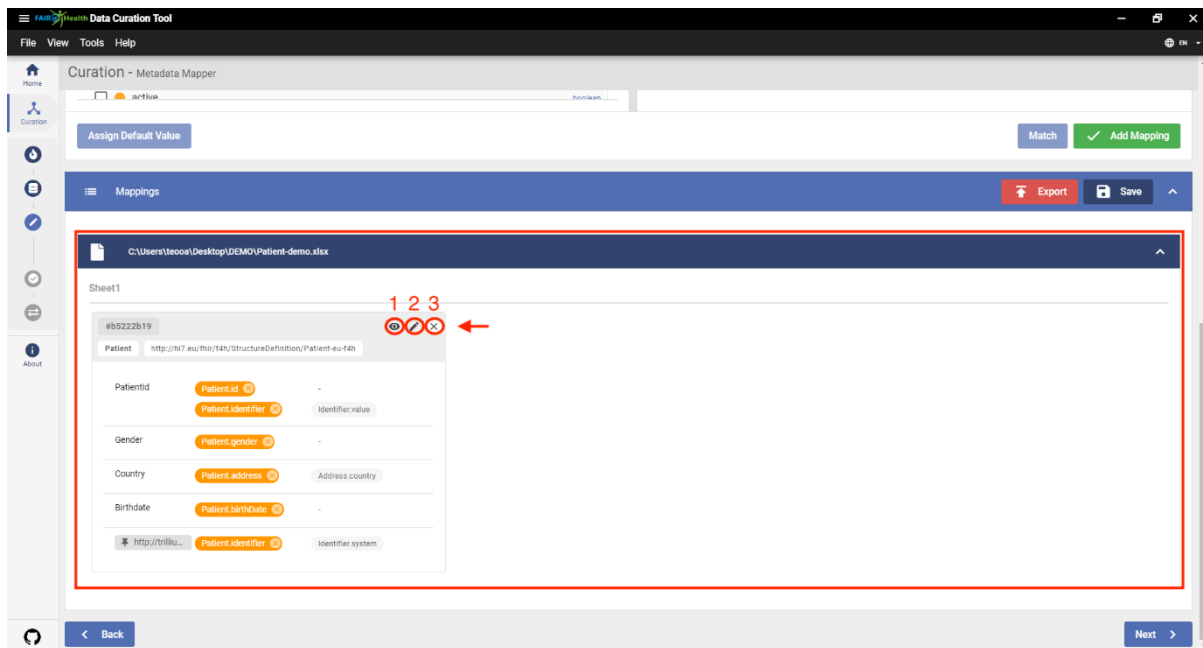


Figure 22. Actions on the Created Mapping

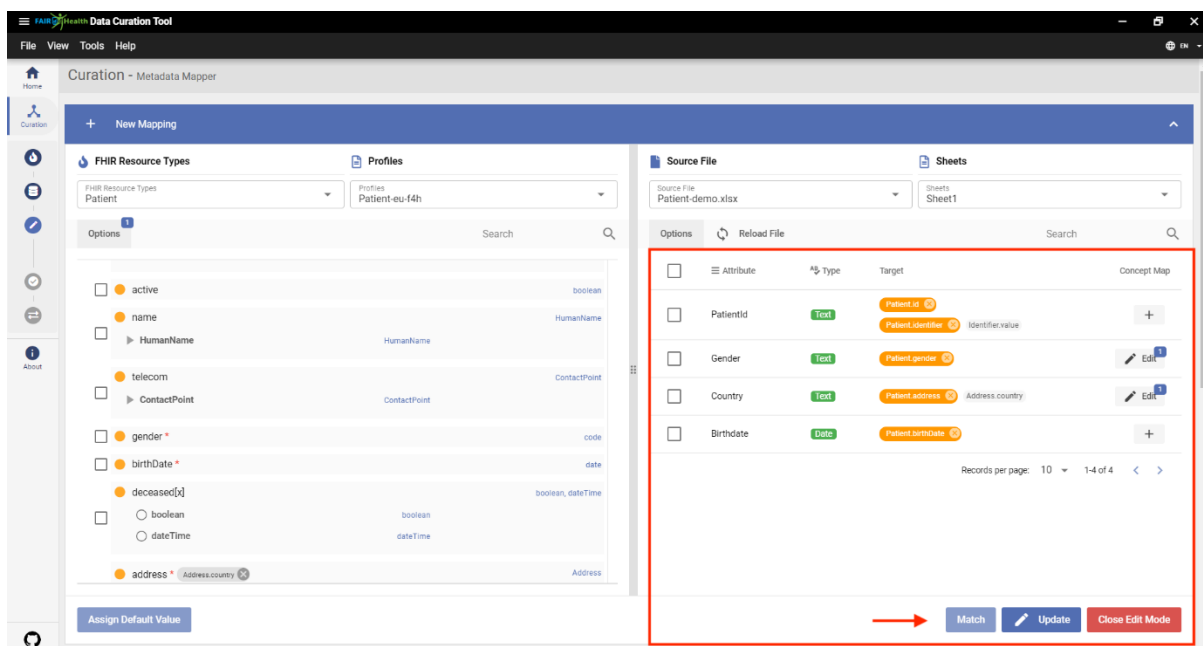


Figure 23. Mapping Edit Mode

2.1.3.8. Save & Export the Mapping

In this section, you will be able to save or export the mappings you made.

As shown in Figure 24, in the right side **Save** button, you can save your mapping into the application, which means into the application local directory (Figure 25).

