Explori
User manual

Quickstart

1. Clone the Git repository.
   `git clone <url>`
2. Change into project directory.
   `cd <project_directory>`
3. Start Explori application (this may take a while the first time). If you are using Windows, you will need to start a git shell or WSL to be able to execute it.
   `./app.sh --start`
4. Open the Explori webinterface at http://localhost
5. Stop Explori application
   `./app.sh --stop`
6. Remove Explori application (everything!)
   `./app.sh --remove`

Mission statement

Explori is a tool meant to make extracting valuable information from Object Centric Event Logs (OCEL) as easy as possible. It displays easily interpretable Directly Follows Multigraphs (DFM) and enhances them with alignment information and performance metrics. It also allows fast and interactive filtering to reduce the complexity of discovered models.

Managing your OCELs

On starting Explori for the first time, you would be navigated to the sessions page. If you have used Explori before then it will automatically load the previous session, i.e. the last used OCEL (along with the saved settings and computed results) and display the discovered DFM.

On the sessions page you can view the available OCELs, delete the available OCELs, upload a new OCEL, select the OCEL and start the new session, or preview an OCEL uploaded as Comma-Separated Value (CSV) file. A demo OCEL is provided with Explori to start exploring.

Uploading and selecting an OCEL
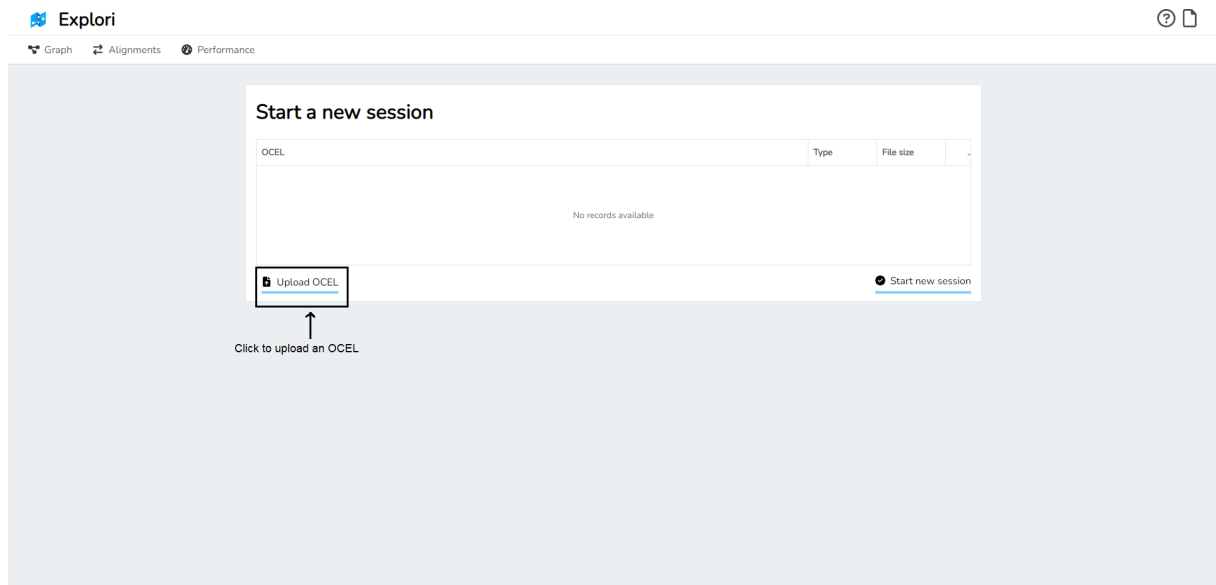
To upload an OCEl, click on the 'Upload OCEL" button
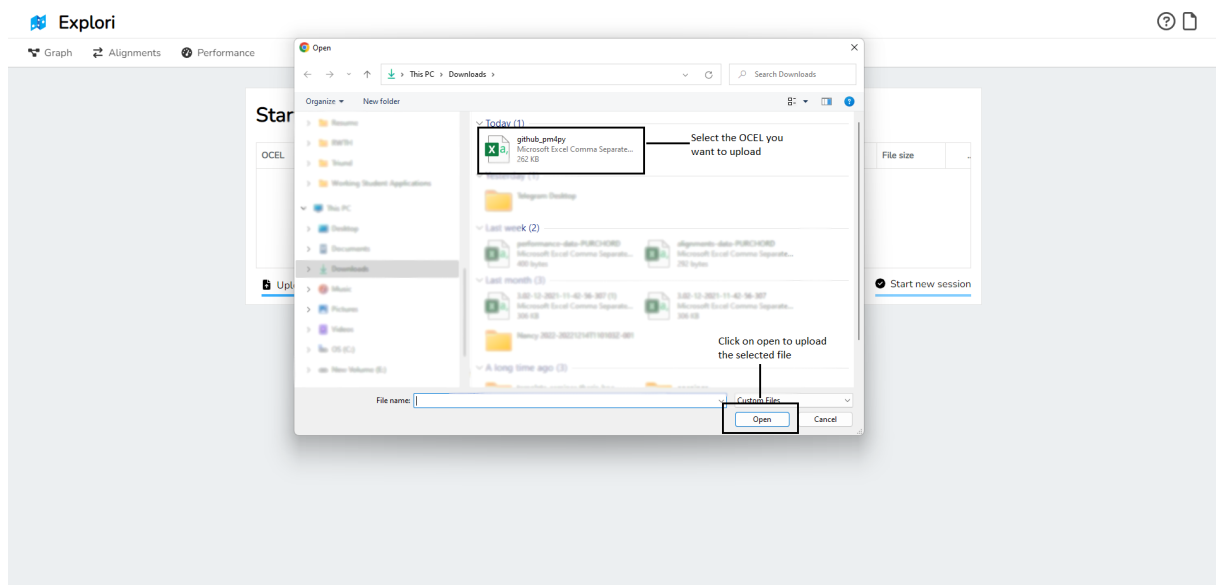
Select the desired OCEL using your file explorer
and click on "Open" to complete uploading

