A Tool to Synthetically Generate Data for Process Mining

Abstract
This document summarizes instructions how to use our tool to synthetically generate data for process mining.

1. Link to the tool

The tool presented in the paper can be found at https://deich.pa.informatik.uni-kiel.de/process-analytics-toolbox/.

2. Maturity of the tool

We refer to the Technology Readiness Level to classify the maturity of our tool (TRL, ISO 16290:2013). The technological concept of our tool has been published in [1], and the proof-of-concept was demonstrated at the CAiSE 2022 Forum. After the tool demonstration in 2022, we implemented most of the feedback that we received. Since that, we have been using the tool in various research projects to validate the tool and thus improve the proof-of-concept. Eventually, we rate the maturity of the tool as TRL level 4.

3. Link to a demonstration video

A video presenting how to use our tool can be found at https://github.com/Process-Analytics-Group/R-YZ-BPM-Demo-Submission/blob/main/docs/bpm2023-demo.mp4 within this repository.

4. Tutorial

This section introduces the components of our tool and how to use them: it allows you to generate your own process models, import and export them in industry standard formats, understand and choose simulation settings, and download the final event logs.

Proceedings of the Demonstration Resources Track, Best BPM Dissertation Award, and Doctoral Consortium at BPM 2023, BPM 2023, Utrecht, Netherlands, September 11-15, 2023

4.1. Login And Overview Of The Tool

The Process Mining Tool component is a web-based interface accessible by all browsers. The login page (see. Fig. 1) serves as the entry point to the web interface. To access the tool, please use the following account:

- Username: BPM2023TestUser
- Password: BPM2023TestUser

Figure 1: The login interface of our tool.

After successful login, you will be automatically directed to the overview page, as shown in Fig. 2. This page is the entry interface of the application, providing access to various tools and functionalities called widgets.

The interface is slit into distinct sections: (1) and (2) address the navigation of the page. Section (1) allows you to return to the widget overview, while by clicking on data (2), you can access the central repository of already created process models.

The widget overview is further divided into two main sections. The left-hand section, (3), comprises a collection of editors with the **Petri-Net Editor** (4) being particularly relevant for this use case. These editors offer a range of functionalities for modelling processes.

The right-hand part consists of the simulation widgets (5). These widgets allow you to perform simulations based on the saved process models. The **Event-Log Generator** widget (6) generates event logs using a Petri-net. However, please note that the same functionality as (6) can also be directly accessed and performed within the **Petri-net Editor** (4).

By clicking on the data button, you will be directed to the repository of Petri-nets and (IoT) location nets that have already been saved. For now, a Petri-net called *BPM 2023 Demo* has been created. If you want to add new Petri-net, you can use either the **Petri-net Editor** or

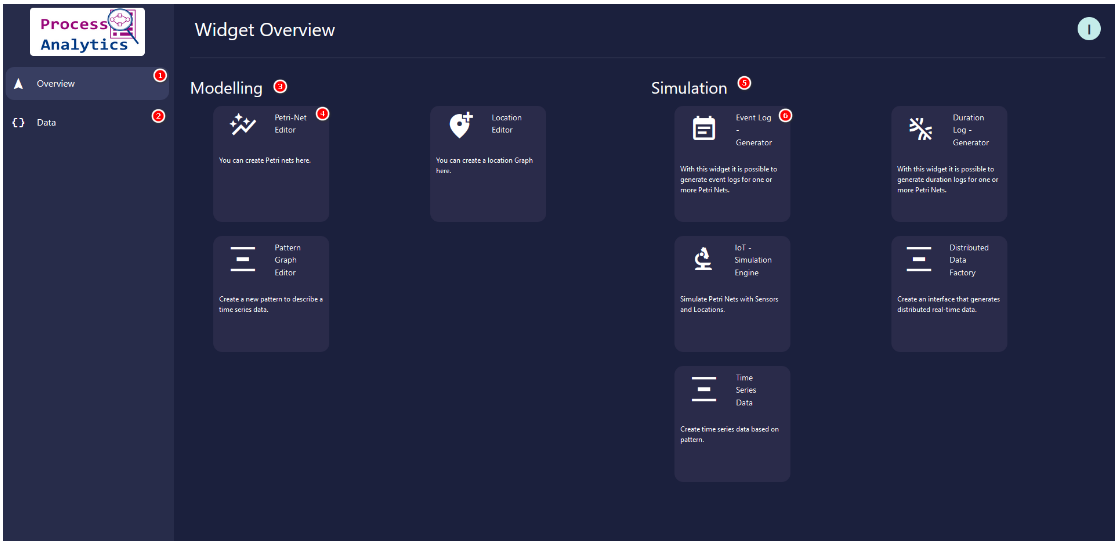

Figure 2: The widget overview page.

