RepGraph

User Manual

1. Introduction

This tool will aid in the visualisation and analysis of meaning representation graphs (MRG) produced according to one of the following frameworks DMRS, EDS, PTG, UCCA, and AMR. These MRG's need to be in JSONLines format as described in MRP.

There are four visualisation formats:

- Hierarchical - The MRG is constructed focusing on node spans.
- Tree - The MRG is constructed focusing on directed edges.
- Flat - The MRG is constructed focusing on token order.
- Planar - The MRG is constructed in a similar manner as Flat but highlights the planarity of a graph.

Furthermore, the analysis tools available include:

- Displaying subsets based on adjacent or descendent nodes.
- Searching for graphs that contain a specific sub-graph pattern.
- Searching for graphs that contain a specific set of nodes.
- Comparing two graphs based on their nodes and edges.
- Finding the longest directed or undirected path.
- Determining if the graph is connected.
- Determining whether the graph is planar.
- Overall Dataset Statistics

This user manual will give you the necessary knowledge to use this tool efficiently and correctly.

2. Getting Started

The User Interface has been created to allow for a quick and efficient way to make use of the visualisations and analysis tools offered.

The following is a step-by-step guide on how to start using the tool and a general overview of the many components present.

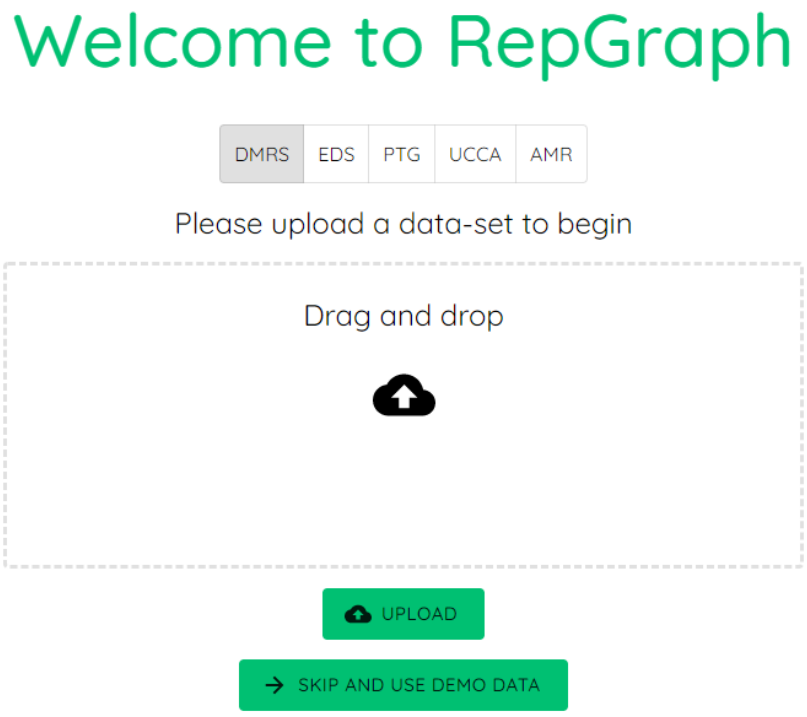
2.1. Upload a File

The first step to using the tool is to upload a data-set of your choosing. This data-set needs to follow a DMRS format and each meaning representation graph in the data file needs to be on a new line.

An example of a parsed EDS format data is as follows:

{"id": "20001001", "flavor": 1, "framework": "eds", "version": 1.0, "time": "2020-04-28", "source": "wsj", "provenance": "Redwoods Ninth Growth (ERG 1214)", "input": "Pierre Vinken, 61 years old, will join the board as a nonexecutive director Nov. 29.", "tops": [10], "nodes": [{"id": 0, "label": "proper_q", "anchors": [{"from": 0, "to": 28}]}, {"id": 1, "label": "compound", "anchors": [{"from": 0, "to": 14}]}, {"id": 2, "label": "proper_q", "anchors": [{"from": 0, "to": 6}]}, {"id": 3, "label": "named", "properties": ["CARG", "NUM"], "values": ["Pierre", "sg"], "anchors": [{"from": 0, "to": 6}]}, {"id": 4, "label": "named", "properties": ["CARG", "NUM"], "values": ["Vinken", "sg"], "anchors": [{"from": 7, "to": 14}]}, {"id": 5, "label": "measure", "anchors": [{"from": 15, "to": 23}]}, {"id": 6, "label": "udef_q", "anchors": [{"from": 15, "to": 23}]}, {"id": 7, "label": "card", "properties": ["CARG"], "values": ["61"], "anchors": [{"from": 15, "to": 17}]}, {"id": 8, "label": "_year_n_1", "properties": ["NUM"], "values": ["pl"], "anchors": [{"from": 18, "to": 23}]}, {"id": 9, "label": "_old_a_1", "anchors": [{"from": 24, "to": 28}]}, {"id": 10, "label": "_join_v_1", "properties": ["TENSE"], "values": ["fut"], "anchors": [{"from": 34, "to": 38}]}, {"id": 11, "label": "_the_q", "anchors": [{"from": 39, "to": 42}]}, {"id": 12, "label": "_board_n_of", "properties": ["NUM"], "values": ["sg"], "anchors": [{"from": 43, "to": 48}]}, {"id": 13, "label": "_as_p", "anchors": [{"from": 49, "to": 51}]}, {"id": 14, "label": "_a_q", "anchors": [{"from": 52, "to": 53}]}, {"id": 15, "label": "_nonexecutive_a_unknown", "anchors": [{"from": 54, "to": 66}]}, {"id": 16, "label": "_director_n_of", "properties": ["NUM"], "values": ["sg"], "anchors": [{"from": 67, "to": 75}]}, {"id": 17, "label": "loc_nonsp", "anchors": [{"from": 76, "to": 84}]}, {"id": 18, "label": "mofy", "properties": ["CARG", "NUM"], "values": ["Nov", "sg"], "anchors": [{"from": 76, "to": 80}]}, {"id": 19, "label": "def_explicit_q", "anchors": [{"from": 76, "to": 80}]}, {"id": 20, "label": "of_p", "anchors": [{"from": 76, "to": 80}]}, {"id": 21, "label": "def_implicit_q", "anchors": [{"from": 76, "to": 80}]}, {"id": 22, "label": "dofm", "properties": ["CARG", "NUM"], "values": ["29", "sg"], "anchors": [{"from": 81, "to": 84}]}], "edges": [{"source": 9, "target": 4, "label": "ARG1"}, {"source": 5, "target": 8, "label": "ARG2"}, {"source": 17, "target": 22, "label": "ARG2"}, {"source": 11, "target": 12, "label": "BV"}, {"source": 2, "target": 3, "label": "BV"}, {"source": 20, "target": 18, "label": "ARG2"}, {"source": 10, "target": 4, "label": "ARG1"}, {"source": 1, "target": 3, "label": "ARG2"}, {"source": 7, "target": 8, "label": "ARG1"}, {"source": 21, "target": 18, "label": "BV"}, {"source": 17, "target": 10, "label": "ARG1"}, {"source": 20, "target": 22, "label": "ARG1"}, {"source": 1, "target": 4, "label": "ARG1"}, {"source": 5, "target": 9, "label": "ARG1"}, {"source": 13, "target": 10, "label": "ARG1"}, {"source": 15, "target": 16, "label": "ARG1"}, {"source": 6, "target": 8, "label": "BV"}, {"source": 10, "target": 12, "label": "ARG2"}, {"source": 19, "target": 22, "label": "BV"}, {"source": 13, "target": 16, "label": "ARG2"}, {"source": 0, "target": 4, "label": "BV"}, {"source": 14, "target": 16, "label": "BV"}]}

