HEDDA.IO

Excel Add-In

End User Documentation

oh22

2024-10-24

smart
casual
datadesign

Contents

1 Overview

HEDDA.IO is a comprehensive Data Quality solution that enables easy integration of standardization, cleansing, matching, and enrichment tasks into existing processes.

By integrating an Excel Add-In into our solution Users get to directly work with data within a familiar environment, Excel being one of the most widely used tools for data management and analysis across industries.

With the Excel Add-In for HEDDA.IO Users can load Executions, view execution statistics, and easily export data back into Excel, enabling immediate editing without switching between solutions.

This is our step-by-step guide on how to use our Excel Add-In.

2 General Flow / Usage

1. Get the HEDDA.IO Add-On.
2. Add the HEDDA.IO Excel Add-In to the Ribbon Bar.
3. Add an Environment.

2.1 HEDDA.IO Add-On

For a HEDDA.IO, please contact us directly under https://hedda.io/get-started/.
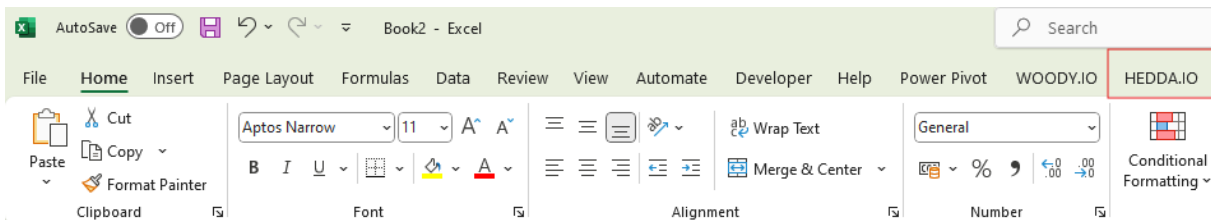
2.2 HEDDA.IO Excel Add-In

Open Microsoft Excel 2016 or latest, Ms Office 365 Excel on Windows or Microsoft Excel Online (supported versions).

Open the Home Menu in the Excel Ribbon at the top of the Excel interface. In the Home Menu, go to My Add-ins on the right side and click on it. A pop-up will appear with all your Office Add-Ins. Search for HEDDA.IO.
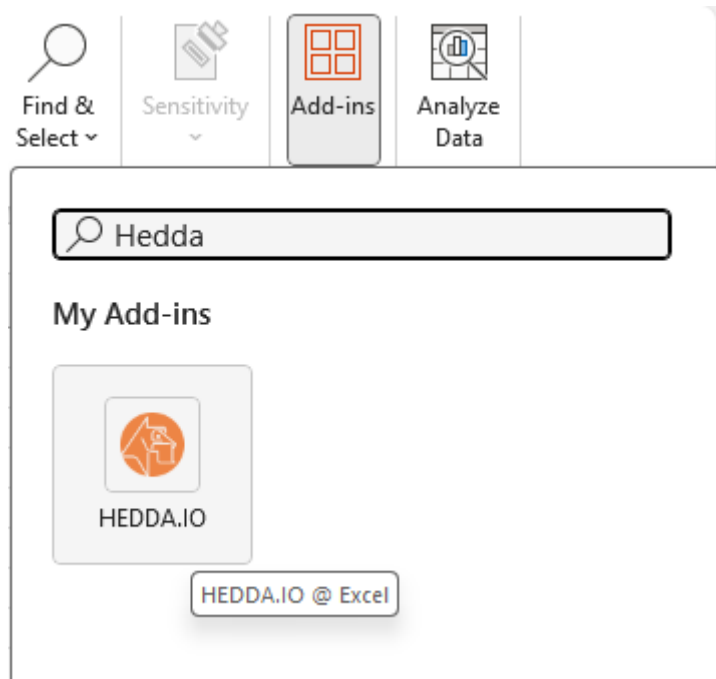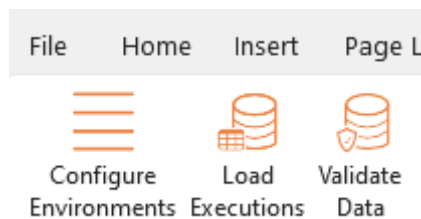
Figure 1: "Add HEDDA.IO Add-In"

Upon selecting the Add-In, HEDDA.IO will appear in the Excel Ribbon:

**Figure 2:** "HEDDA.IO Excel Add-In Ribbon"

Clicking on the HEDDA.IO in the Excel Ribbon will display new buttons to your Ribbon Menu: Configure Environments, Load Executions, and Validate Data.


Figure 3: "Add-In Buttons"

Upon completing the installation, the first step is to start with `Configure Environment`. In this case, an Environment is a HEDDA.IO instance to which you can connect in order to authenticate in order to retrieve the Projects which you have access to. Clicking on the button will trigger a task pane on the right side.
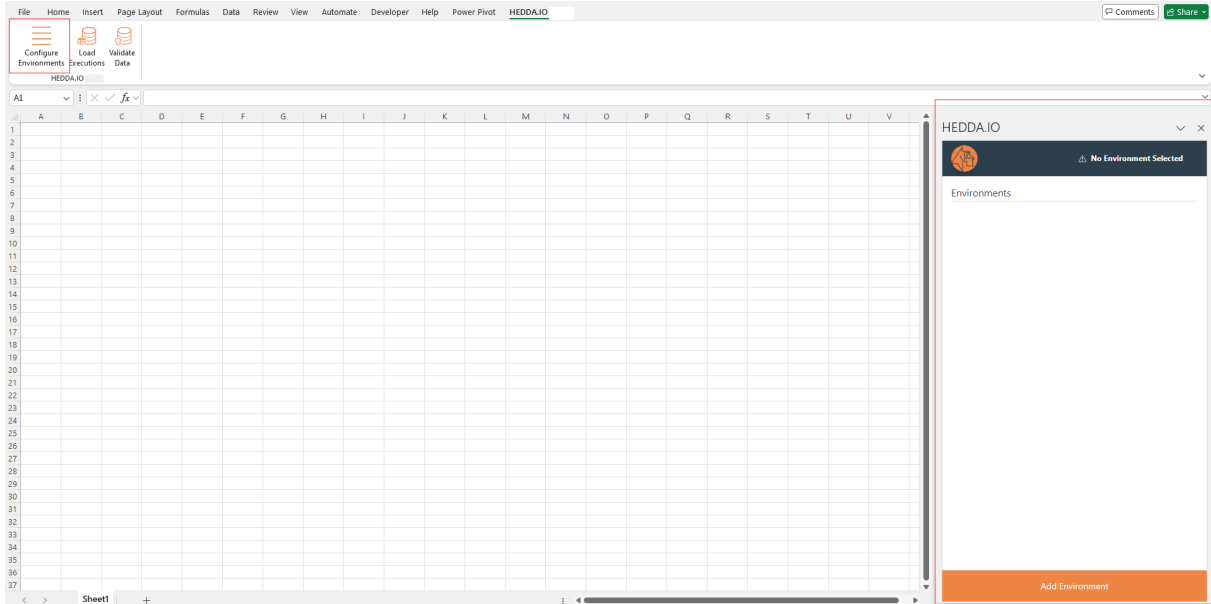

Figure 4: "HEDDA.IO Task Pane"