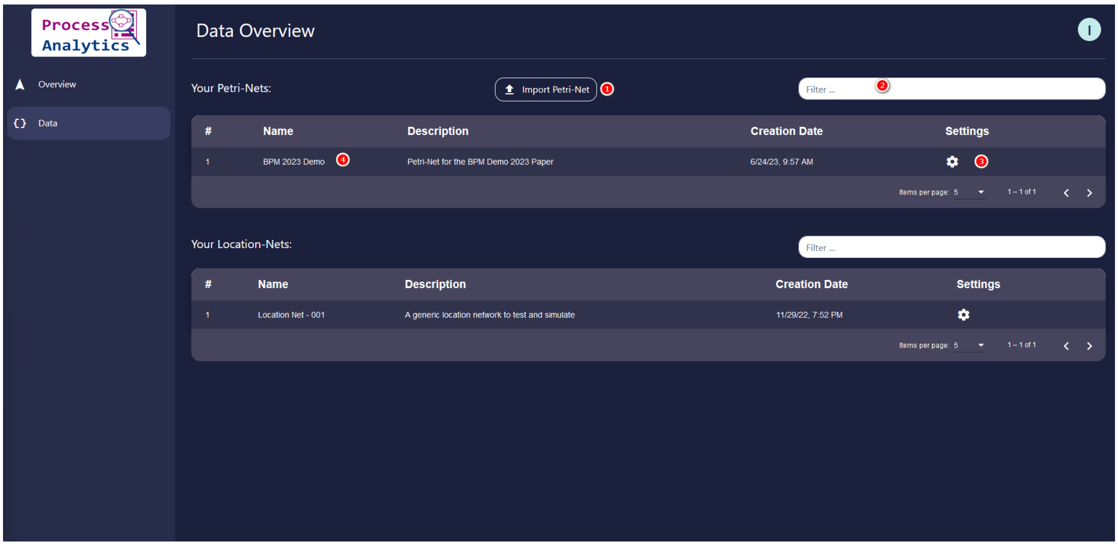

Figure 3: The data overview page.

import a process model in the .pnml-format by using the **Import Petri-net** button (1). For more information on importing, see Fig. 4.

For exporting, a choice of export formats can be accessed by clicking on the gear-button (3). For more information on exporting, see Fig. 5.

Next to the Import button, you will find the Filter text field (2). This text field provides a way to filter the list of Petri-nets, enabling you to display only specific elements based on your filtering criteria.

By clicking on a specific table entry (4), you will be taken to the **Petri-net Editor**, where you can make modifications, perform further analysis, or continue working with the selected Petri-net (see. Fig. 6).