Available OCELs

List of all the available OCELs and starting new session with any OCEL

Deleting the OCEL

By pressing the bin icon next to an OCEL, it can be deleted. This will open up a prompt to confirm that you are sure about deleting the OCEL. Once you delete an OCEL, Explori deletes all the information associated with it ,i.e., the discovered DFM, alignments and performance metrics. So uploading this OCEL again will start the discovery process and conformance checking again.

For mounted OCELs, the deletion option does not exist. Instead, we provide a button with an eraser icon to clear the cache which then in turn also allows for a re-discovery. Mounted OCELs can also be identified by the symbol in front of the name which is either a cloud for uploaded OCELs or a hard drive for mounted OCELs.

Example for a mounted OCEL.

Selecting OCEL to start the session

Once the new OCEL is selected, Explori stores this information and starts process discovery, computes alignments and performance metrics. On the next start it will use this newly selected OCEL as default and load all the associated metrics with it.

Previewing a CSV OCEL and configuring settings

CSV Settings

Preview of the selected CSV OCEL

id	ocel:timestamp	ocel:activity...	ocel:type:case:concept:name	ocel:type:org:resourc...	ocel:type:case:repo
0	2018-08-24 12:03:19+00:00	created	['https://api.github.com/repos/pm4py/pm4py-core/issues/1/timeline']	['Javert899']	['pm4py/pm4py-core']
8	2018-09-03 15:29:33+00:00	created	['https://api.github.com/repos/pm4py/pm4py-core/issues/2/timeline']	['fmannhardt']	['pm4py/pm4py-core']
9	2018-09-12 07:15:13+00:00	commented	['https://api.github.com/repos/pm4py/pm4py-core/issues/2/timeline']	['s-j-v-zelst']	['pm4py/pm4py-core']
13	2018-09-13 11:42:48+00:00	created	['https://api.github.com/repos/pm4py/pm4py-core/issues/3/timeline']	['edugonza']	['pm4py/pm4py-core']

Objects
ocel:type:case:concept:name, ocel:type:org:resource, ocel:type:case:repo ———— Set the object fields from CSV

Activity
ocel:activity ———— Set the activity field from CSV

Timestamp
ocel:timestamp ———— Set the timestamp field from CSV

Id
id ———— Select the ID field from CSV

Seperator
, ———— Select the separator

After selecting a CSV OCEL from the available OCEL, it can be previewed and different settings as shown above can be configured

It might be the case that some fields are already set by default. This is because Explori tries to predict which columns belong to which field. Since this is not perfect and the selection might be wrong, double-check for correctness or just change the fields to your preference.

Mounting your OCELs into Explori

If your OCEL is particularly big, uploading it can take a lot of time and will lead to a redundant copy of the OCEL. This can be avoided by directly mounting the OCEL into Docker. To do so, you need to

1. Simply start the Explori application as mentioned in the Quickstart guide and additionally provide a path to a folder, containing event logs, which should be mounted into the application.
   ```
   ./app.sh --start path/to/event/logs
   ```

The mounted OCELs should now be listed on the "new session" page.