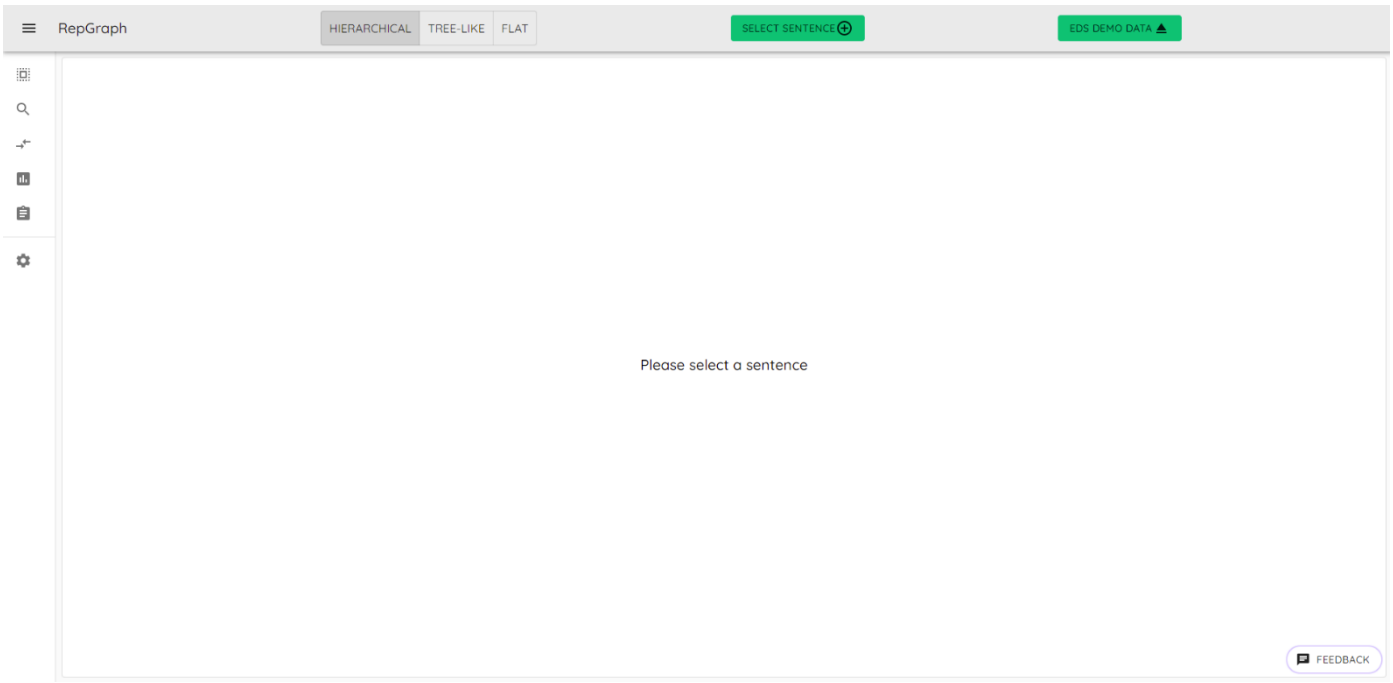
The user will be shown an upload page as follows:



Once on this page, the user can simply drag and drop their data file into the box or make use of the upload file dialog to select their file from their local directory. The file dialog only looks for the following extensions: “**.mrp” , “.dmrs” , “.eds” , “.ptg” , “.ucca” , “.amr”.** The user can also simply click the skip and use demo data button to view sample data of each framework.

When the file is present in the upload box, the user must click the upload button which will then send the file to the server to be processed and stored.

Once a file is uploaded, the user will be taken to the main page.