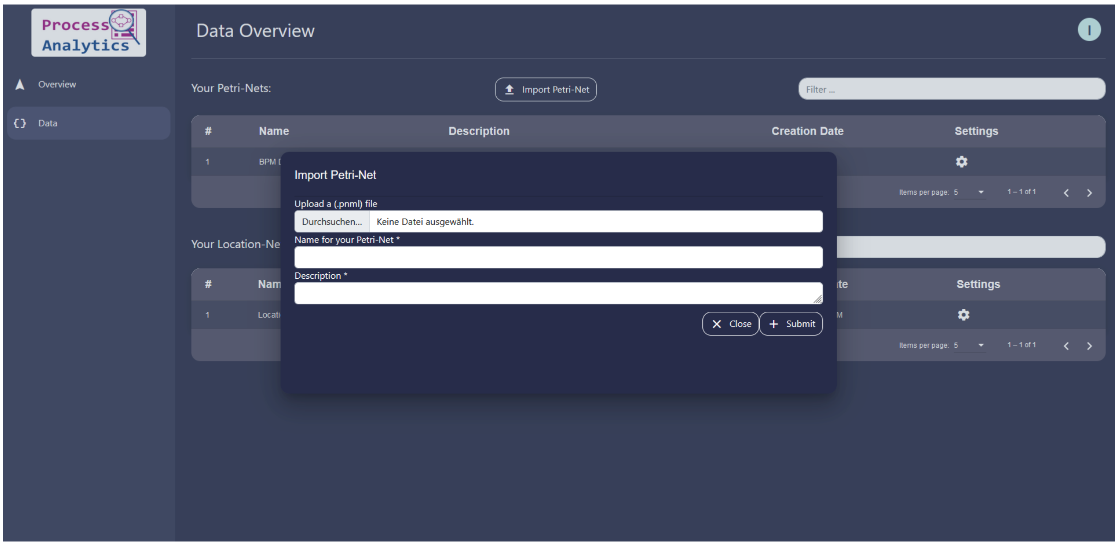

Figure 4: Petri-net import dialog.

When importing Petri-nets in the Petri-Net Markup Language (PNML) format, it is required to specify a name and a description for the Petri-net (see Fig. 4).

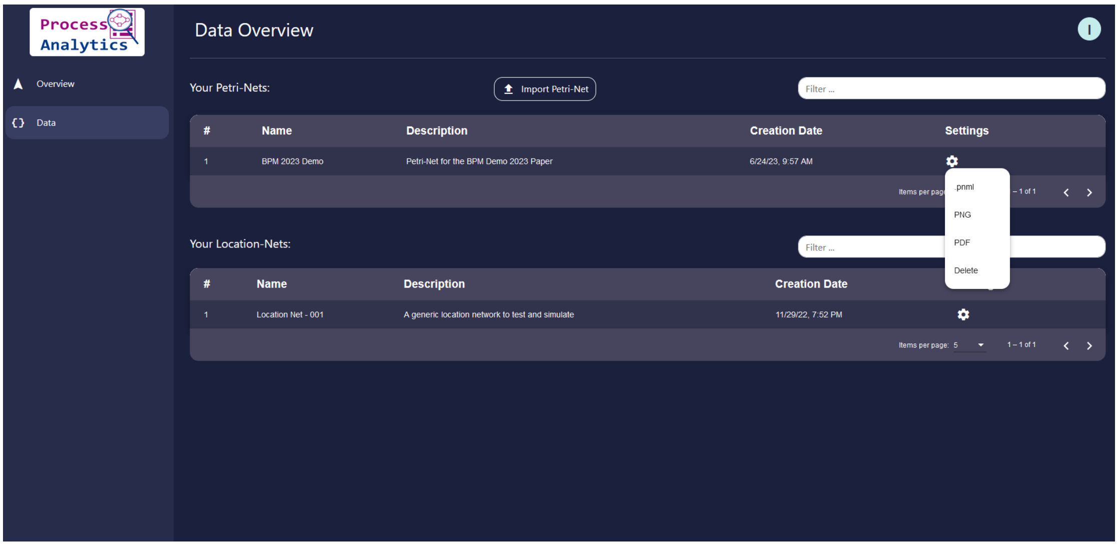

Figure 5: Petri-net export dialog.

At present, you have the capability to export the Petri-net in various formats, including PNML, PNG, or PDF files. Notably, both PNG and PDF exports offer customization options for adjusting the color scheme of the Petri-net.

4.2. Creating And Loading a Petri-Net

This section describes how to model processes with the **Petri-net Editor**. It provides information and guidance on using this editor to effectively build and design processes.