The task pane presents you with the HEDDA.IO icon, that will always bring you to the Home Menu, displaying the same functions, as in the Ribbon Menu: Configure Environments, Load Executions, and Validate Data.
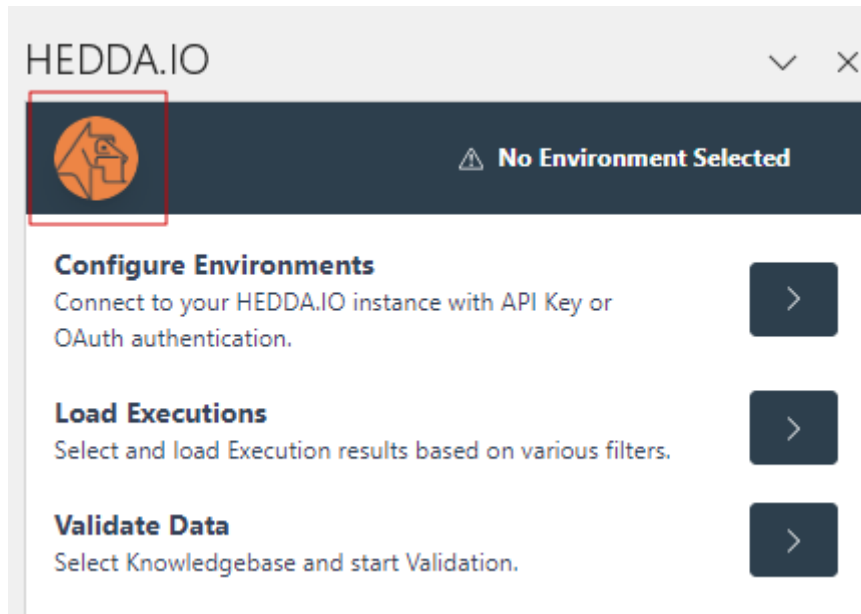
Figure 5: "HEDDA.IO Home Menu"

On the right side you will see the selected Environment, as soon as you have added one. On the bottom you will find the call-to-action button.
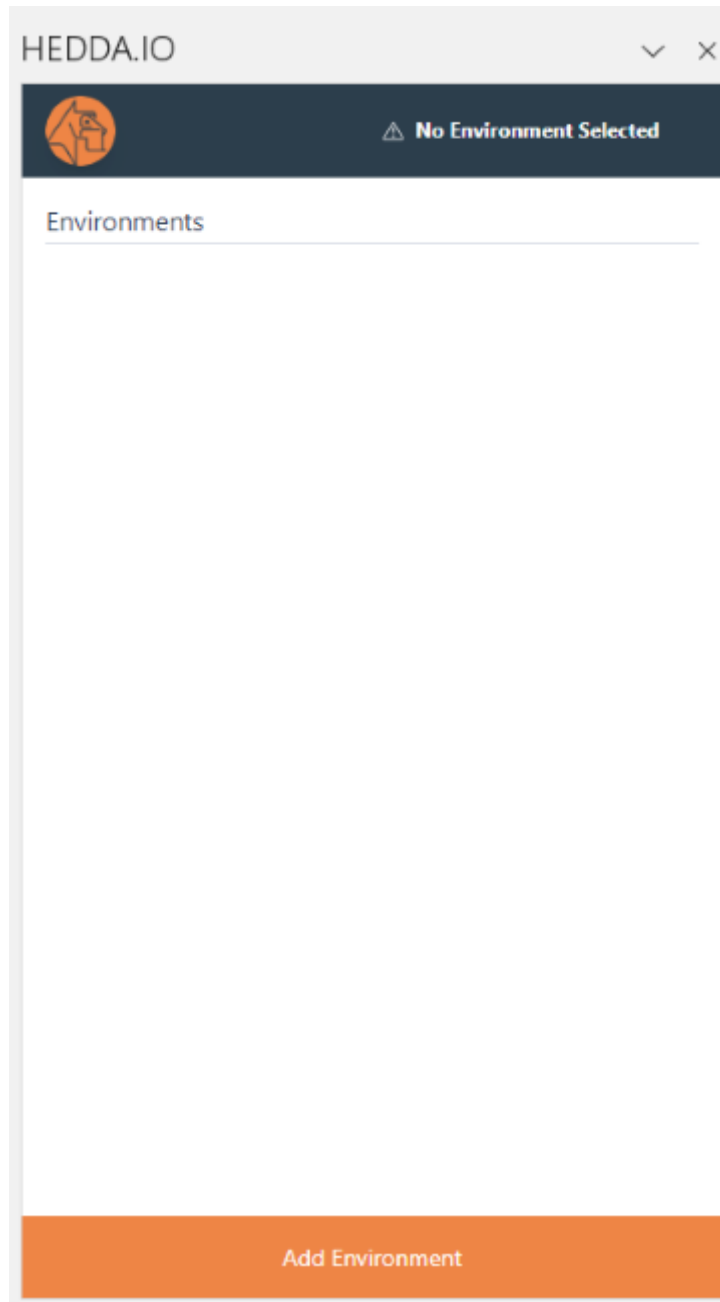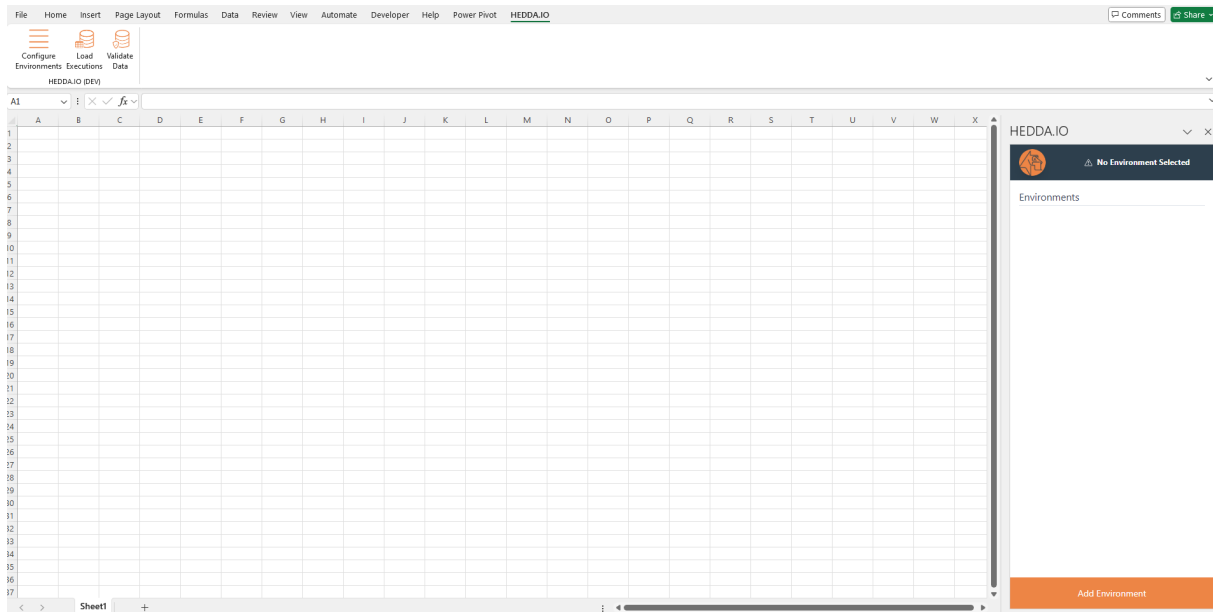
Figure 6: "Task Pane Closeup"

3 Configure Environments

3.1 Adding Environments

The first step is to add your Environment. After clicking on `Configure Environments` in the Ribbon Menu, the Add-In will appear on the right of your Excel spreadsheet helping you to identify the Environment.

Figure 7: "Configure Environments"