Navigating the DFM

Once you started a new session, you will be met with the following screen:
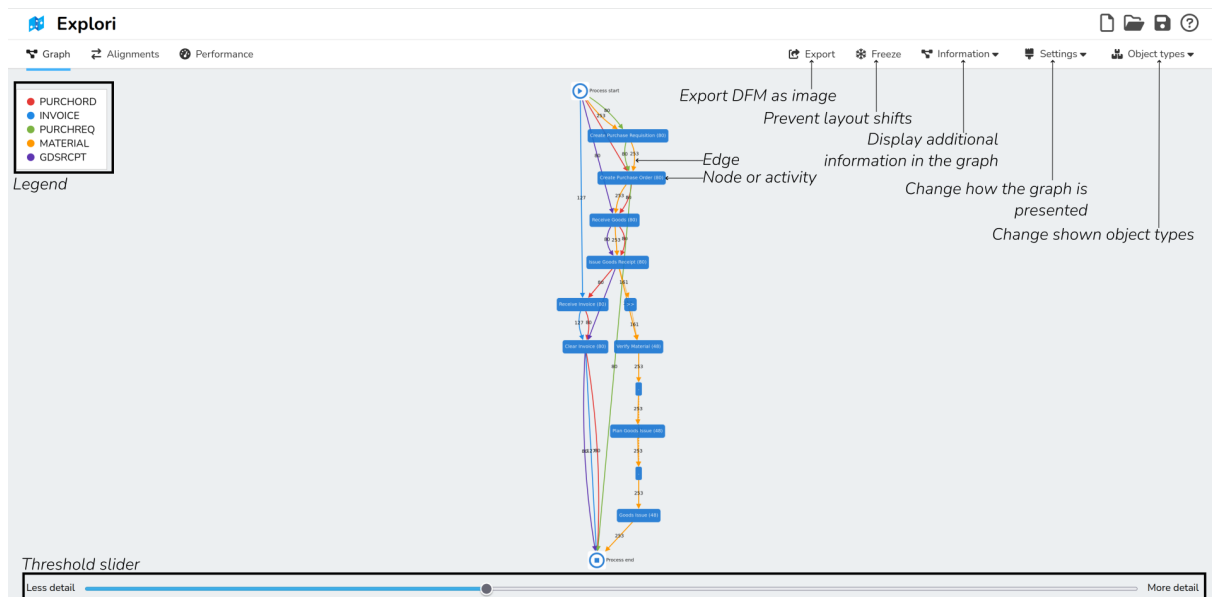
The main component of the screen is the graph presented in the middle. Each node is an activity that occurred in the event log, while the edges show the order in which those activities were performed. For example, the node "Create Purchase Requisition (80)" shows that the activity "Create Purchase Requisition" has been executed 80 times. Similarly, the node "Create Purchase Order (80)" indicates that the activity "Create Purchase Order" was executed 80 times. The green edge between those nodes shows that 80 times, a PURCHREQ (read "purchase requisition") object was first involved in a "Create Purchase Requisition" activity and a "Create Purchase Order" activity directly afterwards. Analogously, 414 material objects were first involved in the creation of a purchase requisition, which resulted in a purchase order being created afterwards. If you have used Directly Follows Graphs before, it might seem very strange to you that the activities were performed 80 times, but the "MATERIAL-edge" shows 414 transitions. This is because, on average, 5.2 MATERIAL objects were involved in both activities.

As shown in the example, there can be multiple edges between two nodes. That means that there are multiple object types that were involved in both activities. The object types are indicated by the color of the edges. The legend in the top left corner shows you which color belongs to which object type.

You can also pan the graph around by holding the left mouse button and dragging the mouse, as well as zoom in and out using the scroll wheel. Furthermore, nodes can be dragged-and-dropped into new positions. However, the node positions are recalculated each time the graph is updated, which might be when the filtering is changed, when

background calculations were updated, or even when you click a node (more on that later). Therefore, it is advised that you enable the "freeze" mode which disables layout calculations by clicking the "freeze" button.

Filtering the graph

In many cases, an OCEL will contain both the "usual" process, and a lot of outlier cases that make the model very complex. This is somewhat related to the pareto principle, so in layman terms, you can assume that roughly 20% of the log causes 80% of the complexities in the model.

Therefore, it can be helpful to filter the graph to only show the most frequent behavior. This can easily be done in Explori by changing the "threshold slider" at the bottom of the screen. The threshold is equivalent to the proportion of objects that are shown in the model, so a threshold of 70% means that at least 70% of the objects contained in the log are included in the DFM.