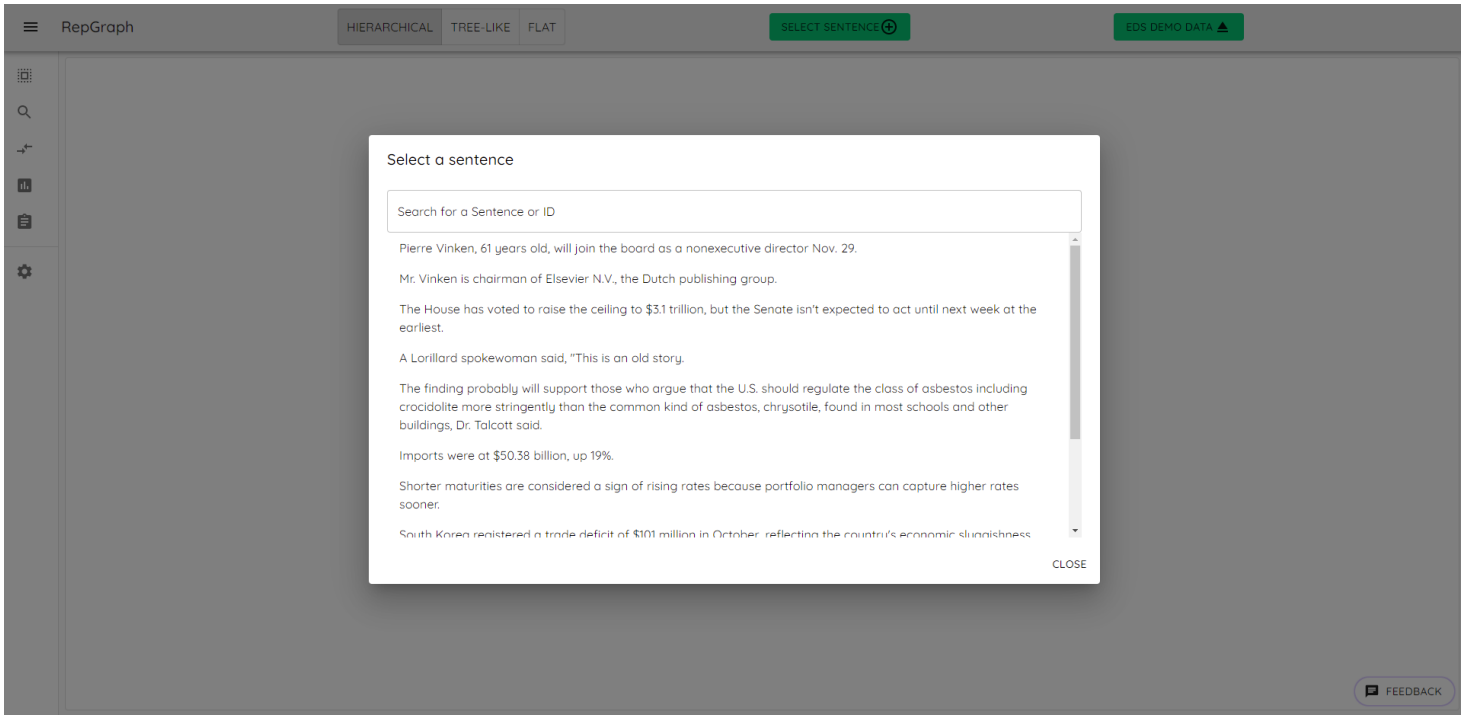


To reupload a new file, the user must click on the data-set button in the top right corner.

2.2. Select a Sentence

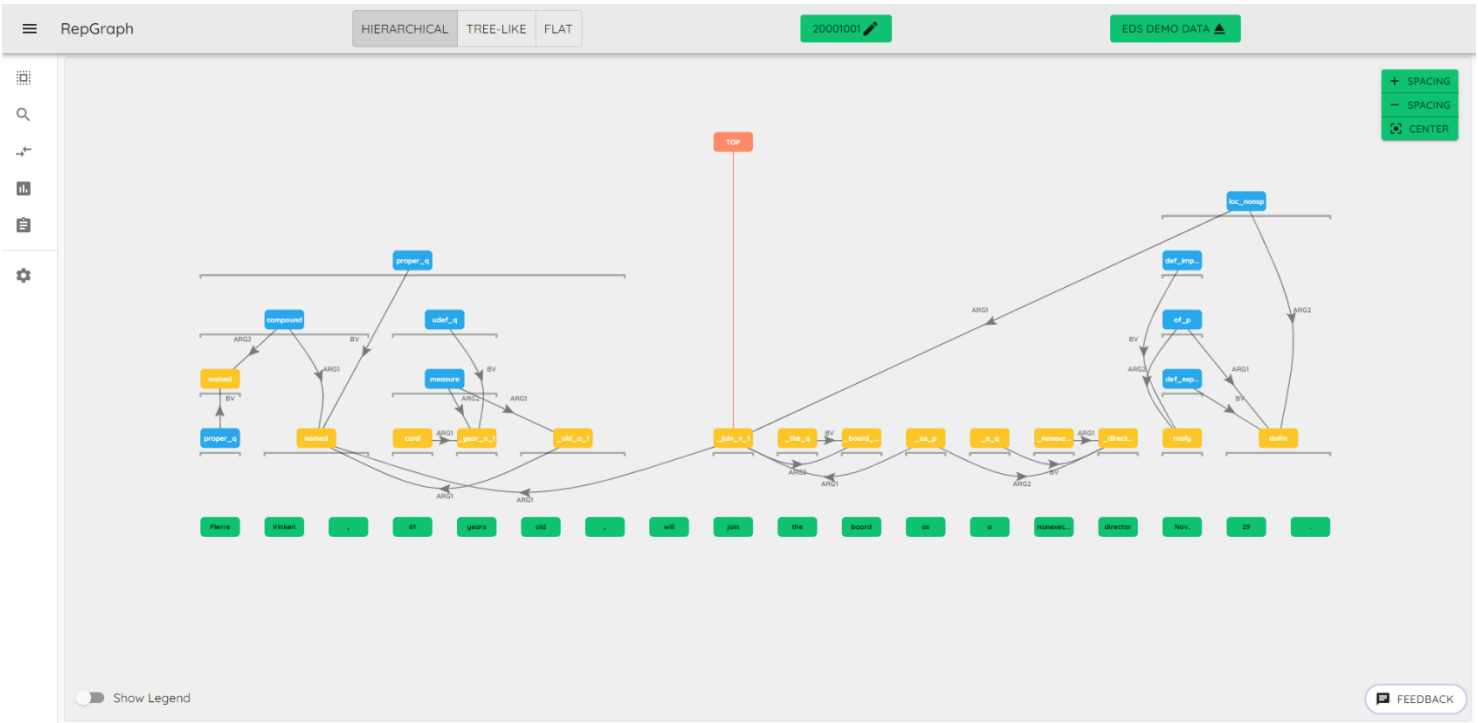
To select a sentence for visualisation, the user must click the sentence selection button.

When the sentence selection button is clicked, the user will be shown a list of sentences from the dataset.



From here, the user can scroll to find or search for (either by sentence or graph ID) the sentence they want visualised.

To visualise a sentence, the user must simply click on the sentence in the list. Once a sentence has been clicked on, the list dialog will disappear, and the visualisation will appear.



When a sentence has been chosen, all visualisations and analysis tools will act upon the chosen sentence.

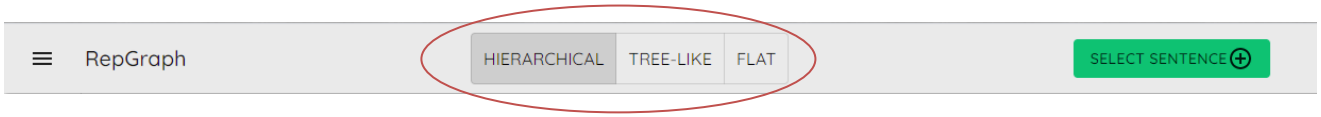
2.3. Visualisation Formats

The tool offers 4 different visualisation formats to view meaning representation graphs.