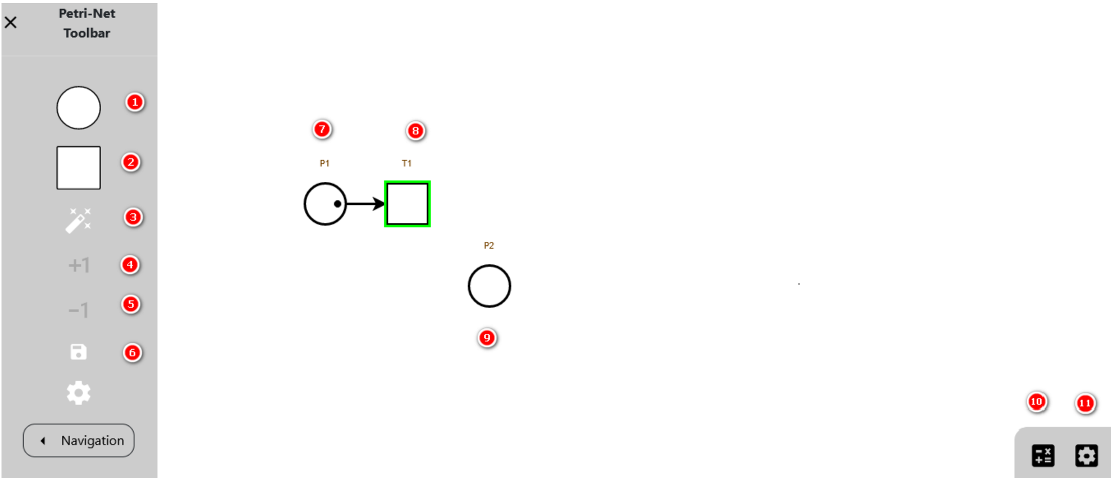

Figure 6: Initially Modeling a Petri-net with the Petri-Net Editor.

The editor interface is divided into two main areas. The left-hand sidebar contains editing elements and can be collapsed and expanded by clicking on the (x) symbol. The main screen represents the editor workspace. Within the editor workspace, you can interact with various elements. (1) represents a place, which can be dragged and dropped into the editor. Similarly, (2) represents a transition that can also be dragged and dropped into the editor.

During the process of modeling, if the graph starts to appear unclear or cluttered, you can utilize the layout algorithm provided by (3) to rearrange the elements into a more organized structure. Once you have selected a place in the editor, you have the ability to increase the number of marking using (4) or decrease them using (5).

To save the Petri-net to the database, you can use (6). The save icon will appear 'active' if there are unsaved changes or 'inactive' (grayed out) if there are no pending modifications. Additionally, (7) and (8) represent instances of a place and a transition. To connect the transition T1 (8) with the place P2 (9), hover the mouse pointer over T1 until a green frame appears. Then, using drag and drop, you can draw an edge connecting T1 to P2. It is important to note that two elements of the same type cannot be connected.

When you select an element in the editor, additional icons will appear on the element itself. By clicking on the four arcs, a new element is placed in the direction indicated by the arc, automatically inserting a connection with the previous element. Alternatively, you can use the plus symbol in the top right corner to access and configure specific properties of the selected element.

Both (10) and (11) allow to simulate the process. Through (10), you can start the simulation by clicking on it. After the simulation, a dialog will automatically open, presenting the results. Please note that only saved changes are considered for simulation. Unsaved modifications are not taken into account during the simulation.

Finally, (11) provides additional functionality by opening a dialog where you can configure various simulation options. This dialog allows you to make specific adjustments and settings tailored to your simulation requirements.

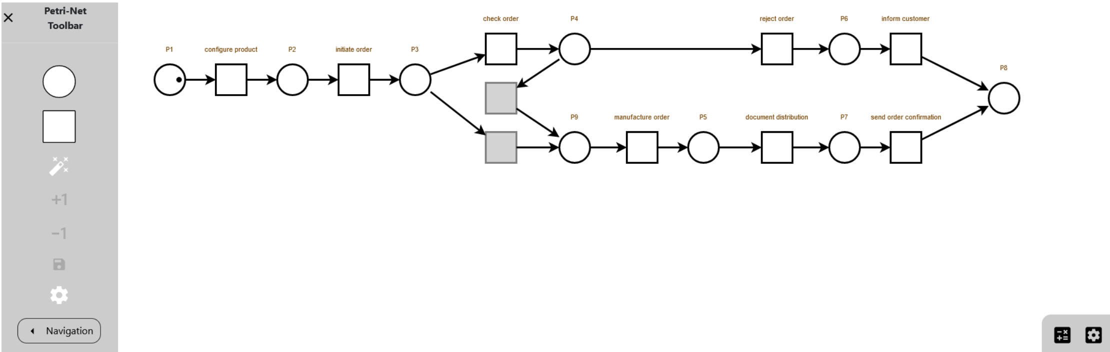

Figure 7: A complete model of the Petri-net.

In the example shown in Fig. 7, our paper's use case was modeled. This process comprises a total of 9 places and 10 transitions, where two of the transitions are silent. This means they do not represent activities in the modeled process, but are required for for syntactically correctness. So far, we assume that the Petri-net is a workflow net.