Or by clicking the left side **Export** button, you can export mapping configuration as a JSON file (Figure 26).

And after this step, as mentioned in section **2.1.2.2**, you can import your saved mappings.

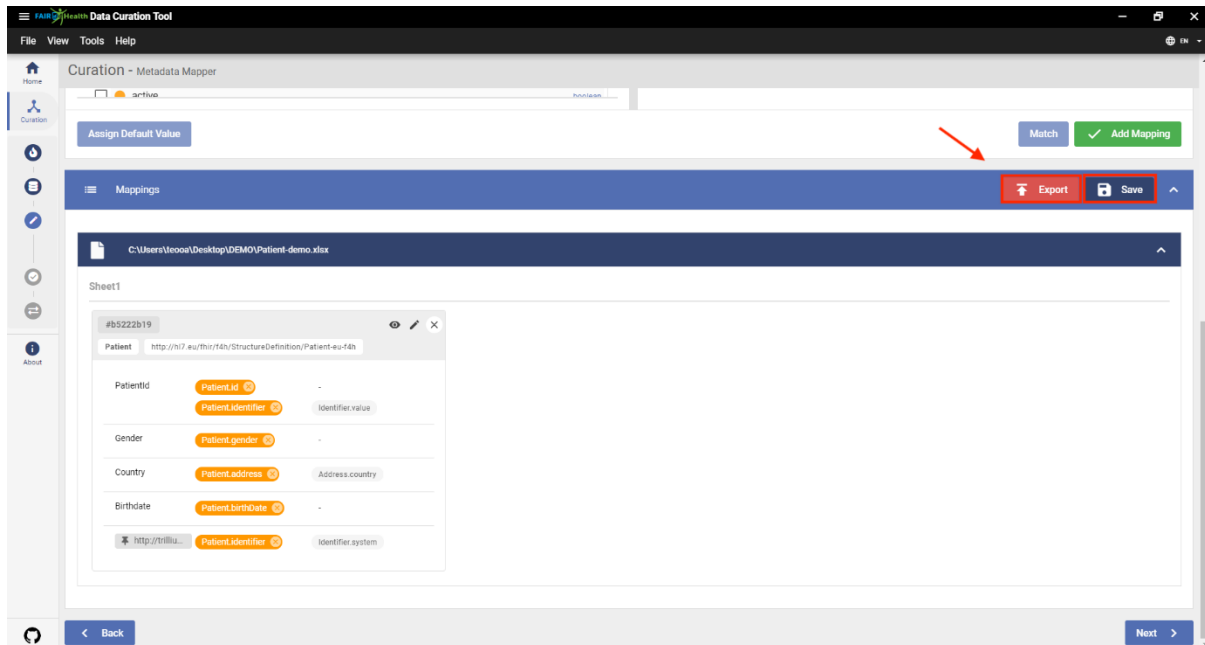


Figure 24. Save & Export the Mapping

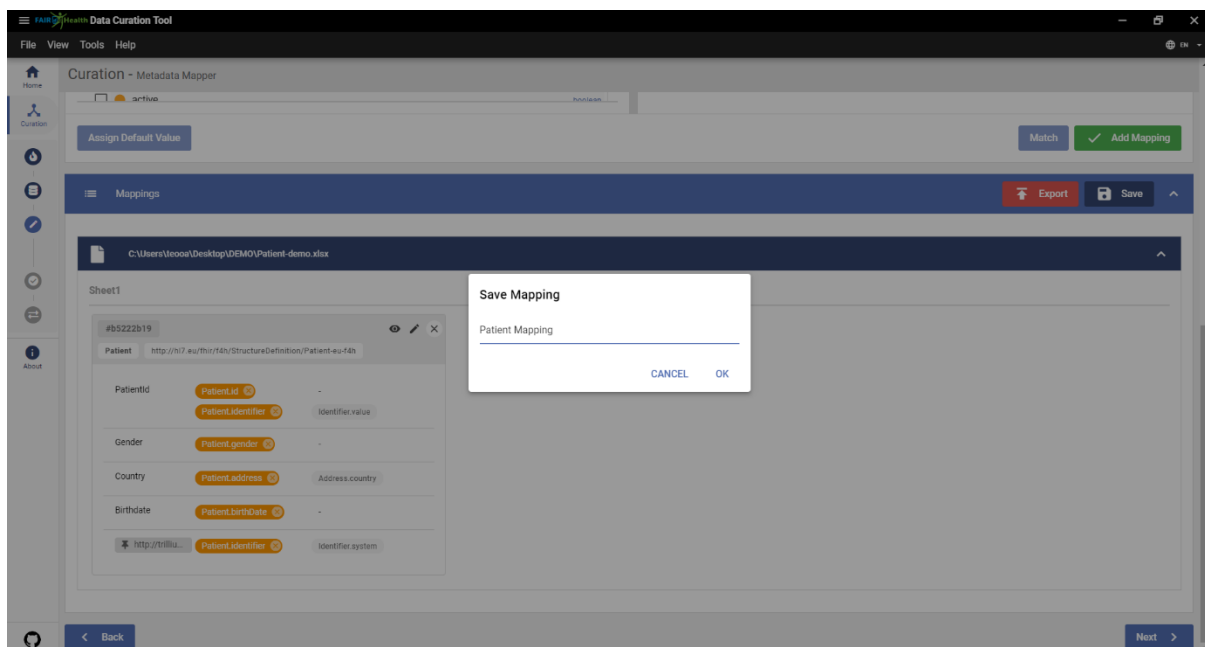


Figure 25. Save Mapping in the Project Local Directory

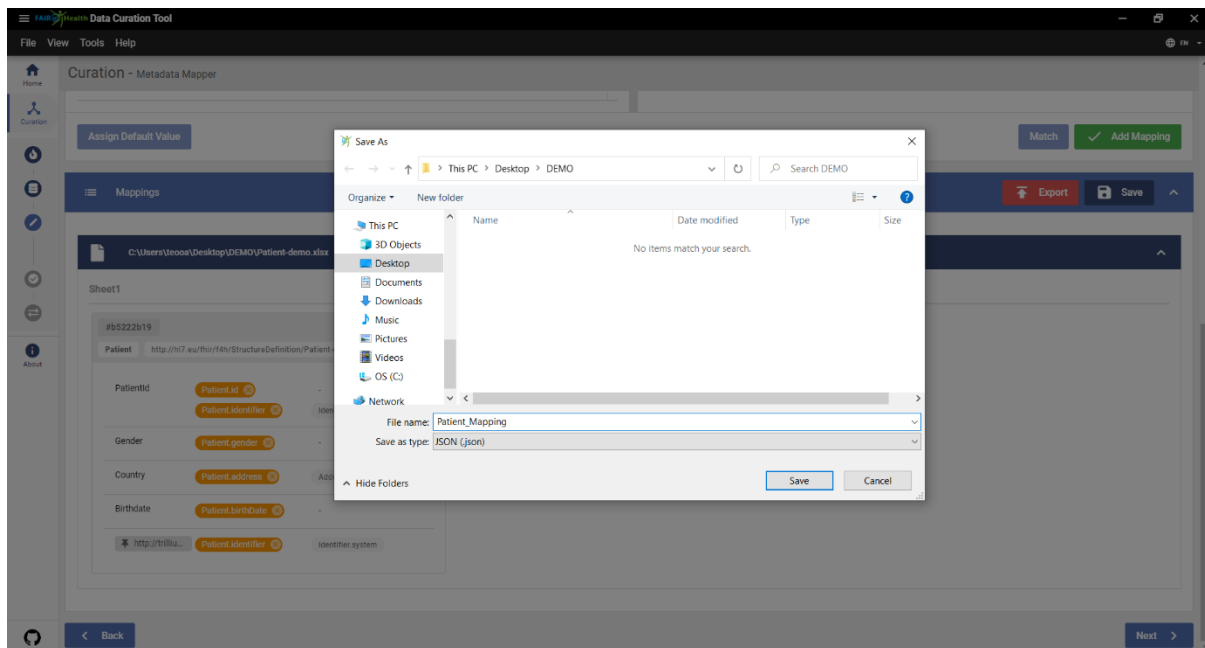


Figure 26. Export Mapping

2.1.4. Validate Created FHIR Resources

2.1.4.1. Create Resources and Validate

At this step, the resource instances are created to validate the mappings made and their suitability to the profiles is performed through the validate operation. Figure 27 shows the validation screen.