- Hierarchical
- Tree
- Flat
- Planar - *Special Case*

The last visualisation “Planar” is a special case and is not viewed in the same way as the others but that will be explained in the ***Formal Tests (4.5.2)*** section.

The three main visualisations are selected on the Top bar of the tool.

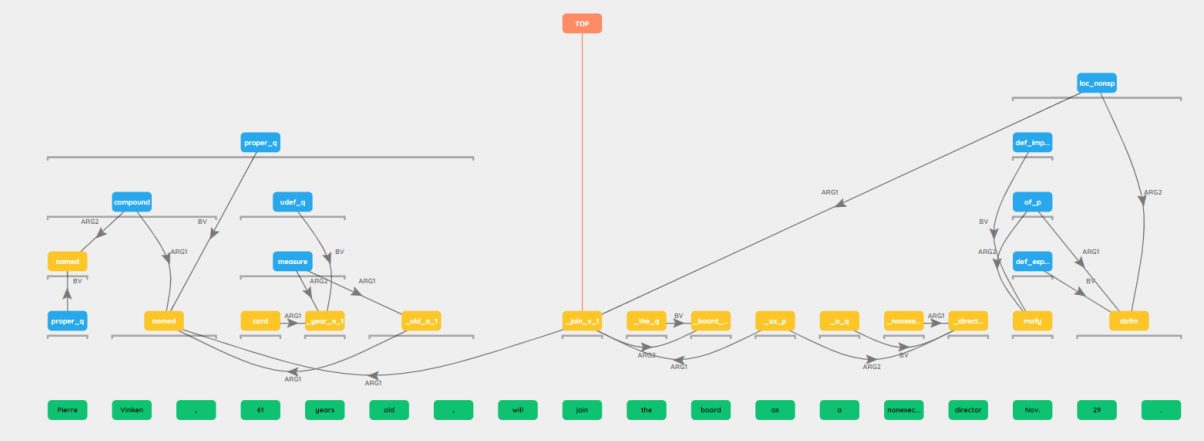


Whichever visualisation is clicked on, decides the visualisation of all graphs throughout the application, including within the analysis tools. This means users can choose whichever format they wish to view all graphs in consistently.

3. Visualisations

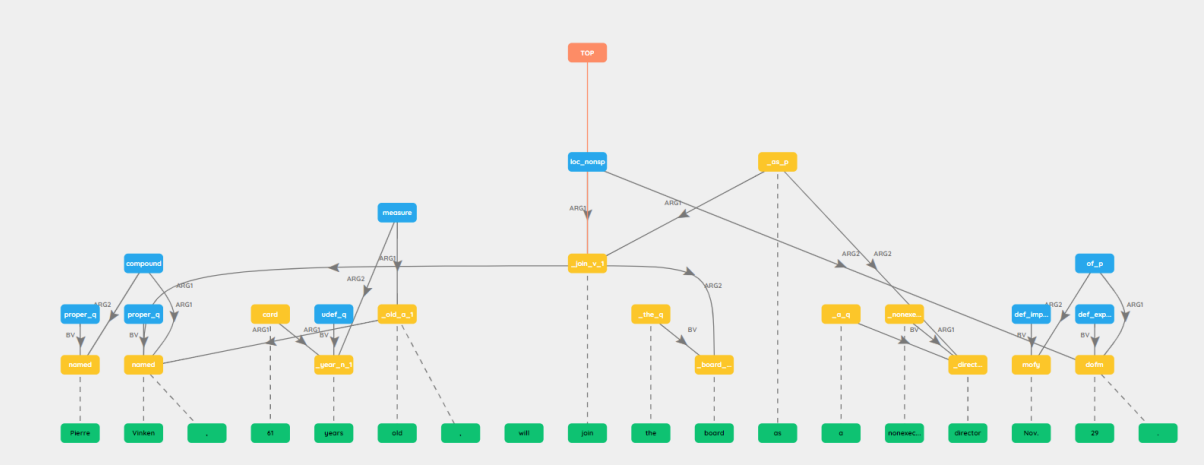
3.1. Hierarchical

The hierarchical visualisation format focuses on the spans of nodes. The width of nodes corresponds with the number of tokens they span.



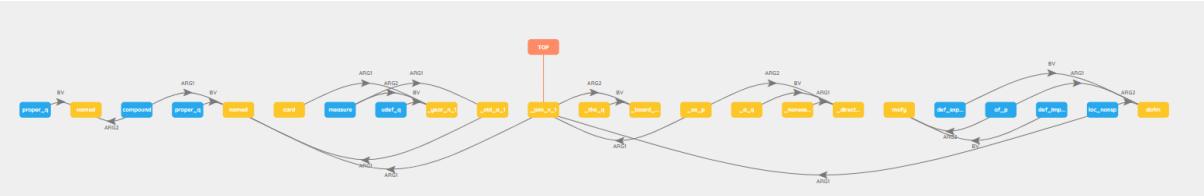
3.2. Tree

The Tree visualisation focuses on nodes with the most descendents i.e it visualises the graphs in a tree like manner with all nodes that are below other nodes have less descendants.



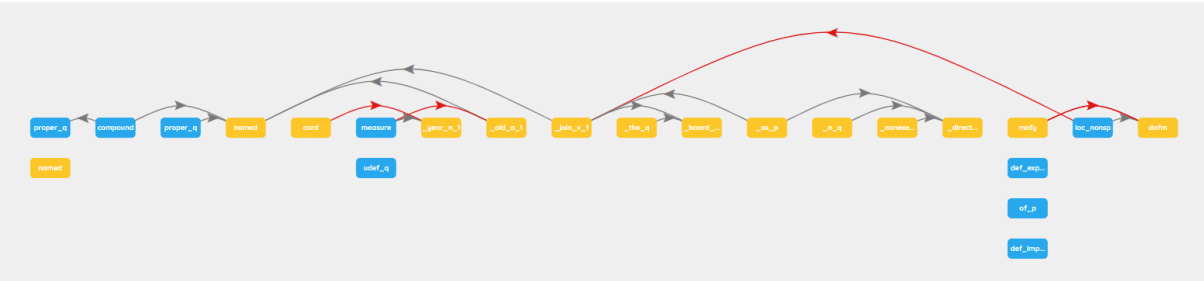
3.3. Flat

The Flat visualisation is the simplest visualisation which just shows the nodes on a flat plane with their edges.