Furthermore, you might only be interested in a subset of object types. In that case, you can click the "Object Types" button in the top right corner. A dropdown menu will open where you can (un-)select object types. Note that at least one object type has to be selected at each point in time.
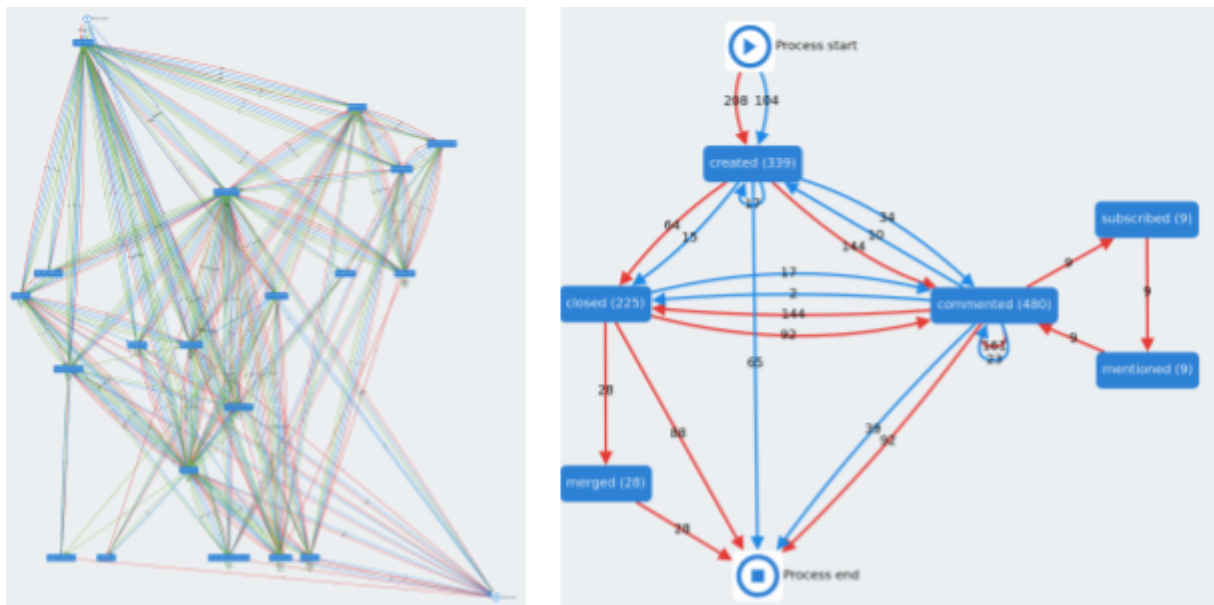
Both models are produced from the same OCEL, the only difference being that the model on the right has a lower threshold and fewer object types.

The Infobox

If you already have played around with Explori, you might already have noticed that you can click on edges and nodes. As a result, the so-called infobox will appear. It contains

information performance metrics, which will be covered in more detail in their dedicated section, as well as all traces going through that node or edge, separated into the traces that are shown (included) in the model and those that are "filtered out" of the DFM.

If you wish to do so, you can disable the infobox in the settings menu in the navbar.

Displaying further information in the graph

The information dropdown in the navbar also allows you to present more information in the graph itself. It contains the groups, regarding alignments, edge labels and the highlighting mode.

The alignment settings will be addressed in the following section which is dedicated to alignments. The edge labels settings allows you to switch between showing the number of objects traversing along the edge, or the minimum, mean, maximum, or total waiting time spent on that edge. The waiting time will be described in further detail in the section of performance metrics.

Finally, the highlighting modes change the widths of the edges depending on the selected criteria, which are relatively analogous to the edge labels. The biggest difference is that we also offer a logarithmic scale for some criteria. Whether it is more helpful than the linear scale depends on the log and personal preferences, so our suggestion is just to try both scales for each event log and use the one you like most.

A DFM with mean waiting time as edge labels and highlighting mode.

Alignments
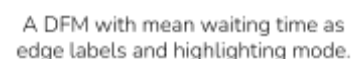
When lowering the threshold, objects start being filtered out as explained above. This leads to cases where the model is not able to fully replay the behavior of the log anymore. To be able to show which behavior specifically was filtered out, alignments are calculated using the threshold and the initially used OCEL.

These calculated alignments can be viewed in two ways:

The first one is to open up the "Information" dropdown and select the type of visualization you want to have for the alignments. "None" is the default and shows no alignments at all. "Simple" will show one edge per log move and one per model move as can be seen in the left picture below. The dotted edge represents a log move which means that in the