You can select the files you want to validate from inside the box number **1**. By default, all are selected.

Finally, you can start the validation process from button **2**.

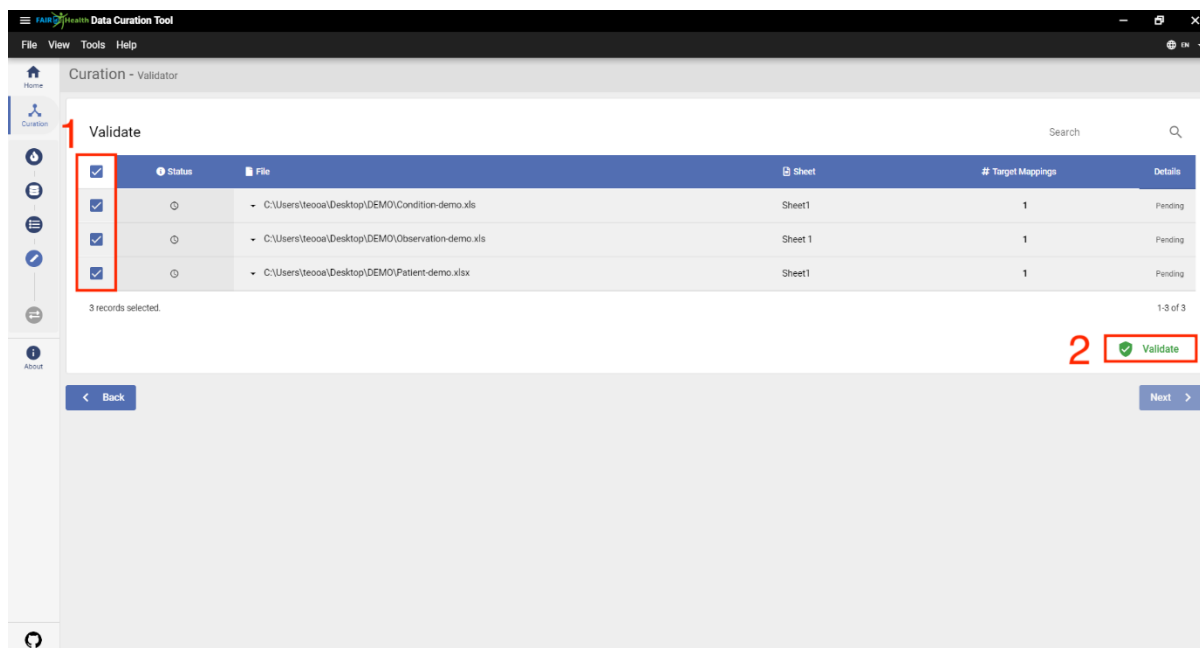


Figure 27. Validation

2.1.4.2. View Validation Details

Once validation is complete (Figure 28), you will be able to see the details file by file. Click the number **1** Details button to see the details of the validation result occurring in that file. As seen in Figure 29, a popup window will appear and you will see the resource types created in that file, and the validation result success-error numbers.