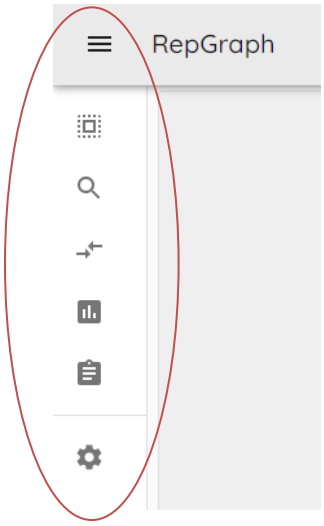
3.4. Planar

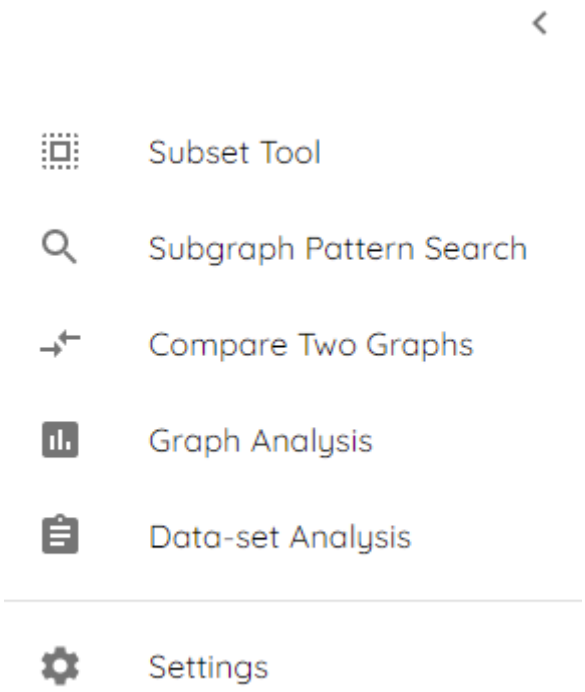
The planar visualisation is a special visualisation showing all nodes in their token reference position and showing all edges in the top plane whilst highlighting crossing edges.



4. Analysis Tools

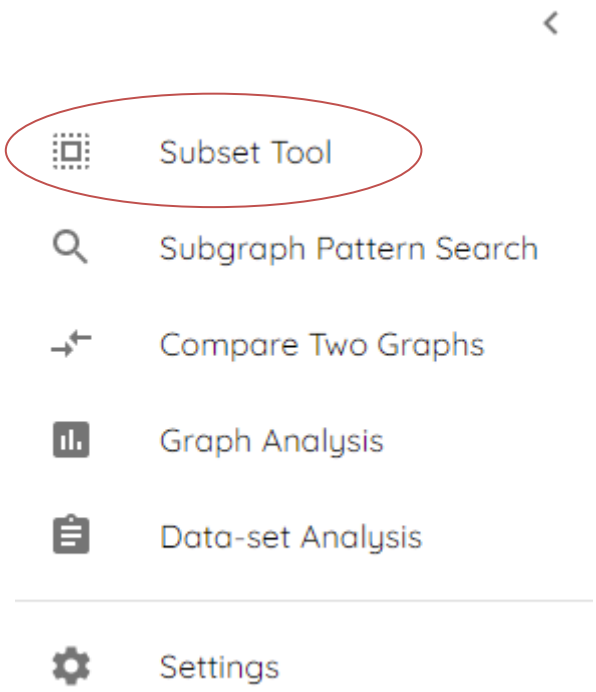
The analysis tools are used to analyse the selected graph or other graphs in the model. To access the analysis tools, the user must either click on the symbol referring to it's respective tool or expand the menu and click on the item in the list.



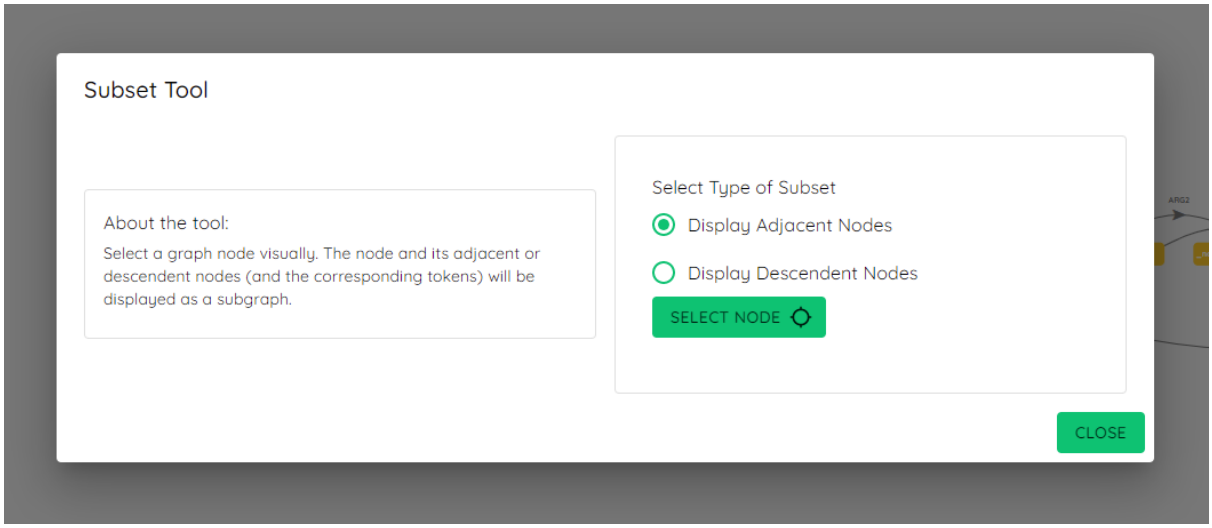


4.1. Display Subset

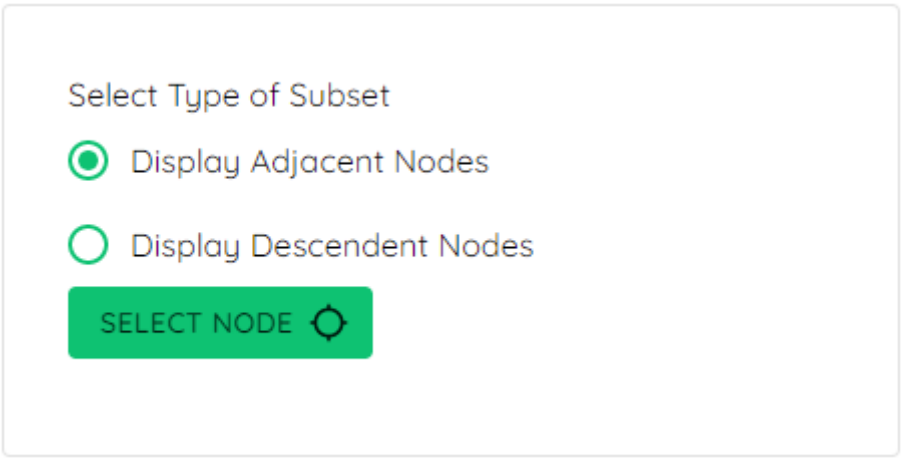
The subset tool is used whenever a user wants to only see a specific portion of a graph or if they want to see how a specific node is connected and/or relates to other nodes.