To add a new Environment, click on `Add Environment` and fill out the required input fields.
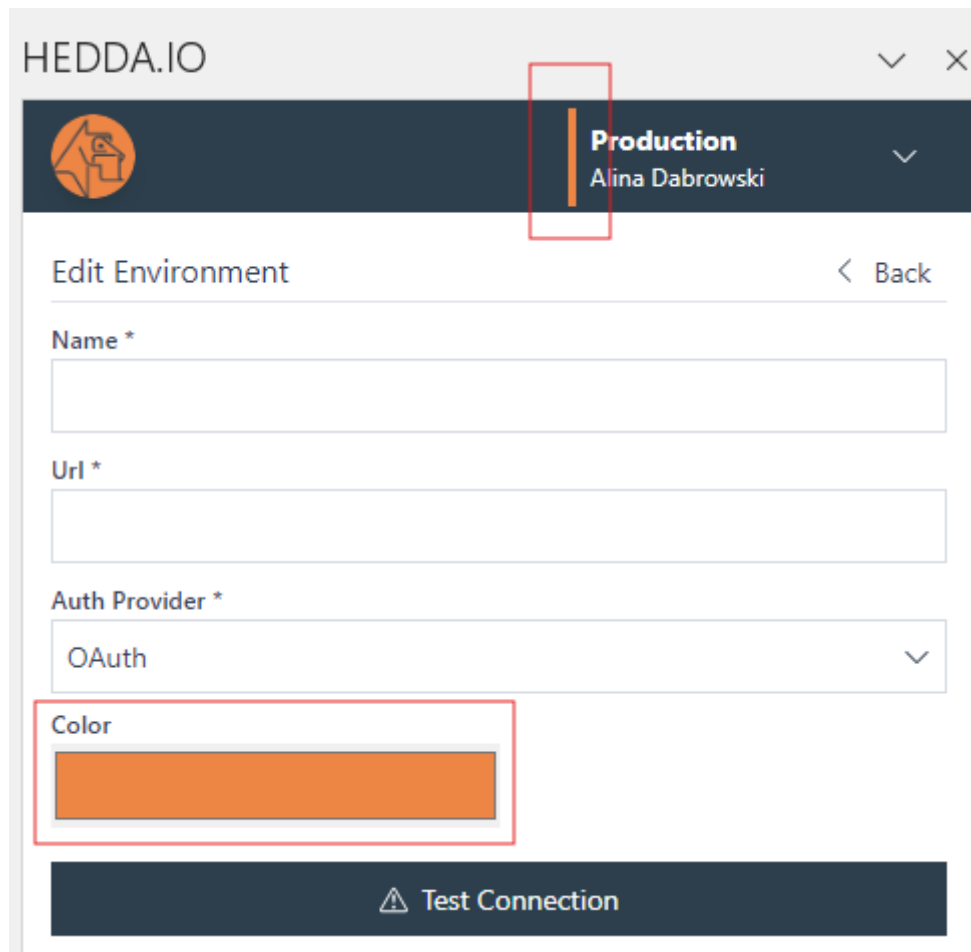
Figure 8: "Add Environment"

You can provide your own name, fill in the Environment URL (the URL you see in the browser when you are on the home page of your HEDDA.IO instance) and choose between an API-Key or OAuth as an authentication method.

API-Key authentication is a method where a unique key is used to authenticate API requests, which you can create in the HEDDA.IO Web App. See Documentation Chapter 3.3.1 API Keys Manager. OAuth, on the other hand, is a token-based system that allows user-specific access and permissions, often involving user consent.


Figure 9: "Auth Provider"

In addition, you can select a color for the strip to mark different systems for an easier overview. For example, red can stand for a productive system with live data.

Figure 10: "Environment Color"

Once all parameters have been entered, the connection can be saved. 'Test Connection' Clicking the 'Test Connection' button checks whether the Authentication works against the Environment and thus 'validates' it. If the validation is successful, the Environment can be saved.
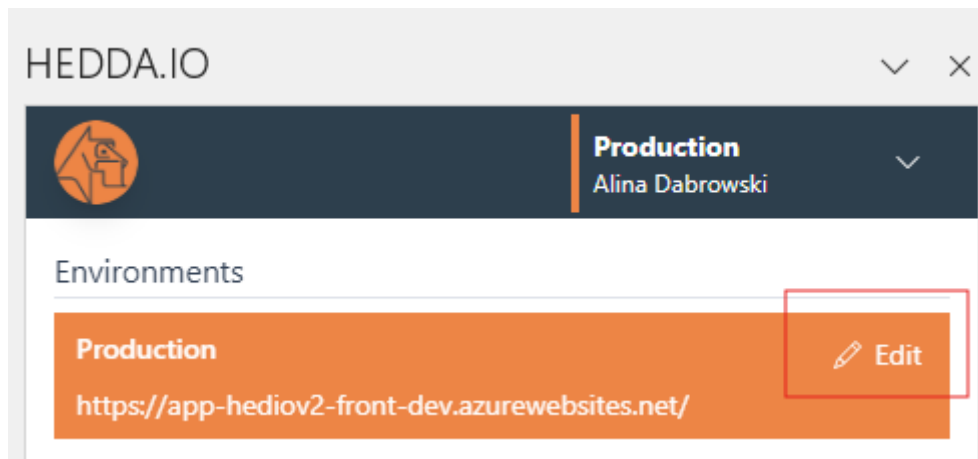
Note: If the Authentication fails, the Environment is invalid and cannot be used.
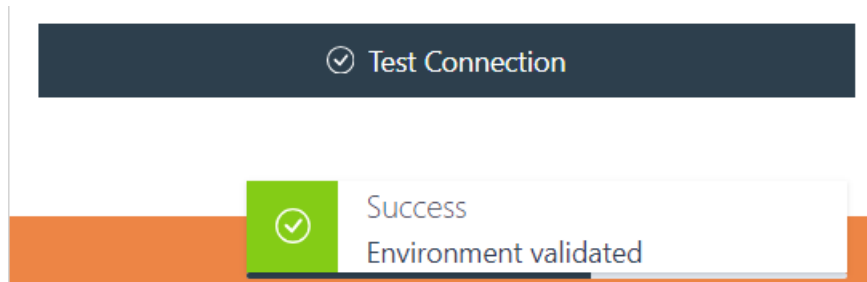
Figure 11: "Test Connection"

3.2 Edit Environments

If you want to edit the Environment, press the `Edit` button.
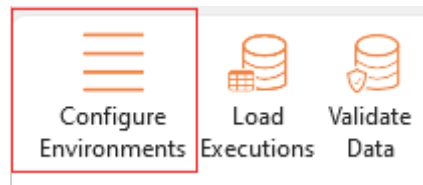

Figure 12: "Edit Environment"

Here you can adjust the parameters, change the color or delete the Environment completely.

3.3 Changing Environments

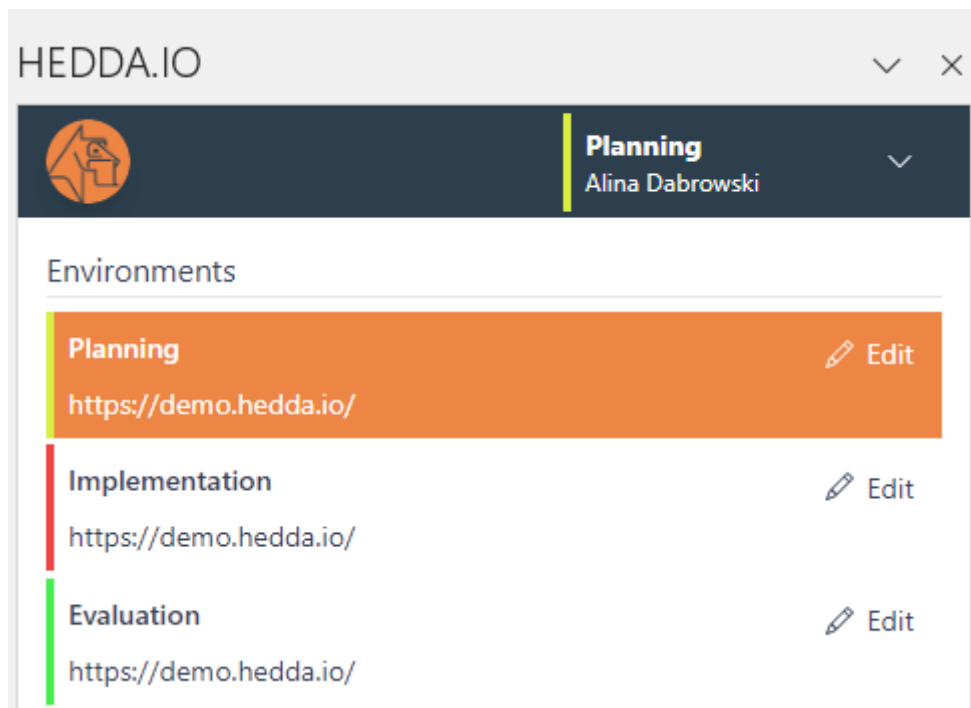If you have created several Environments and want to change from one to another, there are several options:

3.3.1 Ribbon Menu

The first option is to click on `Configure Environments` in the Ribbon Menu.