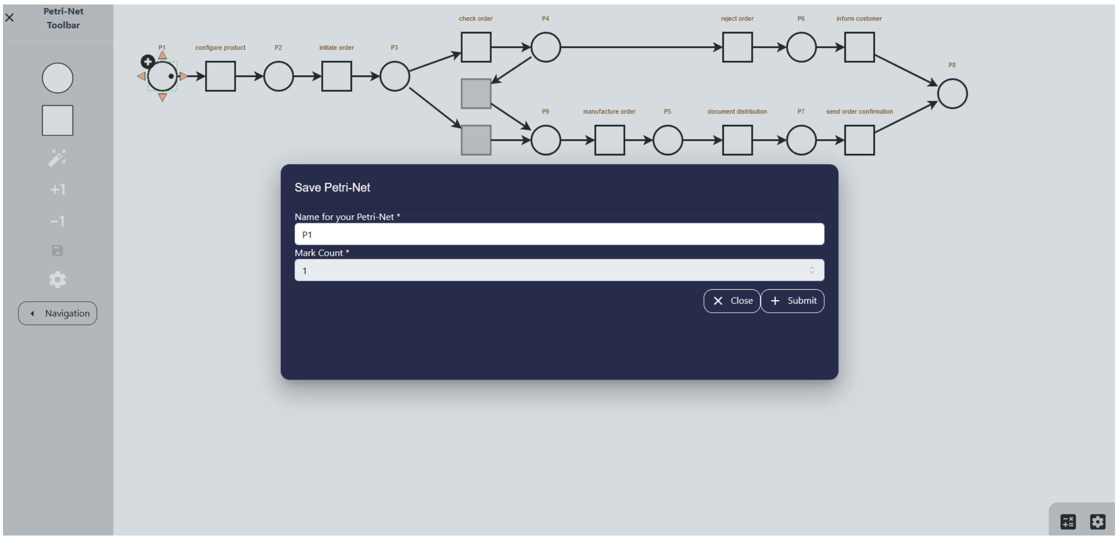

Figure 8: Settings of a place.

You can rename places by selecting a place and clicking the plus symbol on it. Additionally, you can also view and change the number of marking associated with the selected place.

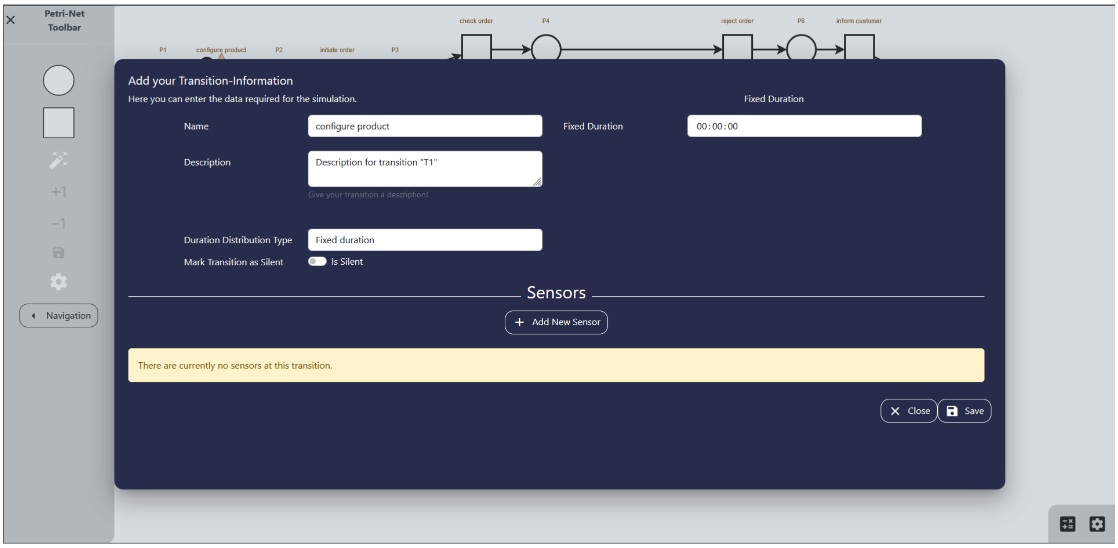

Figure 9: Settings of a transition.

A transition has more setting options than a place. In addition to the name and description fields, also the duration of the transition can be specified. For this, you can select a **Duration Distribution Type** from the options like **Fixed duration**, **Gaussian distributed duration**, and **Duration distributed as equipartition**. Once you make a selection from the drop-down field, the input options on the right-hand side will dynamically change to correspond to the chosen duration distribution type. In an IoT environment, you can enhance transitions by adding sensors to them.