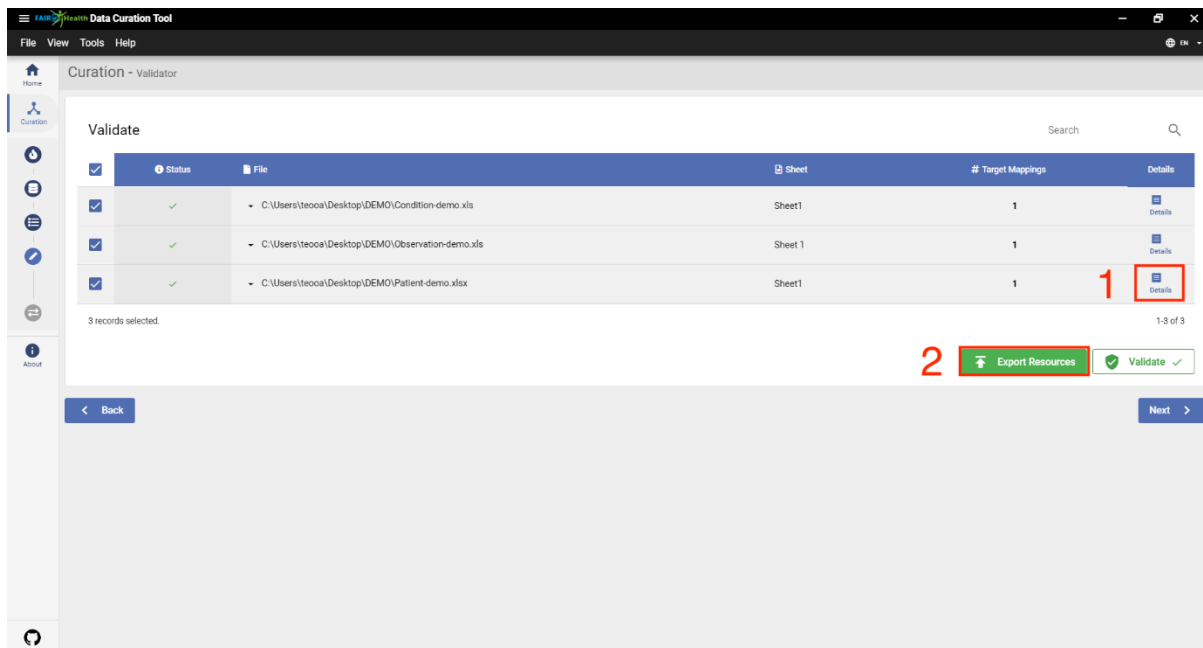


Figure 28. Validation Result Details – 1

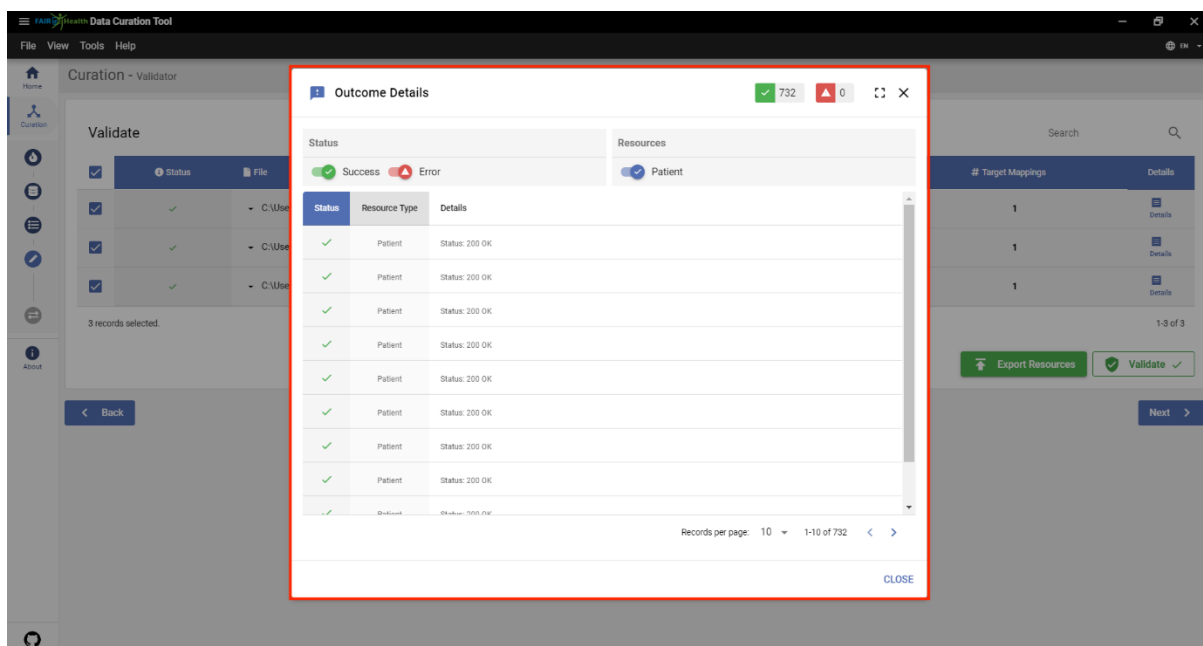


Figure 29. Validation Result Details – 2

2.1.4.3. Export Created Resources

When validation is finished, (shown in Figure 28) click on the #2 Export Resources button to export the resources created before putting them in the FHIR repository.