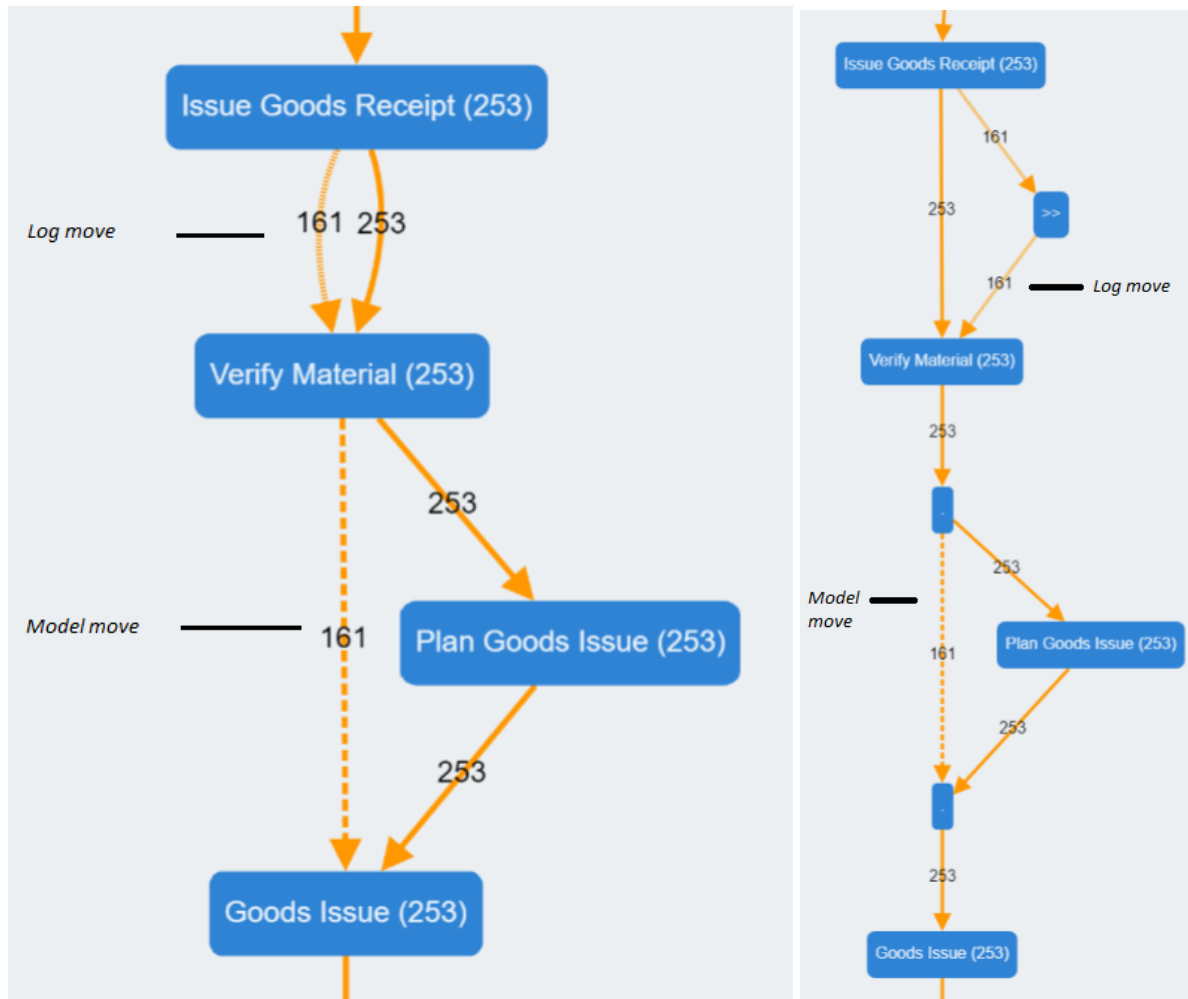original log, an activity is happening which the model is not able to replay because nodes or edges are missing. The dashed edge represents a model move which means that in the model an activity is executed which is not present in the log. "Extended" will show two dotted edges and one skip node ">>" per log move to represent that something is happening in the log which is not possible in the model and one edge and one dashed edge and two nodes per model move to show that an activity or activities are skipped in the log. An example can be seen below on the right side.
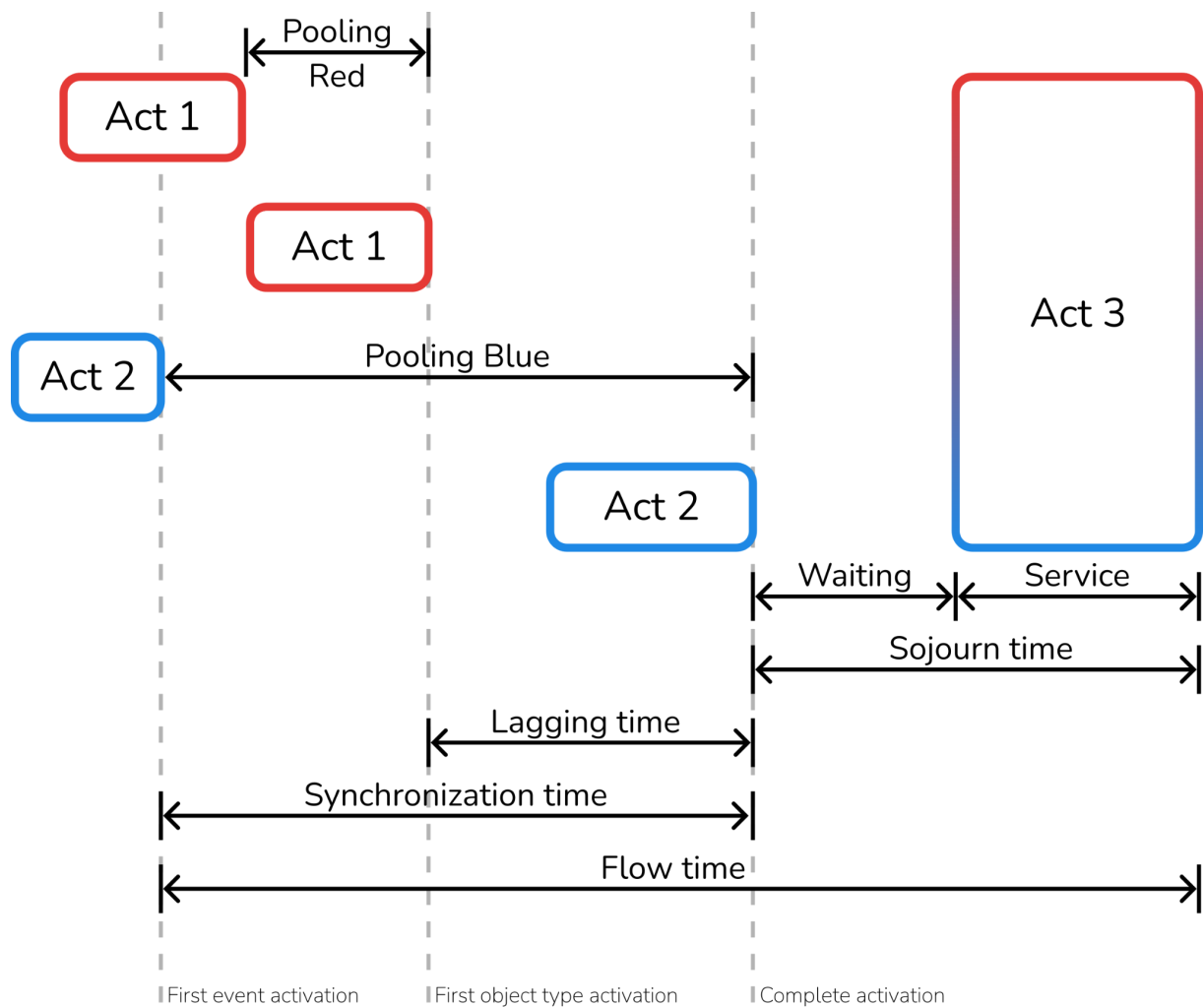