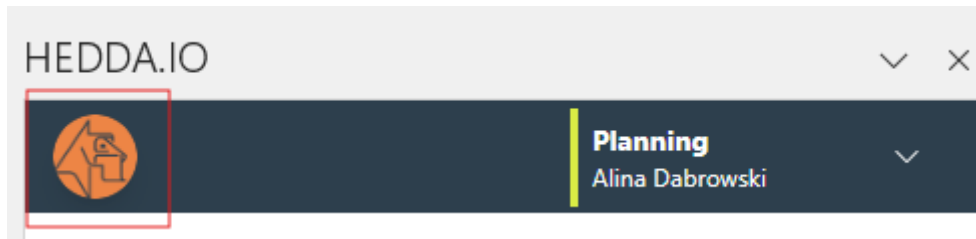**Figure 13:** "Configure Environments Ribbon Menu"

That will take you to all Environments you have added.

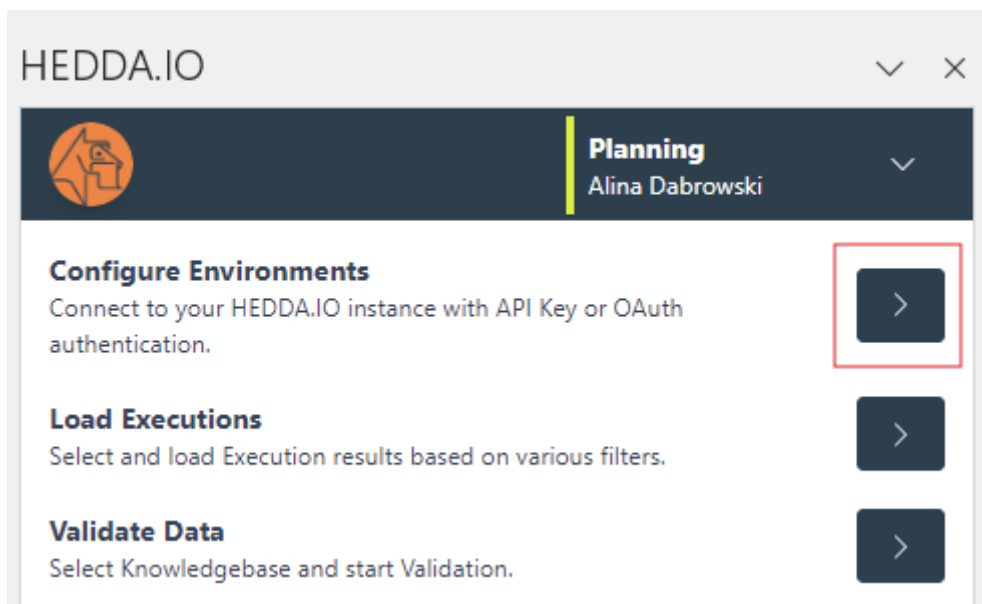**Figure 14:** "Environment Overview"

3.3.2 Home Menu

Another option is to click on the HEDDA.IO icon. This takes you to the Home Menu which contains the same processes as the icons in the Ribbon Menu.
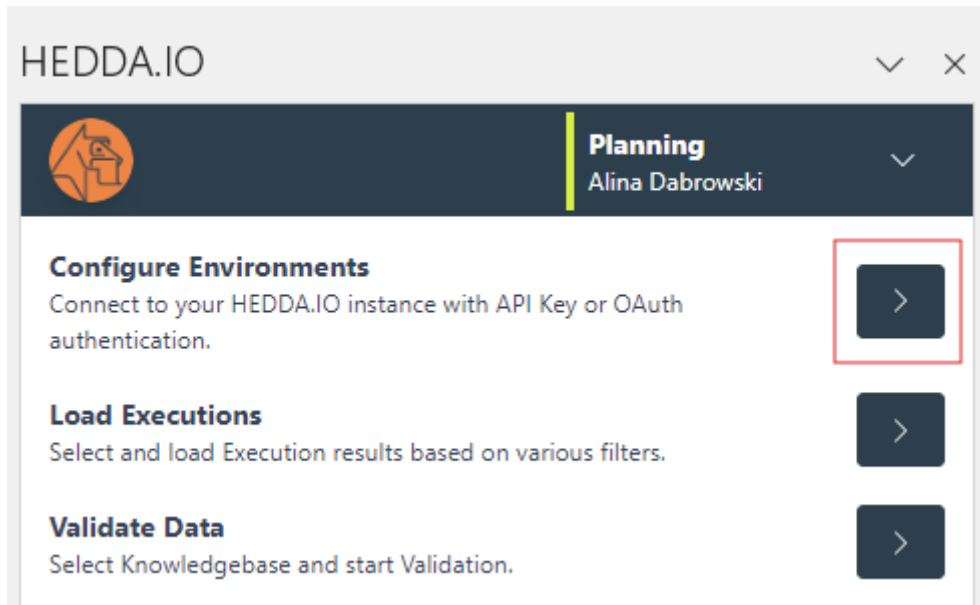
Figure 15: "Home Menu Icon"

If you then select `Configure Environments`, you will be taken to an overview of all the Environments you have added.

Figure 16: "Configure Environments Home Menu"

3.3.3 Drop-Down Menu

The third option would be to use the drop-down menu in the task pane and choose between your Environments.

Figure 17: "Configure Environments Drop-Down Menu"

4 Loading Executions

Executions are results of tests in HEDDA.IO. They contain the results of Runs to which previously defined Business Rules have been applied.

Executions can be loaded and the data from them displayed in Excel.

4.1 Selecting an Execution

After the desired Environment has been added, an Execution can be loaded. The displayed parameters must be filled in for this.

**Figure 18:** "Load Execution Selection"

4.1.1 Project

Projects are essentially workspaces within HEDDA.IO, designed to cater to your distinct needs for utilizing the HEDDA.IO platform. Unless you have a Project established, it is not possible to generate a Knowledge Base. Knowledge Bases are entities responsible for containing comprehensive information about the structure of your data, the regulations it needs to adhere to, Execution metrics, and a variety of additional details.

4.1.2 Version

The Version here refers to the Knowledge Base. We differentiate between a Live Version and an Edit Version.
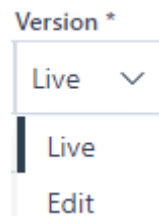
Figure 19: "Version"

While the Live Version is strict read-only, the Edit Version allows you to apply changes and publish those to update the Live Version.

4.1.3 Run

Within the HEDDA.IO application, a Run is simply a container in which Executions and their respective Statistics are stored. Runs add context to Executions so that you can then easily locate the statistics linked to respective Executions, whenever you require.

Suppose, you are initiating HEDDA.IO Executions utilizing distinct databases. If the Executions share the same Run, it would make the result convoluted and hard to locate specific metrics within the multitude of Executions. To prevent that, it is advisable to create a separate Run for each case. This way Execution statistics remain organized and easily accessible. Mappings and Tags can further help create more fine-grained Use Cases.
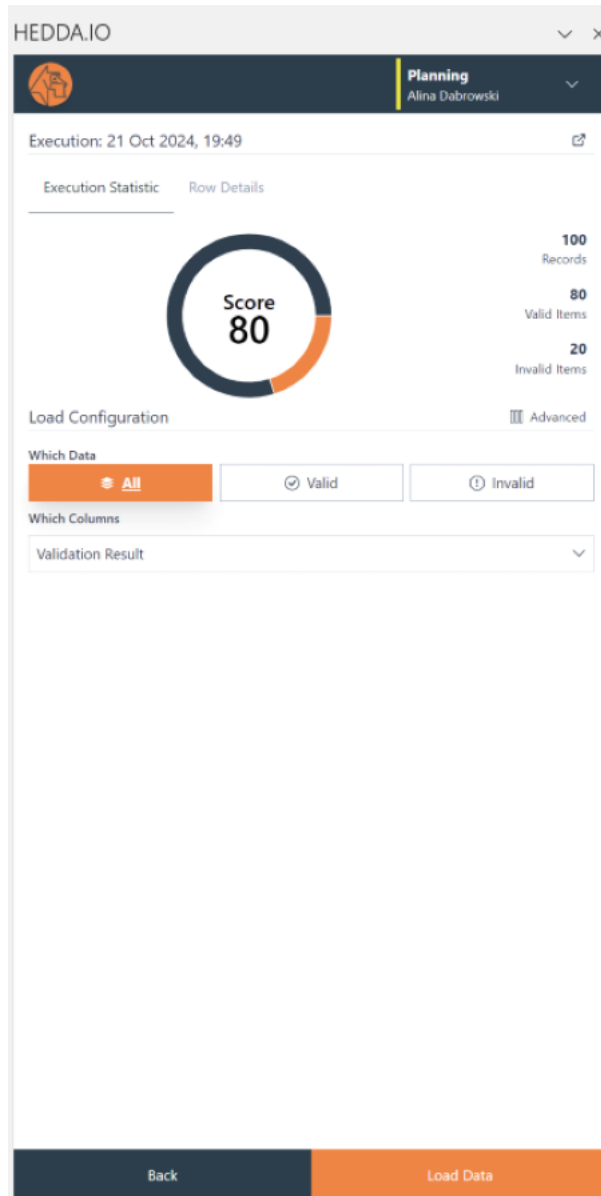
As soon as you have selected the required parameters, you then will be able be pick an Execution and klick on Next, to proceed.