2.1.5. Transform

2.1.5.1. Put Resources into the FHIR Repository

In the last step of Curation (Figure 30), you will see the created FHIR resources listed. If there is a resource that you do not want to transform into your FHIR repository (although you have validated), you can remove it from the box you see number **1**. And then click on the Transform button **2**.

You will see the details of the actions on the left panel: pending, in-progress, error, success, or warning.

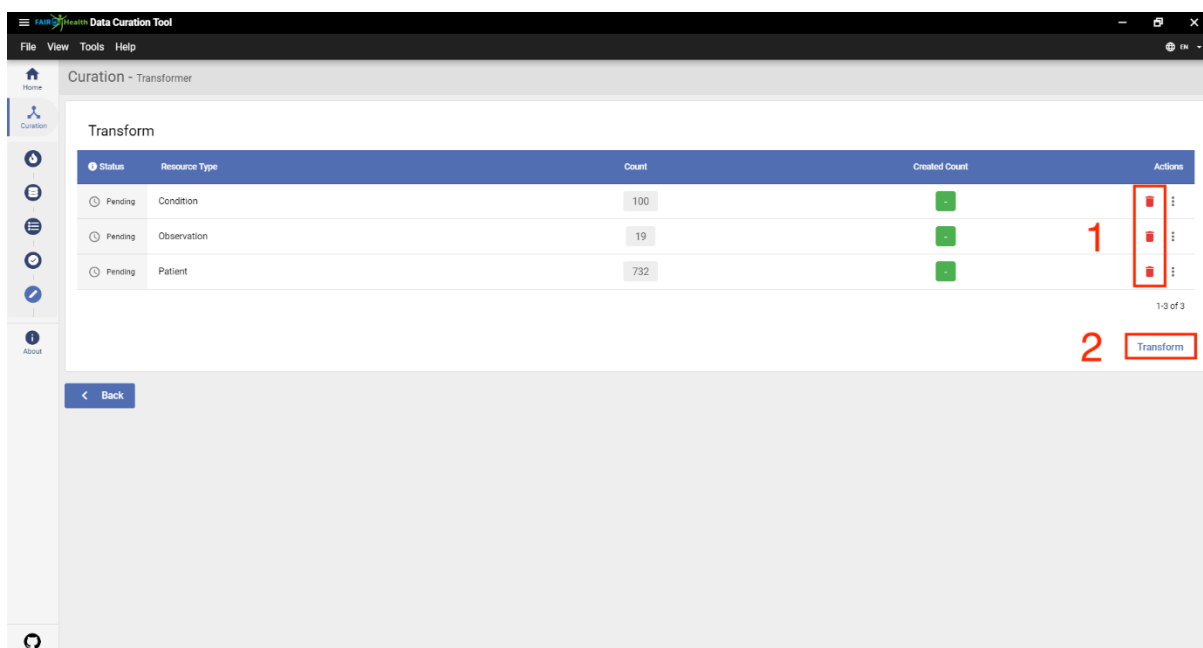
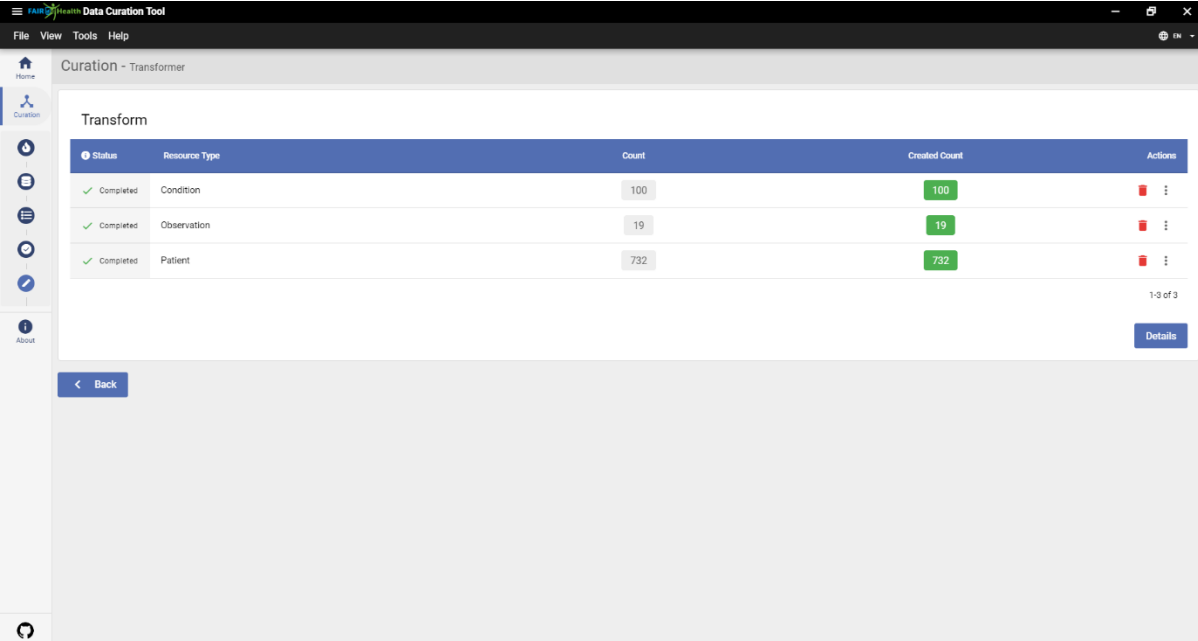