On the left side, the "simple" mode is selected and on the right side, the "extended" mode is selected.

The other option to view alignments can be found under the "Alignments" tab right of the "Graph" tab where the DFM can be viewed. Here, a table for each object type is shown in which the alignments are presented.

MATERIAL								🡅 Export
Trace 0	[Log Alignment]	|EXPLORI_START|	Create Purchase Re...	Create Purchase Or...	Receive Goods	Issue Goods Receipt	Plan Goods
	[Model Alignment]	|EXPLORI_START|	Create Purchase Re...	Create Purchase Or...	Receive Goods	Issue Goods Receipt	>>
Trace 1	[Log Alignment]	|EXPLORI_START|	Create Purchase Re...	Create Purchase Or...	Receive Goods	Issue Goods Receipt	Verify Mate
	[Model Alignment]	|EXPLORI_START|	Create Purchase Re...	Create Purchase Or...	Receive Goods	Issue Goods Receipt	Verify Mate

I< ‹ Page 1 ⏶⏷ of 1 › >I | Results per page 50 ▼ Showing 1 - 4 of 4

Alignment table for the object type "Material" which belongs
to the alignments shown in the graph above.

There are three options for pairs of activities in this table. Either log and model activity are the same which means log and model are aligned, log has an activity and model has skip activity (">>") which is called log move and log has skip activity and model has an activity which is called model move.

This table does also provide export functionality (found in the top right corner of each table) which will download the alignment data as a json file in the same format that the backend returns. The exported data can then be used for further processing.

Performance Metrics

Explori also calculates OCEL-specific performance metrics that can be helpful to identify bottlenecks and other issues in your processes. As shown in the figure below, those performance metrics "anchor" to five different timepoints, those being the time an event start and ended, the time all conditions are met for the event to be executed (also known as the activation time of the event), the first time all preceding event was finished (here described as the "first event activation") and the first time all preceding events of one object type finished (here described as the "first object type activation"). Those timepoints are then used to calculate the performance metrics according to the figure on the next page.

Explori also calculates pooling and waiting times for edges. Those are defined as the respective performance metrics of the target activity if the log was filtered to only contain objects that transition along a given edge.

Explori automatically starts calculating the performance metrics when starting a session and on every change of the filtering threshold. These calculations can take a few seconds, depending on the size of the event log. Therefore Explori will automatically update shown information as soon as the performance metrics are available. These metrics are displayed for each node and edge, and can be viewed by clicking on the "Performance" tab on the top navigation bar. Performance metrics for alignment edges are not shown since the required timestamp cannot be determined for such moves.

The performance tab displays three tables as described:
- Node metrics: It lists the performance metrics for every node. Different metrics can be enabled and the corresponding aggregators can be selected for efficient comparisons.
- Pooling times: It lists the Pooling time for each a node and an object,
- Edge metrics: Similar to node metrics, this table lists the performance metrics of every edge. Different metrics can be enabled and the corresponding aggregators can be selected for efficient comparisons.