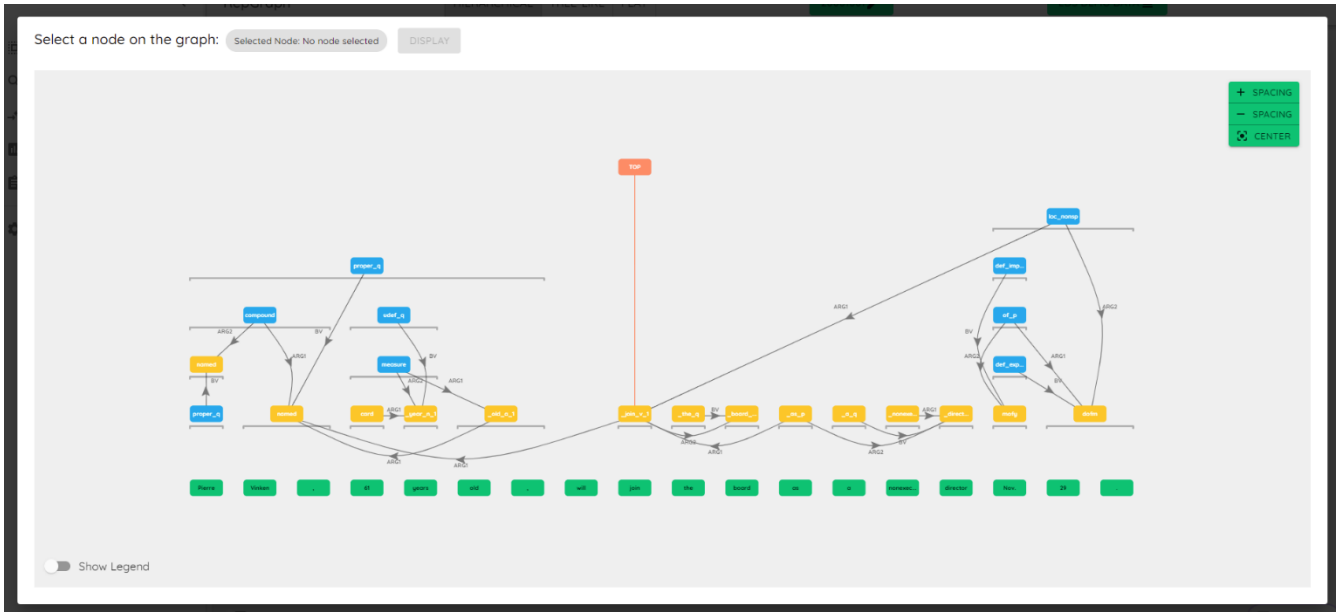
Once the tool is clicked, the list will show the user the tool's options.



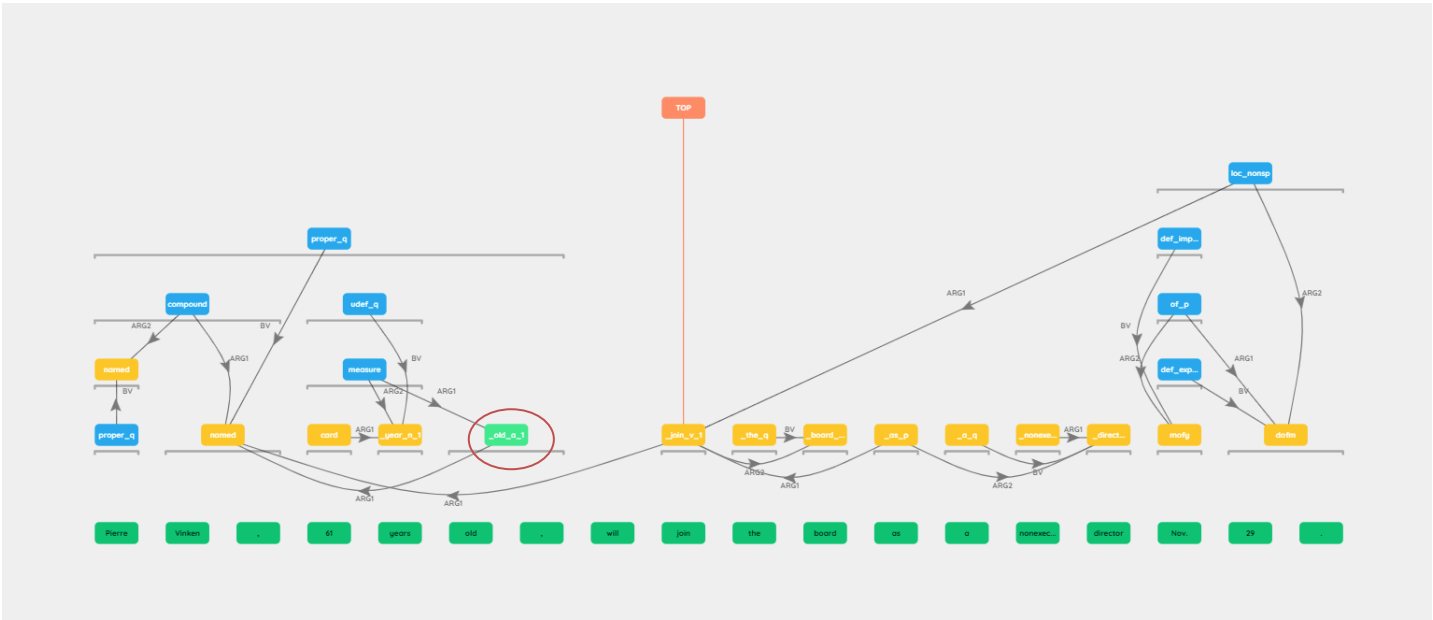
The user can then decide whether they want an adjacent subset (meaning the subset is created based on the selected node's adjacent nodes i.e all nodes that are connected to that node regardless of edge direction) or a descendent subset (meaning the subset is created based on the selected nodes descendants).