Figure 30. Transform Screen



| Status | Resource Type | Count | Created Count | Actions |
|-------------|---------------|-------|---------------|---------|
| ✓ Completed | Condition | 100 | 100 | |
| ✓ Completed | Observation | 19 | 19 | |
| ✓ Completed | Patient | 732 | 732 | |

1-3 of 3

[Details](#)

[< Back](#)

Figure 31. Transform Result

2.2. Tool Settings

In this section, tool-wide settings are described such as change language, adjust window size, developer options etc.

2.2.1. Change Language (Localization)

As shown in Figure 32, select *Tools -> Language* from the taskbar in the upper left panel.

Or you can change it at the top right of the app.

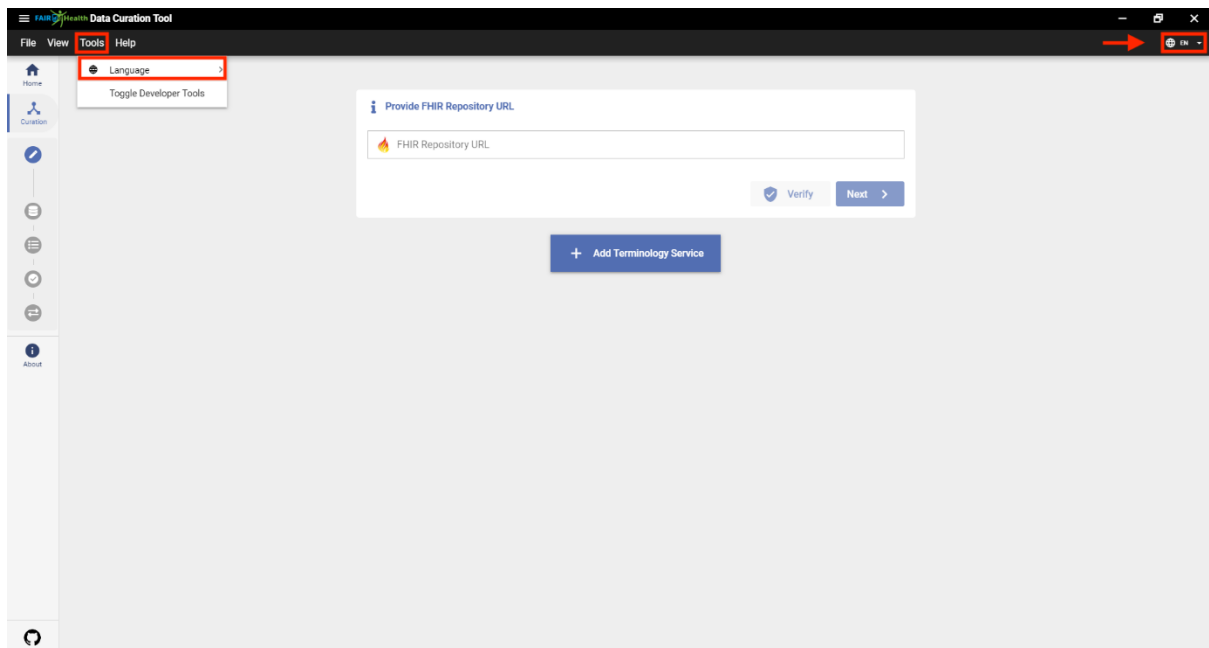


Figure 32. Tool Language Settings

2.2.2. Toggle Sidebar & Full Screen

In **Full Screen** mode, the tool expands the main window to occupy the entire screen. It does the same thing on macOS and extends it to the dock. In order to occupy the entire screen, the window expand button must be clicked, all operating system controls will be hidden, but you can access the main menu if you hover the mouse pointer over the top of the screen. Having these two expansion options in macOS is for ease of use. In Figure 33:

1 you can switch the full screen mode. In Windows, you can also do this from the button in the upper left corner. But in macOS, the application view will look like this.

2 you can toggle sidebar. Its open form is shown in Figure 34.