Node metrics for an OCEL

This table also provides export functionality (found in the top right corner of each table) which will download the performance data as a json file in the same format that the backend returns. The exported data can then be used for further processing.

These performance metrics are also available in infobox in the Graph tab. To highlight the edges of the DFM with the, you can select a metric from information dropdown and the desired aggregator.

Performance metrics in infobox and the drop to highlight edge with selected metric

Help

Explori also offers a help page which is useful if Explori broke and needs to be fixed or you just need to access the backend docs. For these purposes, the help page provides several buttons which allow you to clear the whole cache, reset queries and reset the local storage. It can be accessed by clicking on the "Question mark" button in the top right of the navbar.
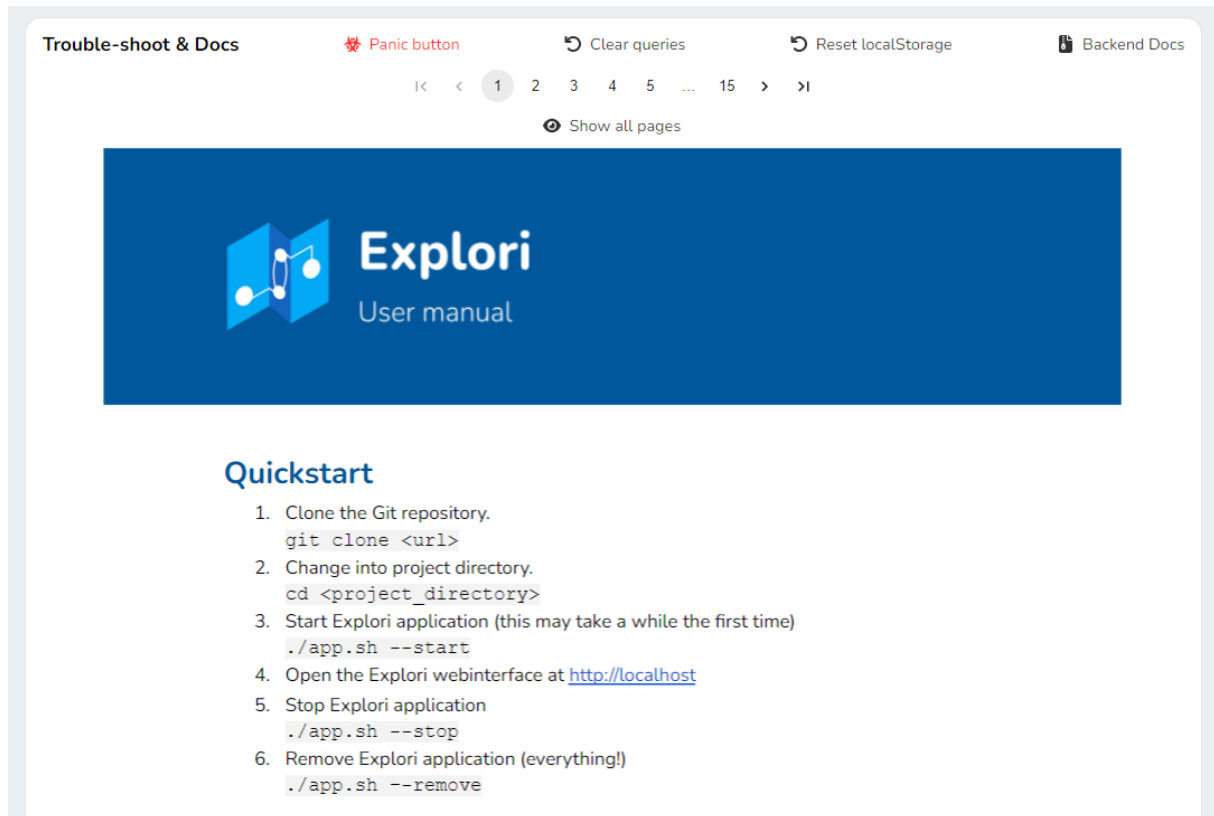
Tooltips

We put tooltips everywhere we thought of, so when you don't know what a button does, hover over it and you will probably get a tooltip which shortly informs you about the function of the button.

Panic button

The panic button is useful when Explori is stuck at executing tasks, keeps throwing errors or is simply broken. When clicking this button, a prompt will pop up to inform you about what the button will do and if you are sure about it. In this case, the whole hard-saved cache will be cleared, including the DFM, alignment, performance metrics, autosaves and if you used csv files, also saved csv column mappings.

Clear queries

The "Clear queries" button resets the DFM, alignments and performance queries to their initial state. This is helpful if Explori did run into an error loop to potentially fix that.

Reset local storage

The "reset local storage" button is useful if you loaded an OCEL which is somehow broken and Explori keeps executing tasks which throw errors because they can't be computed. It resets the local storage item "explori-currentOcel" to the default demo OCEL. Once you go back to the Home page and refresh the page, the demo OCEL should be displayed again. While this does "reset" the frontend part, it normally does not suffice to stop the continuous task queries which are sent to the backend.