4.2 Loading Execution Data

After loading an Execution, an overview with the general execution statistics is displayed. You can see not only the number of records, but also how many of them are valid or invalid.

Figure 20: "Load Execution Details"

There are now several options for displaying the result of the statistics: `All`, `Valid`, `Invalid`. The status `Valid` or `Invalid` depends on whether the data fulfills the previously defined Business Rules or not.

Furthermore, you can select the following columns:

- Data only - only shows the data from the Execution.
- Comparison - compares the data from the Execution with the original data set. The separation is done by naming the columns as well as by color separation. The current data set is displayed in blue, while the original data set is displayed in orange.
- Validation Result - if the Business Rules are fulfilled, the row is displayed as "TRUE". If they are not fulfilled, the validation result is displayed as "FALSE". The columns with the Rulebooks and BBusiness Rules that are valid or invalid are also displayed with "TRUE" or "FALSE".
- Full: all parameters are displayed.

4.2.1 Advanced Mode

Should you require a more advanced configuration, you can choose the `Advanced` mode.
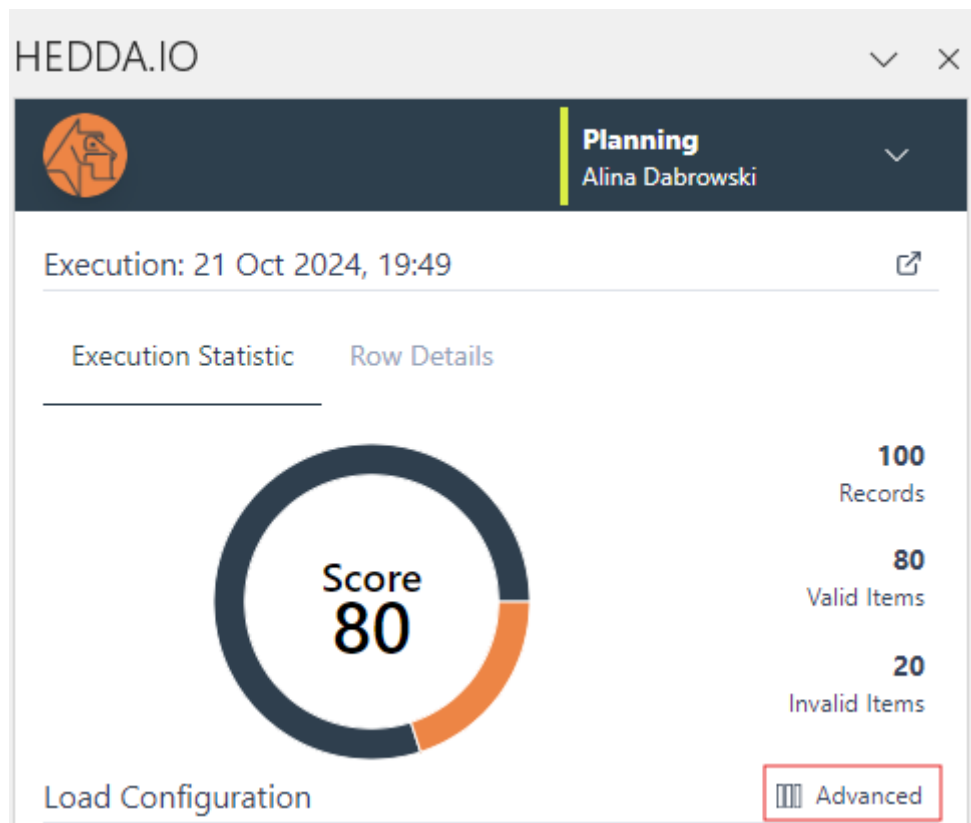
Figure 21: "Advanced Mode"

In `Advanced` mode, you are no longer tied to the predefined parameters and can display very specific evaluations. If you are only interested in invalid data and the Business Rules, you can set the corresponding sliders to adapt the overview to your requirements.
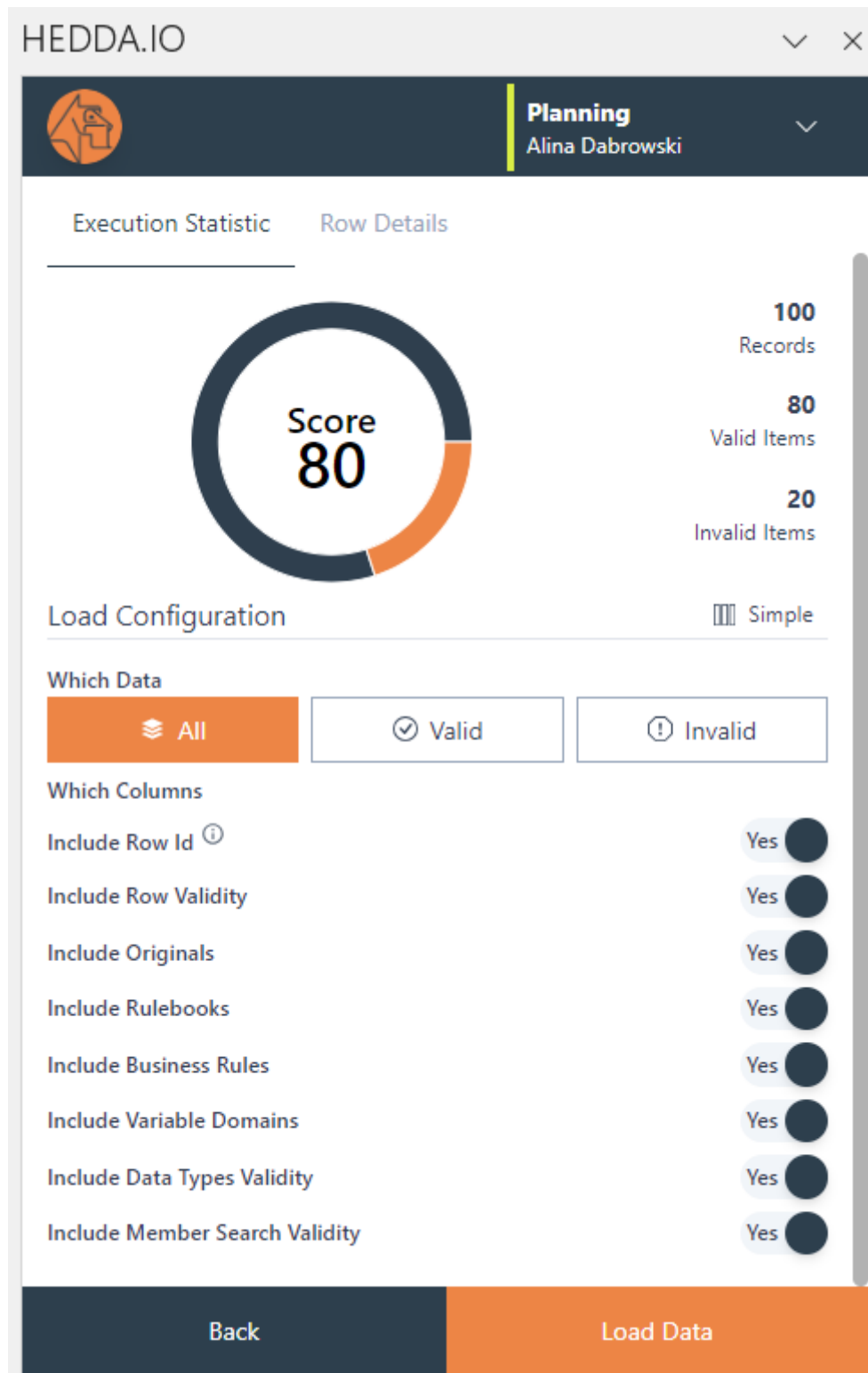
Figure 22: "Advanced Mode Selection"