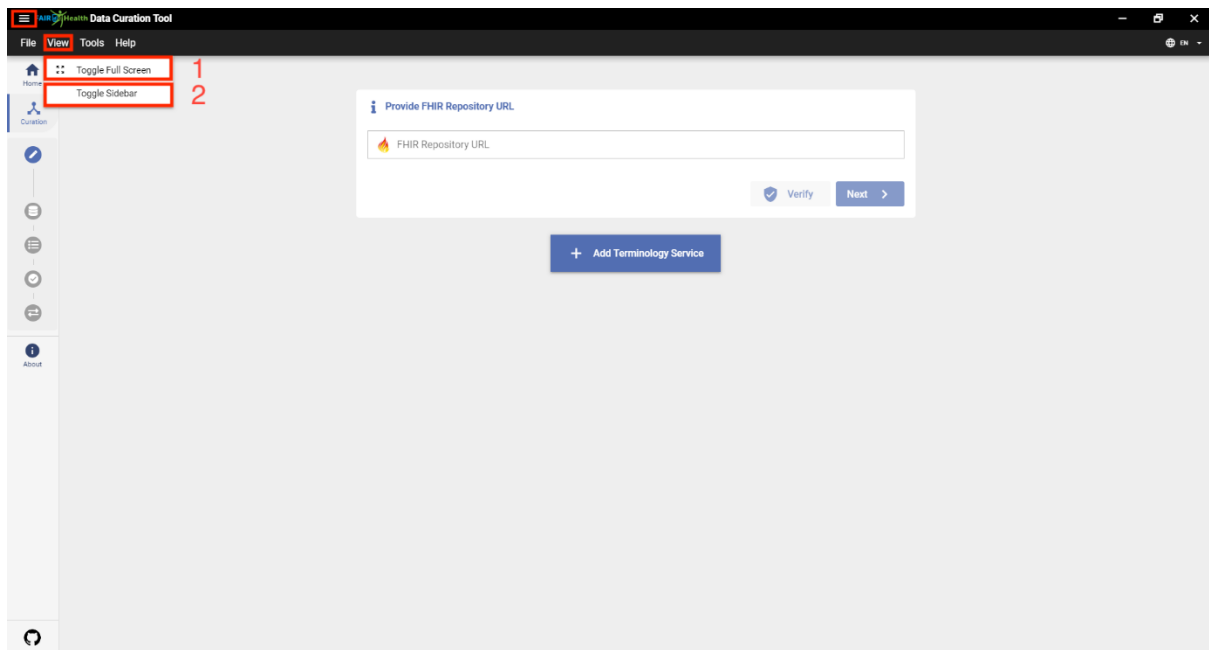


Figure 33. Tool Window View Options

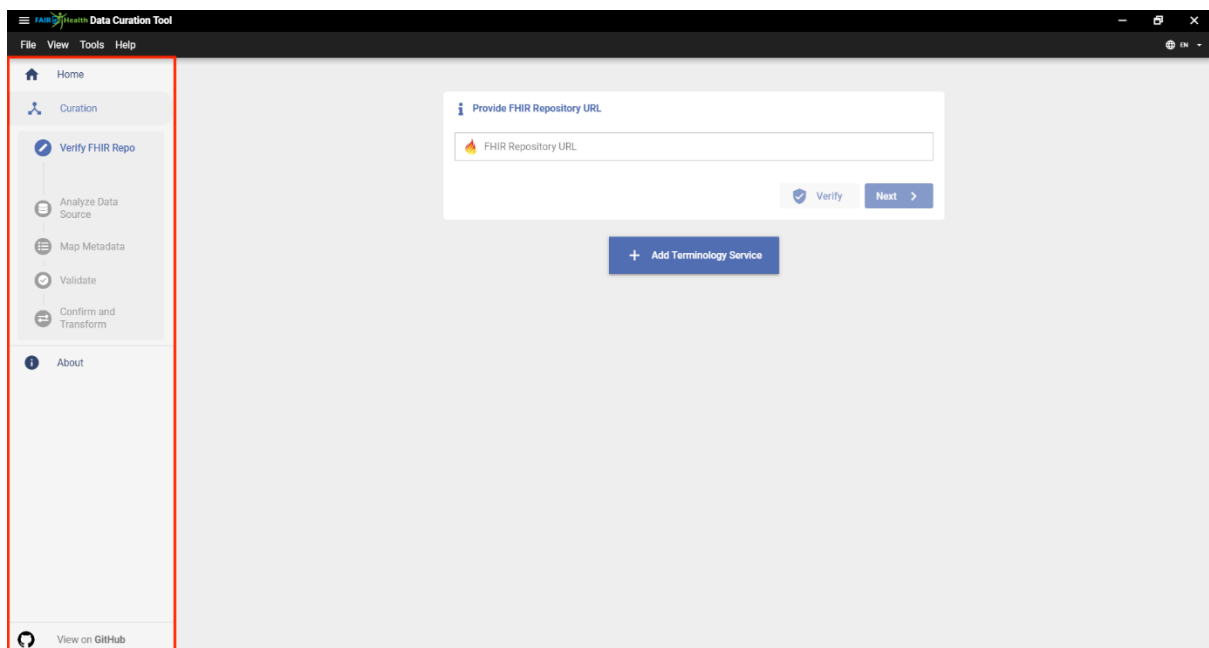


Figure 34. Open Sidebar

2.2.3. Debugging

If you are a developer and the application is not behaving the way you wanted it to, an array of debugging tools might help you find coding errors, performance bottlenecks, or optimization opportunities. The most comprehensive tool to debug the renderer process is

the Chromium Developer Toolset, so to open dev tools select **Tools -> Toggle Developer Tools** (Figure 35).