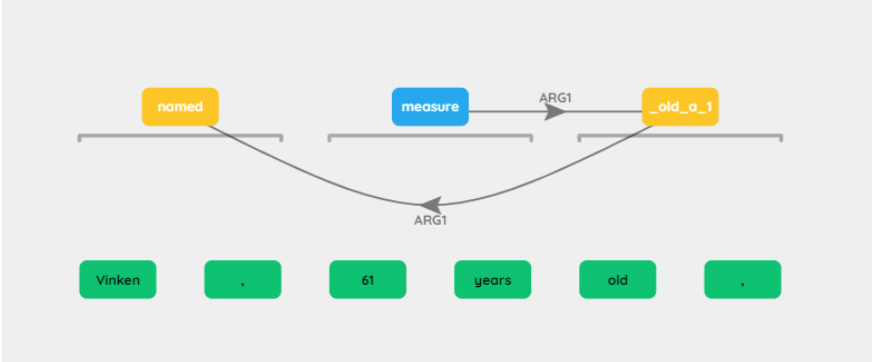
Once the user has selected which type of subset they want displayed, the user must click the "Select Node" button. This will open a dialog of the current graph selected.



The user must then select the node they want the subset to be created from.

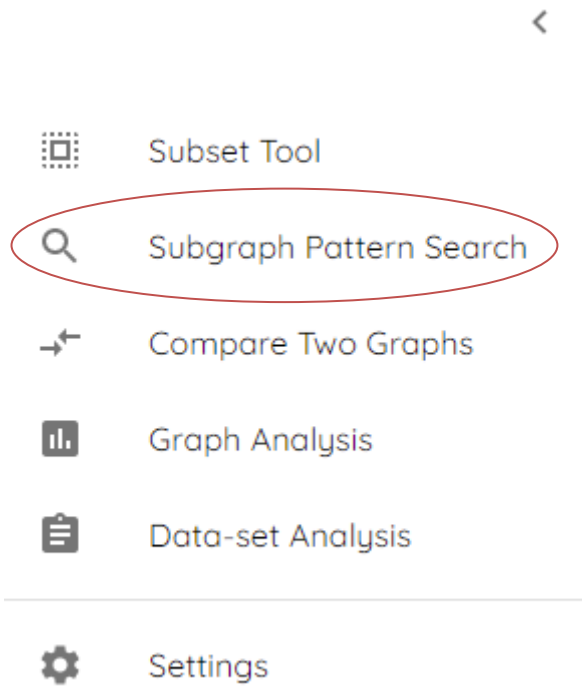


After the user has selected the node they want, they must press the “Display” button to display the subset. This will result in a dialog with the subset visualisation appearing.

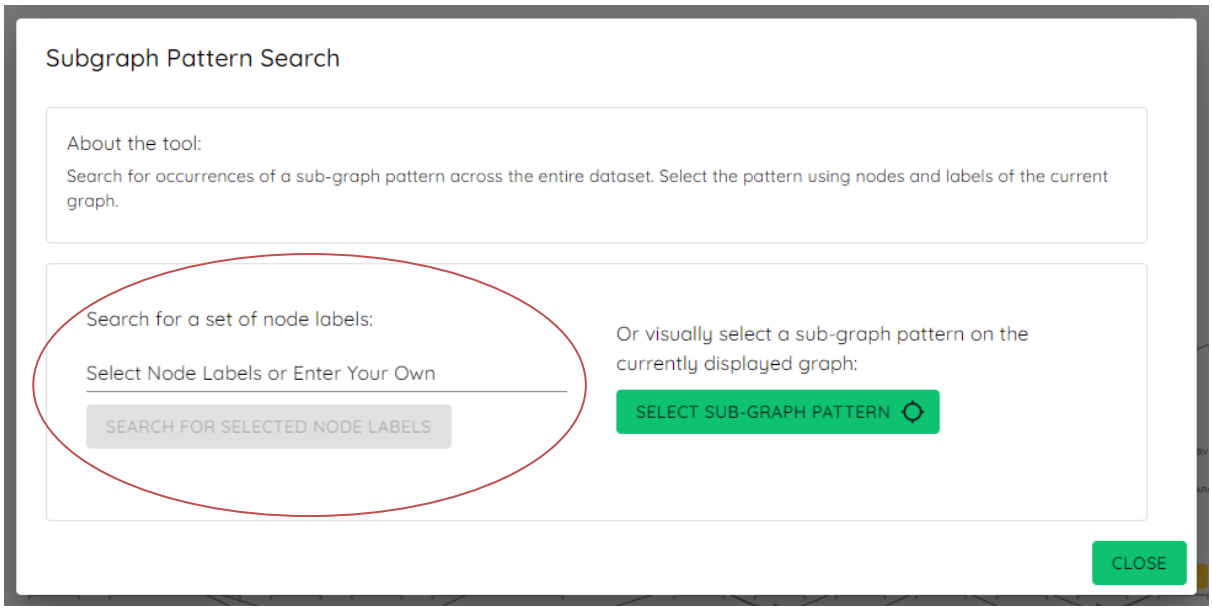


4.2. Search for Set of Nodes

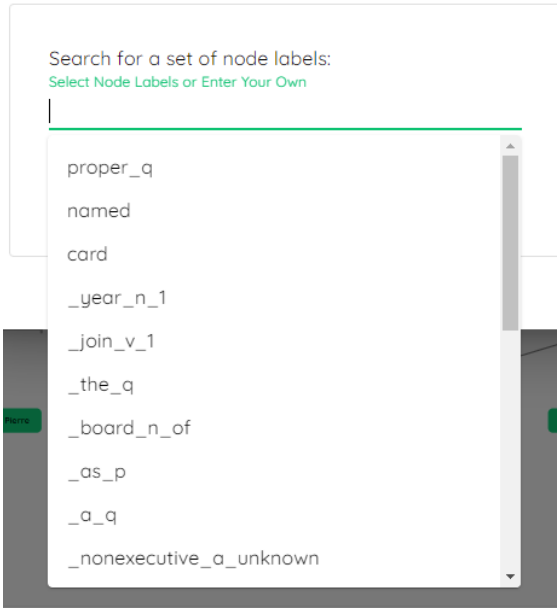
The search for a set of nodes tool is used when the user wants to know which graphs in the dataset have a certain set of nodes.