4.3. Simulation Settings

To modify the simulation parameters, you can use the provided form. This form comprises all the options for configuring the simulation and displaying the resulting output. Within this form, you can select various parameters such as the **Use Case** or the **Execution Amount** among others.

To discard any changes made and close the dialog, click on the **Close** button. Clicking the **Submit** button will save the entered data to the database and set the settings as active.

4.4. Define a Sub-Process For The Simulation

It is possible to simulate only a subset of your process model (a sub-process).

To create a sub-process, CTRL-click a place in the Petri-net as your starting point. CTRL-click another place as your end point. The tool will then determine the shortest path between these two places and highlight the resulting sub-process by coloring the arcs blue. If no sub-process is selected and you click the simulation button, the full process will be simulated. However, if a sub-process is selected, only the sub-process will be simulated.

Figure 10: The settings that are needed for the simulation of the event logs.

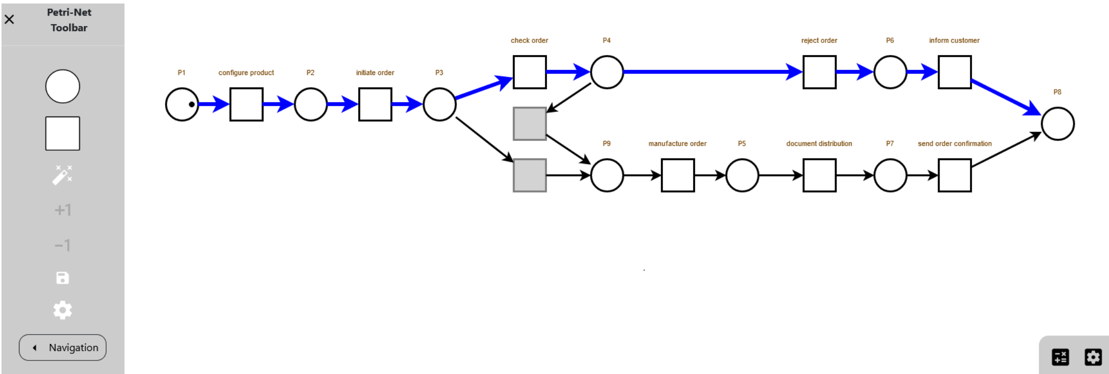

Figure 11: A selected sub-process that can be used for the simulation.

It is important to note that a sub-process must always be defined between two places. Selecting a transition as the end of the sub-path is not possible. This decision was made according to the definition of a workflow net.

4.5. Simulation Results

This section focuses on the simulation results of the process. Once you click the simulation button and the back end successfully computed the results, the simulation results dialog will automatically open. This dialog shows the outcome of the simulation, allowing you to analyze and download the results.

In the simulation results dialog, the generated event logs are represented in a table format. The table includes several fields, namely:

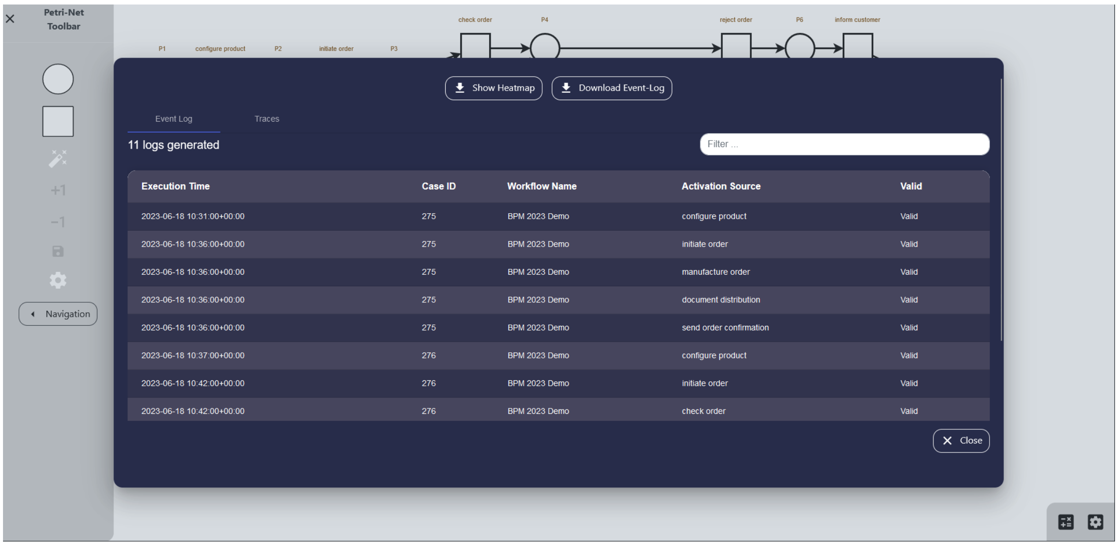

Figure 12: Dialog of the synthetically generated event logs.

Execution Time: This field indicates the duration of the corresponding transition, representing how long it took to complete.

Case ID: The Case ID field allows you to differentiate between different cases in the simulation. Since it is possible to simulate a workflow multiple times, each trace can be identified by its unique Case ID.

Workflow Name: The Workflow Name field specifies the name of the workflow associated with the event log entry. This field is particularly useful when simulating multiple workflows simultaneously.

Activation Source: The Activation Source field displays the name of the transition that was triggered at that particular time, providing information about the origin of the event.

Valid: As noise can be incorporated during the simulation based on the chosen settings, the Valid field helps distinguish correct entries from noisy ones. In case of noise, the Valid field will have the value "Noise" assigned to indicate its nature.

You can use the filter textbox to refine the displayed entries based on specific criteria. For example, you can filter for valid entries only or restrict the display only to entries from a specific workflow.

Additionally, there is a **Show Heatmap** button available, which allows you to visualize the frequency at which different transitions were fired, providing insights into their occurrence patterns.

Once you have generated an event log, you can download it using the **Download Event Log** button. This button provides the functionality to save the event log in either CSV (Comma-Separated Values) or XES (eXtensible Event Stream) format.

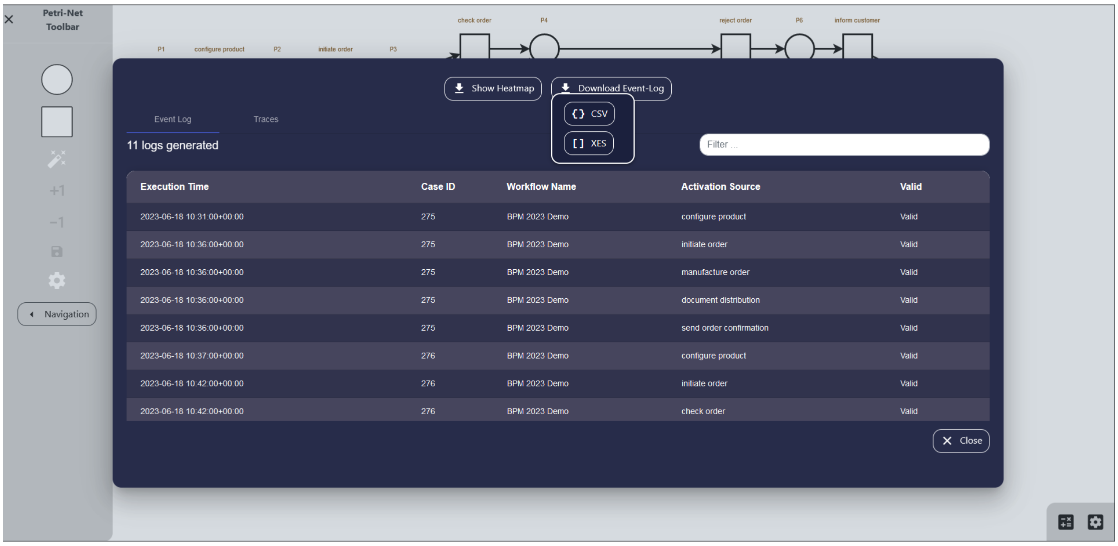

Figure 13: Export options of the event log.

References

[1] Y. Zisgen, D. Janssen, A. Koschmider, Generating Synthetic Sensor Event Logs for Process Mining, in: J. De Weerdt, A. Polyvyanyy (Eds.), Intelligent Information Systems, volume 452, Springer International Publishing, Cham, 2022, pp. 130–137. doi:10.1007/978-3-031-07481-3_15.