After the Execution has been loaded, the result appears as an Excel export on the sheet:
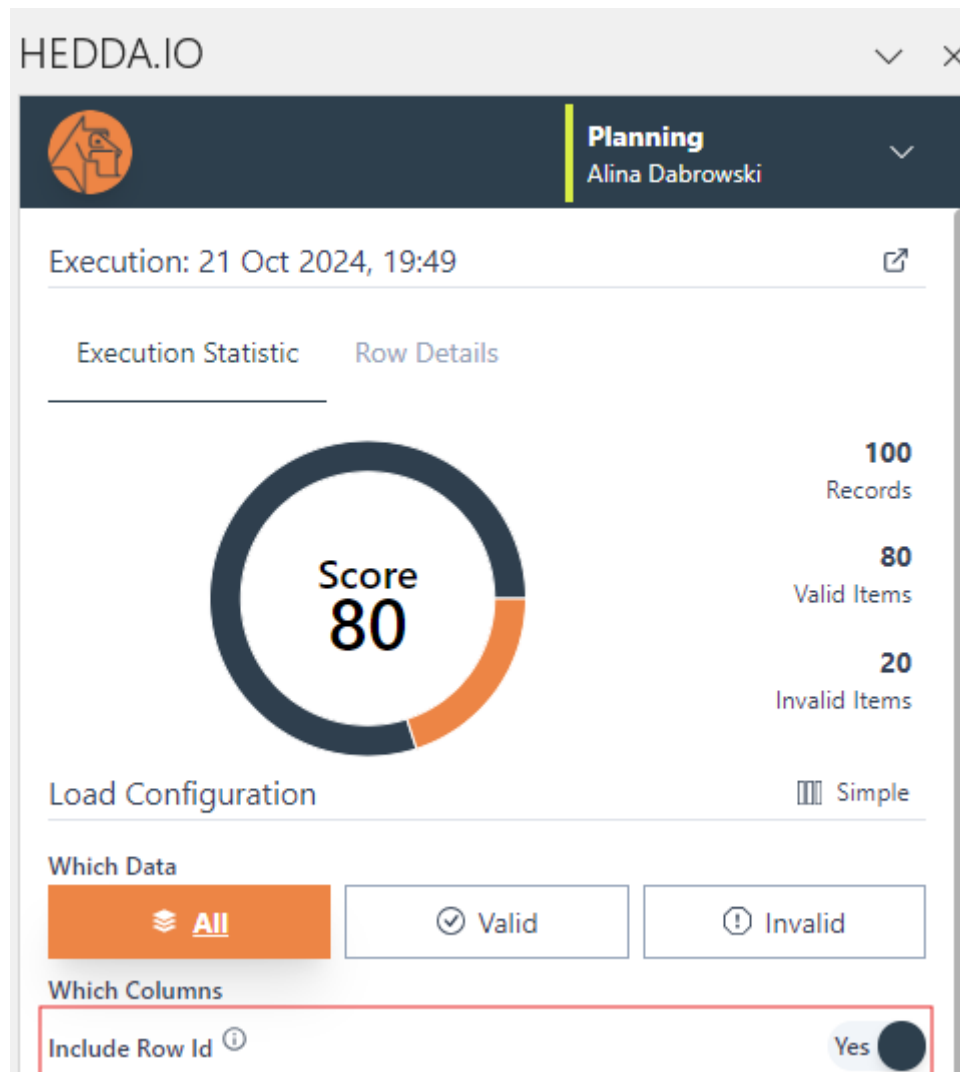

Figure 23: "Loaded Data with 'Full' Configuration"

4.3 Row Details

Including a Row ID column in the Advanced settings enables Row Results.

Figure 24: "Include Row ID"

If you have included a Row ID column, you can now select a row in the Excel sheet and proseed to Row Details. This view displays an overview of the result of the Execution in this Row. The domains with the current values from the Execution and the associated original values are displayed in summary. Furthermore, it allows you to see more precisely whether and why the result is valid or invalid.

4.3.1 Domain Values

In the Row Details, the Domain Result provides an overview of the domains that match the column headings in the Excel sheet. You can also see the comparison between the loaded data as Value, and the original data set as Original Value.

Figure 25: "Domain Values"

4.3.1.1 Toggle Variables On the right-hand side, you also have the option of making Variable Domains visible, which are then also displayed in the Domain Result overview. Variable Domains aren't

loaded from the source data but are added by HEDDA.IO and they are used during data validation in Rulebooks.
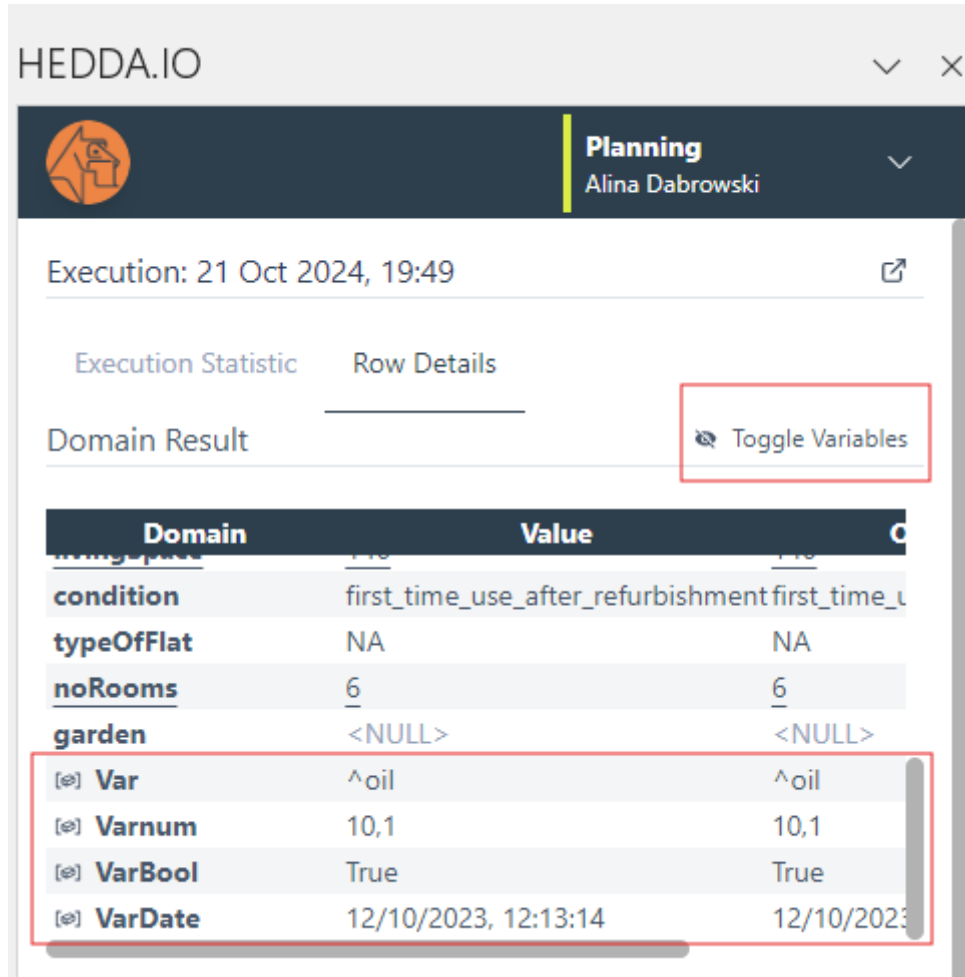
Figure 26: "Toggle Variables"

These are special domains (columns in the results data) that are not automatically filled with data from the data source. A default value can be entered, or the value can also be filled or changed using Business Rules. If you want to know how to create and manage domains, you can find this information in the HEDDA.IO End User Documentation 2.7 Domains.