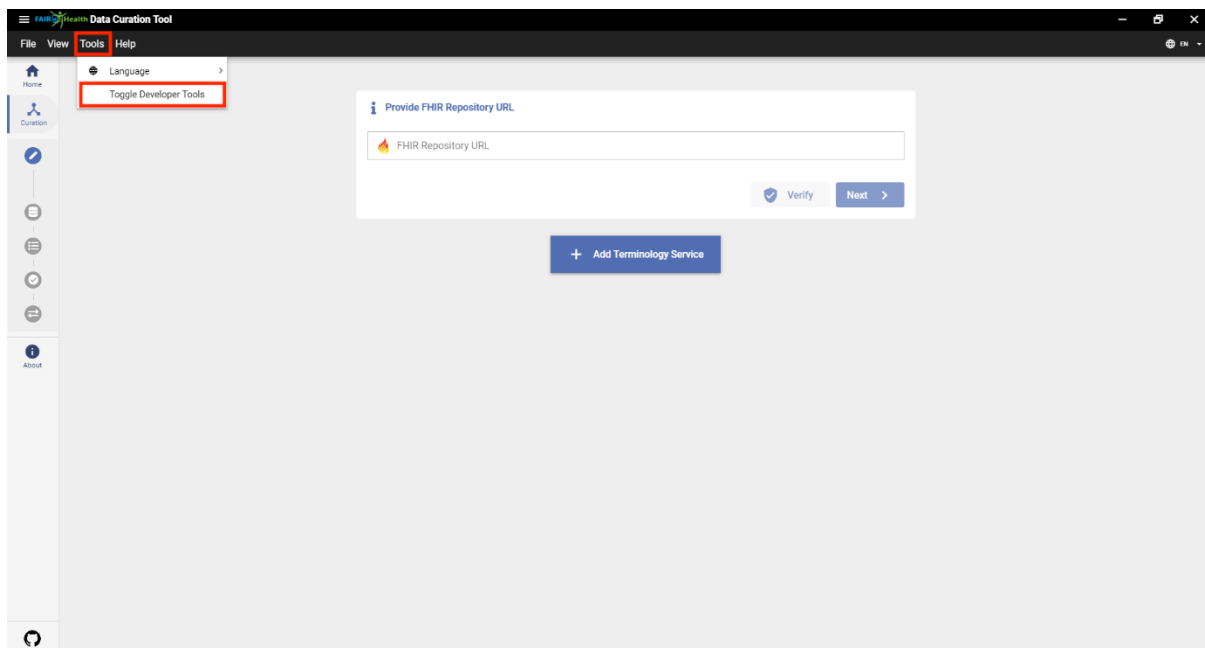


Figure 35. Open Developer Tools

2.2.4. Log Locations

Data Curation Tool log file contains information about events that have occurred from app start to end. These events including crash reports and validation & transforming errors are logged out by the tool and written to the file.

The tool writes logs to the following locations:

- ❖ on **macOS**: ~/Library/Logs/FAIR4Health Data Curation Tool/log.txt
- ❖ on **Windows**: %USERPROFILE%\AppData\Roaming\FAIR4Health Data Curation Tool\logs\log.txt
- ❖ on **Linux**: ~/.config/FAIR4Health Data Curation Tool/logs/log.txt

2.2.5. Help

If you have any questions, suggestions or any issues, you can contact us on the project GitHub page and explain your issue (Figure 36. Help - GitHub Repository).

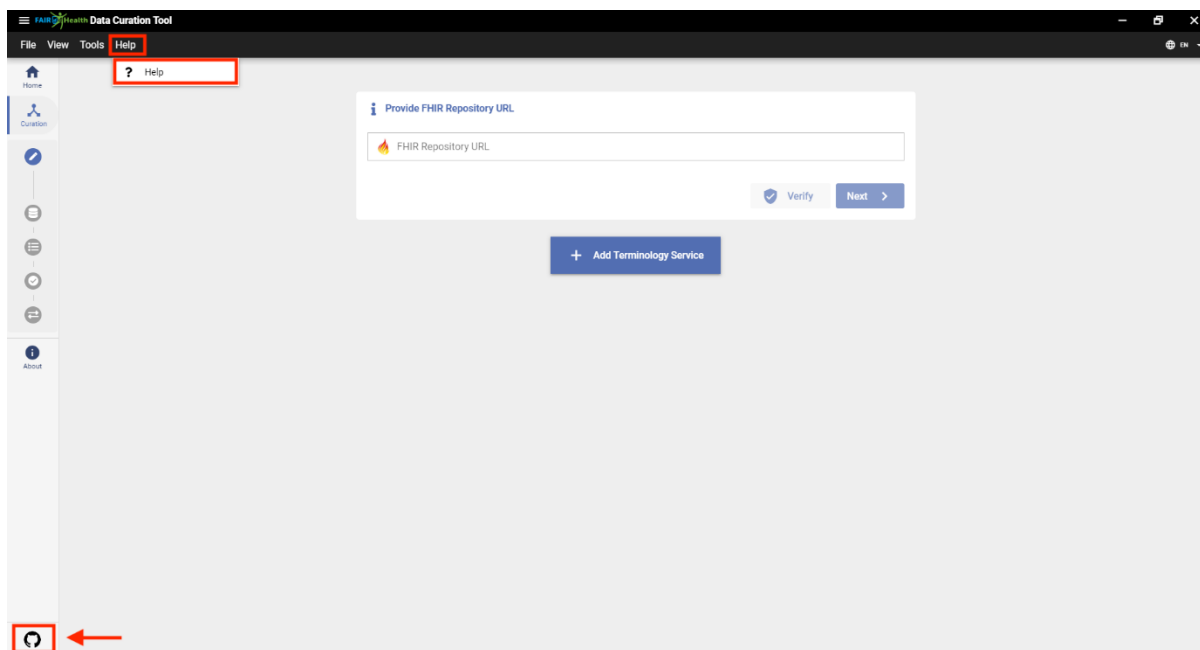


Figure 36. Help - GitHub Repository