Backend docs

The "backend docs" button leads you to the backend swagger interface of FastAPI which lists all endpoints and shows an example of how to use each of them.

Docs

The Help page does also display this user manual as a pdf which is split by pages. At the top, some pagination options exist to quickly traverse through the pages. Additionally, the "Show all pages" button allows the user to see all pages at once and just scroll down to see them all like in a normal pdf file.

Manual Sessions

While Explori automatically creates and updates autosaves for each OCEL you loaded, it also supports the use of manual sessions. These are sessions which you need to manually create and where you can have multiple sessions for one OCEL which lets you quickly switch between two or more states.
The buttons to use manual sessions can be easily found in the navbar on the right side.

The left button with the "Open folder" icon is used to restore sessions and the right button with the "Save" icon is used to store sessions.

Storing sessions

If you want to store manual sessions, navigate to the "Save" button in the navbar. On the page, a table is displayed with all information which will be saved when you decide to store the current session. At the top, an input field is shown which is used to decide on the name of the session. Per default, this will be the name of the currently used OCEL. Be aware that an already stored session with the same name will be overwritten. Next to the input, there is the button to store the session. Once pressed, the session will be stored and you will be redirected to the Home page.

Stored sessions are saved on the disk, so will also be available after shutting Explori down.

Key	Value
OCEL	uploaded/demo_ocel.jsonocel
Threshold	75
Selected object types	["PURCHORD","INVOICE","PURCHREQ","MATERIAL","GDSRCPT"]
Graph orientation	Top to down
Highlighting mode	none
Legend position	top-left
Alignment mode	none
Edge labels (Metric)	count
Edge labels (Aggregation)	mean

Name: demo_ocel 💾 Store Session

What will be saved:

The "Store Session" page with information which will be saved, name input field and button to store session.

In the following, more information about the information displayed in the table:

● OCEL:

- The name of the OCEL which you are currently using, with prefix to differentiate between uploaded and mounted OCELs.
- Threshold:
  - The currently selected threshold to filter the DFM in a range of 0-100 (ref to explanation above).
- Selected object types:
  - List of object types you currently have selected.
- Graph orientation:
  - Corresponds to the graph direction setting found under the "Settings" dropdown. Either "Top to Down" (default) or "Left to Right".
- Highlighting mode:
  - Corresponds to the highlighting setting found under the "Information" dropdown. Per default "None".
- Legend position:
  - Corresponds to the legend position setting found under the "Settings" dropdown. Either "None" or one of the four corners, per default "top-left".
- Alignment mode:
  - Corresponds to the show alignments setting found under the "Information" dropdown. Either "None" (default), "Simple" or "Extended".
- Edge labels (Metric):
  - Corresponds to the edge metrics setting found under the "Information" dropdown. Per default "Counts".
- Edge label (Aggregate):
  - Corresponds to the edge metric aggregation setting found under the "Information "dropdown. If "Counts" is selected as the metric, then "sum". Per default for the other two options "mean".

Restoring sessions

If you want to restore a session, head to the "Download" button in the navbar. On the page, you will see a table containing all saved sessions, whether they are autosaves or manual saves.

Session name	Last change date	Used OCEL	Th...	Object types	Alignm...	Metric...	Aggreg...	
p2p-normal	2/3/2023, 12:30:38 PM	p2p-normal	75	PURCHORD,INVOICE	simple	waiting_t	max	🗑
autosave-fa13da45-527a-5572-b351-0d6674ebfd2	2/3/2023, 12:30:31 PM	p2p-normal	75	PURCHORD,INVOICE	simple	waiting_t	max	🗑
autosave-5958420c-96cc-58b5-9cd8-3a7b254e7ff9	2/3/2023, 12:28:51 PM	demo_ocel	75	PURCHORD,INVOICE	none	count	mean	🗑

⬇ Restore Session

An example of the "Restore session" page. The data is sorted by "Last change date", so the entry at top will always be your currently used OCEL.

To restore a session, select the session which you want to restore and press the "Restore Session" button at the bottom center. This will then fetch the saved session and redirect you to the Home page.

It is also possible to delete sessions by clicking on the button with the bin icon for the corresponding entry. While it is possible to delete autosaves, we generally don't recommend it unless you know what you are doing. On pressing the delete button, a prompt will pop up with more information about what will be deleted and the consequences of it.

Development

Setting up the development environment

The source code can be found on GitHub (). The README in the project contains detailed explanations on how to set up your own development environment to add your own features or change already existing ones. Furthermore, if you are just interested in using Explori and not development, we suggest that you use the Docker deployment method.

Backend documentation

The Explori Backend has a Swagger interface that allows you to interact with it without the need to write code. You can access it at http://localhost:8080/docs.