4.4 Validation Result

The `Validation Result` displays the overall validity status of the Row with the possibility to see the result for each Rulebook in the drop-down and upon selection for each Business Rule within.

If there is a mismatch, the `Validation Result` is displayed as invalid. From the Rulebook drop-down menu, you can also see which Rule was not fulfilled. Rulebooks and Rules which weren't executed

are displayed greyed out.
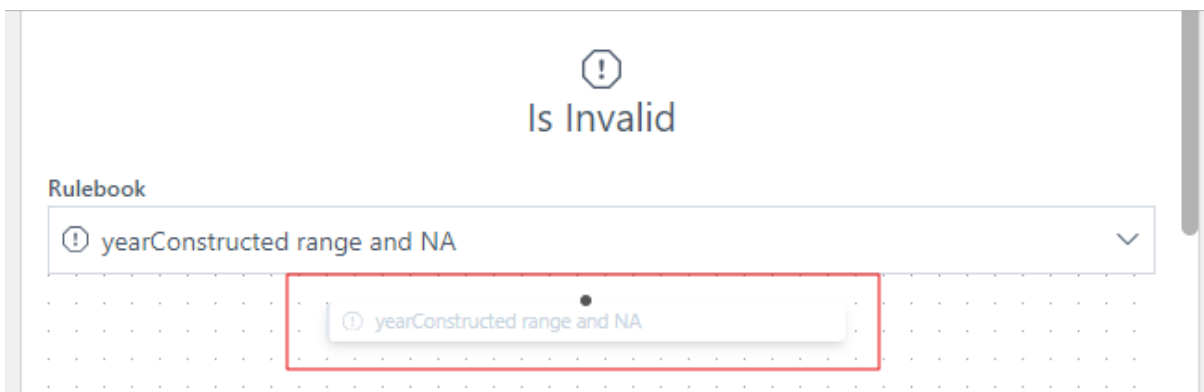
Figure 27: "Invalid Validation Result"

It is also possible that several Business Rules were not fulfilled. These are displayed accordingly.

Figure 28: "Multiple Invalid Validation Results"

It is important to remember that different validation processes run through the Business Rules in the background. Some are simple, others have multiple nodes. As an example, the Rulebook "heatingType NA" in HEDDA.IO looks like this:

Figure 29: "heatingType NA Rulebook"

The display is then mirrored in the Excel Add-In:

Figure 30: "heatingType NA Rulebook Excel Add-In"

4.5 Open Execution in HEDDA.IO

To take a closer look at the Execution with the Rulebooks and Business Rules, you can open it directly in HEDDA.IO.
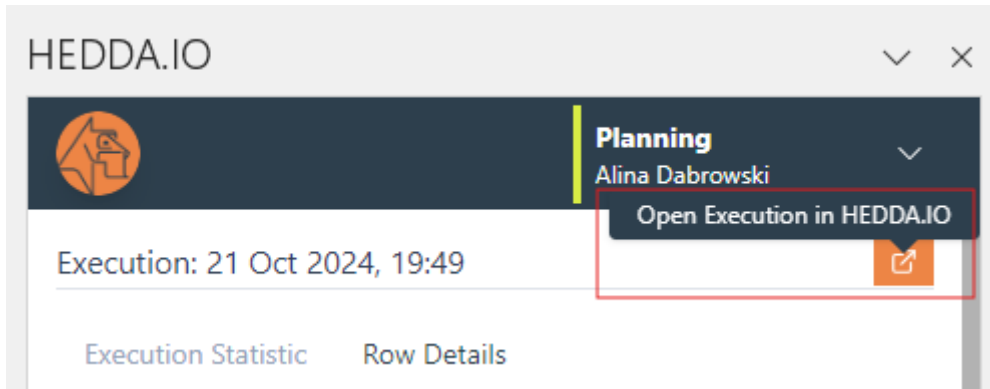
Figure 31: "Open Execution in HEDDA.IO"

4.6 Opening Excel Files with Data from HEDDA.IO

If you have saved an Excel file with Executions, or if somebody sent you one, you can load it again with Excel. However, you will notice that only the data in a sheet is displayed there. If you want to return to the Add-In, you have to select the HEDDA.IO Add-in in the Excel Ribbon again. Selecting `Load Executions` in the Ribbon Menu that appears underneath, the HEDDA.IO Add-In automatically suggests loading an Execution. If you click on `Fetch Execution`, you are automatically redirected to the Execution Statistic and can continue working with it.
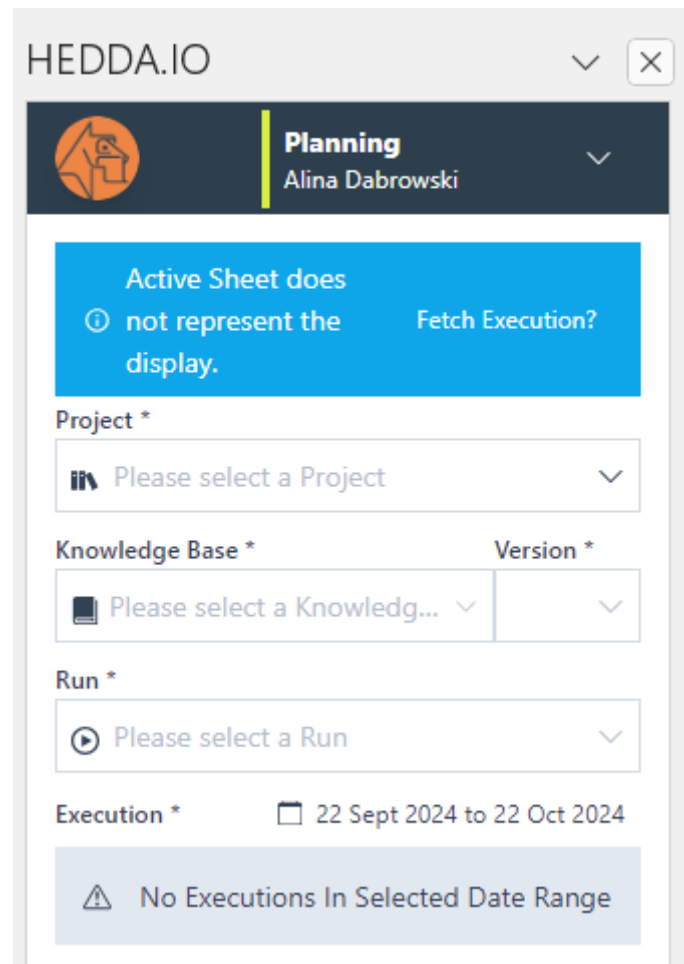
Figure 32: "Fetch Execution"

5 Validating Data

The Excel Add-In also offers users the option of validating their own data against Business Rules created in HEDDA.IO.

To serve as an example, a Project "Audit" with a "Travel Cost Audit" Knowledge Base was created in the Live Version in HEDDA.IO. The Domains containing the names of the employees and the travel costs will be our columns in Excel. "Cost Check" is created as a Rulebook with the Business Rules that the "Total Costs" should be greater than 0 and less than 1000. Everything that does not fulfill the Business Rules is displayed as invalid.
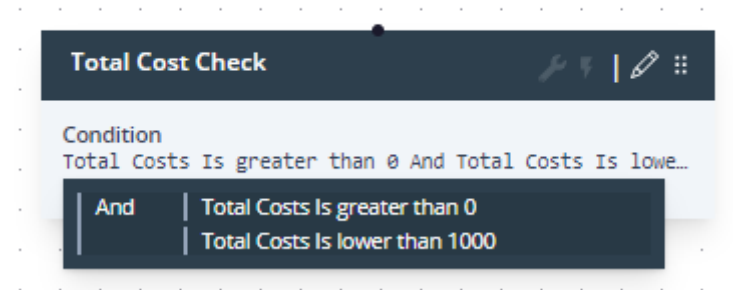
Figure 33: "Rulebook Total Cost Check"

5.1 Validating Data from Sheet

Once you have prepared the data you want to validate in an Excel sheet, click on `Validate Data` from the Ribbon or Home menu.

Figure 34: "Validate Data Button"

The configuration window will then open. Here you enter what you want to validate your data against. With `Select Data` you can either validate the entire table, if you have previously formatted it as such, or just an individual selection of data. It is important that column headings match the Domain names.

> Note: If you want to keep the original data, it is recommended that you copy it to another sheet and perform the Validation there.

Once the Validation started, the process for `Load Execution` is automatically triggered to carry out the Execution with the previously defined parameters. The result of the Validation is displayed as an Execution, where you can take a closer look at the result of the Execution Statistics and the individual Row Details.

Figure 35: "Configure Validation"

An invalid position is displayed during Execution. This means, that a Business Rules has not been fulfilled. In this case, the Total Cost Check has failed, since the amount should not have surpassed 1000. If you now want to load the Configuration, select the corresponding parameters and click on `Load Data`. This will trigger an Excel Export of the data onto the sheet.
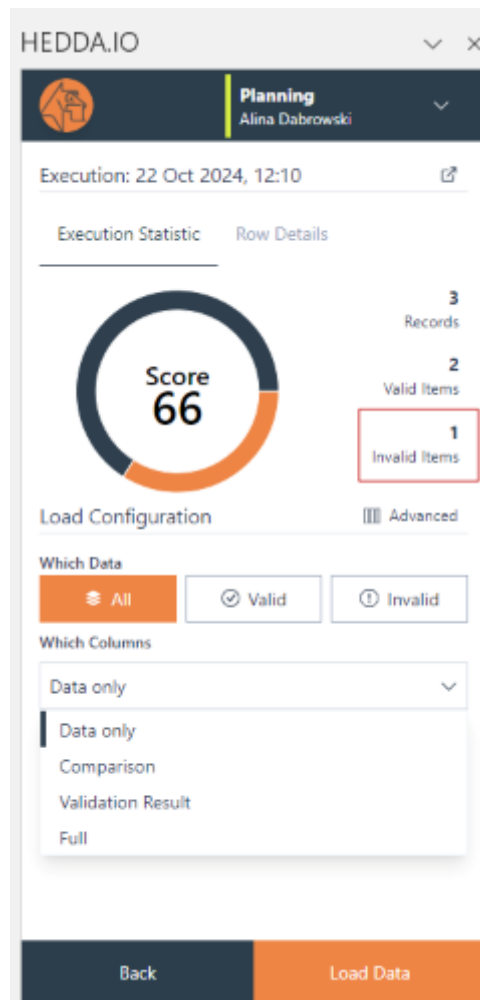
Figure 36: "Validation Execution"

In the view, you can see on the Excel sheet which result is invalid and which Rulebook and which Business Rules are affected.
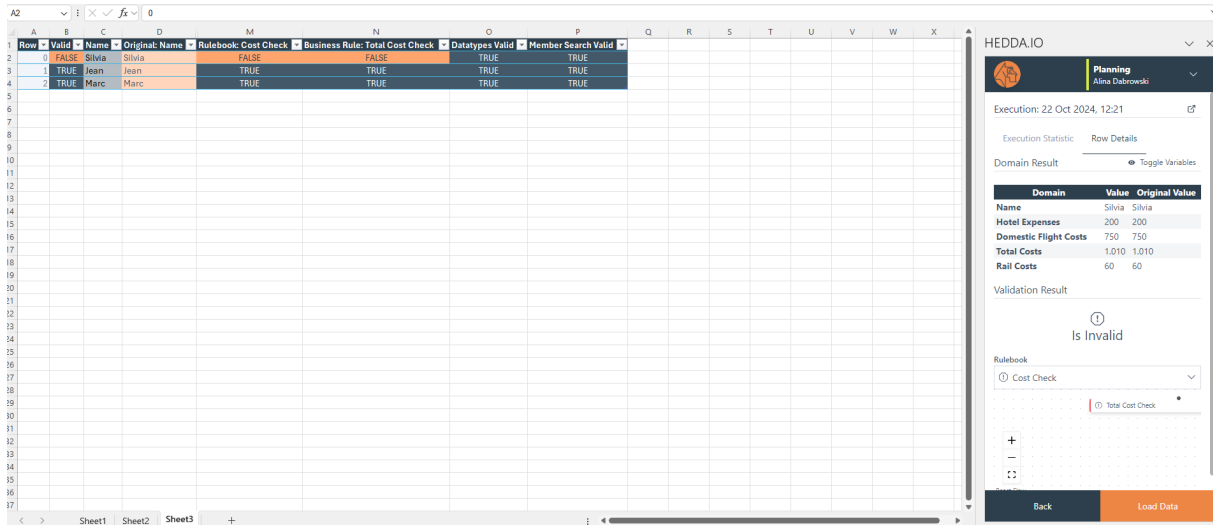
Figure 37: "Validation Result"

6 Glossary

6.1 Business Rules

Business Rules can be designated as Dataflow Rules, enabling HEDDA.IO to treat them as logical operations only. This configuration excludes them from the validation results and associated statistics, emphasizing their role in data manipulation or transformation without affecting validation outcomes.

It is essential to note that Business Rules can only be formulated within a Rulebook and must be part of the Business Rule sequence inside that Rulebook. A Rulebook can have just one Business Rule sequence, which is capable of forking any number of times.

6.2 Edit Version

While in a Live Version, you can exclusively view or read values, without the capability to alter, erase existing entities, or introduce new ones. To achieve any of the three actions, you must switch to the Edit Version mode.
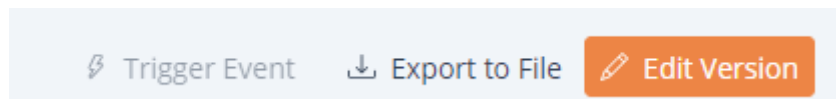
Figure 38: "Edit Version Button"

Now you can add, edit, and delete entities. Create your own Rulebook and Business Rules and test

your data against them. It is important to note, that the changes will not apply until they are published. If they are, the Edit Version will become the Live Version.

6.3 Knowledge Base

Knowledge Bases are entities responsible for containing comprehensive information about the structure of your data, the regulations it needs to adhere to, Execution metrics, and a variety of additional details. More information on how to create, configure and use Knowledge Bases can be found in the HEDDA.IO User Documentation.

6.4 Live Version

In HEDDA.IO, the initial step to utilize a Knowledge Base involves publishing it. As soon as it is published, it goes "Live".

Figure 39: "Publish Version Button"

This functionality serves a dual purpose, allowing the establishment of a version history as well. The primary page of the Knowledge Base is perpetually set in Read Only mode. This signifies that you can exclusively view or read values, without the capability to alter, erase existing entities, or introduce new ones.

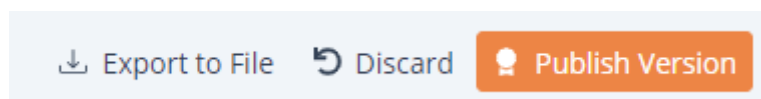To achieve any of the three actions, you must switch to the Edit Version mode, accessible by selecting Edit Version.

6.5 Mappings

In HEDDA.IO, Mapping is an operation that associates Domains with columns from a dataset. It ensures that the data within different Domains and columns are correctly aligned, enhancing the overall consistency and accuracy of the dataset.

6.6 Projects

Projects are essentially workspaces within HEDDA.IO, designed to cater to your distinct needs for utilizing the HEDDA.IO platform.

Unless you have a Project established, it is not possible to generate a Knowledge Base. As a result, initiating a Project stands as the primary step towards making use of HEDDA.IO.

The Project page encompasses several vital Project-wide features, including External Connections, Events, and User Management. Moreover, it presents the list of Knowledge Bases and provides the capability to add new ones.

6.7 Rulebook

To further enhance the validation process for the data that is ingested, HEDDA.IO offers a feature known as Rulebooks. This serves as an encapsulation for a collection of Business Rules, enabling you to organize and arrange them in a coherent order, based on basic logical rules.

Furthermore, it provides an overview and a visual representation of the sequence of Business Rules on a canvas, allowing you to easily visualize the flow and make modifications.

Conditions and subconditions can be incorporated into Business Rules, allowing the utilization of Lookup values or other Domain values to validate Domains.