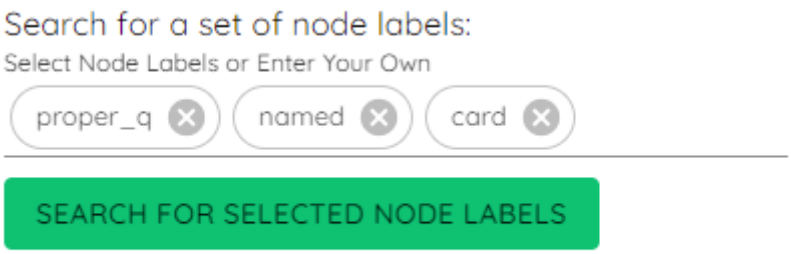
To search for which graphs have a set of nodes, the user must enter which nodes they want searched for.



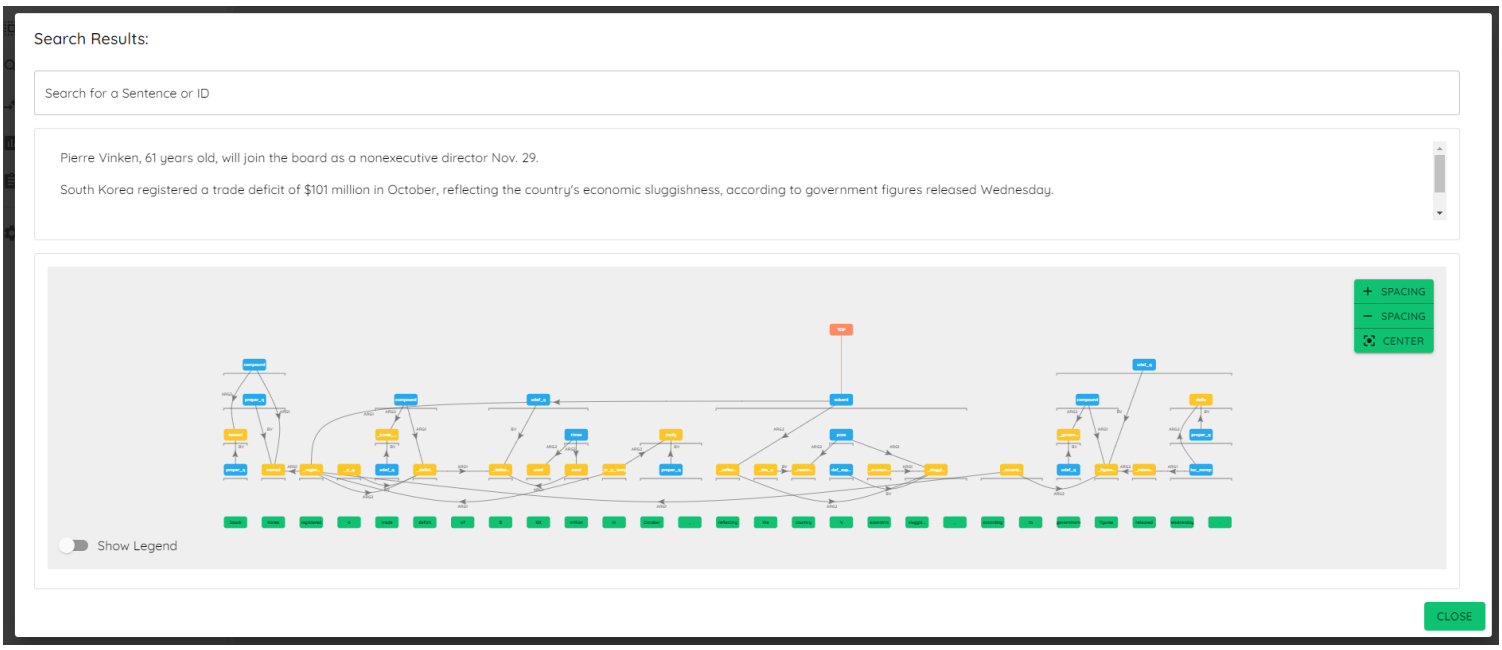
They can do this either by clicking the dropdown menu and using node labels that are present on the current selected graph or they can manually enter which nodes they want by typing into the box.



Once the node labels the user is searching for are loaded in, the search button must be clicked.

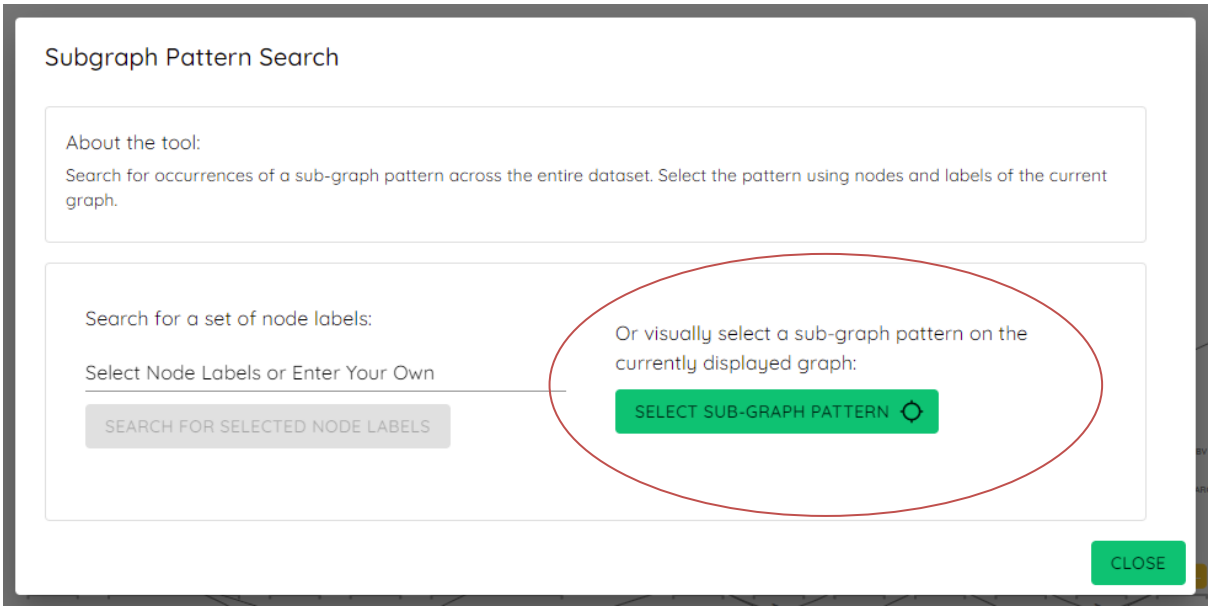


After the button has been clicked, a dialog will appear showing the user the search results along with a space to visualise the results quickly.

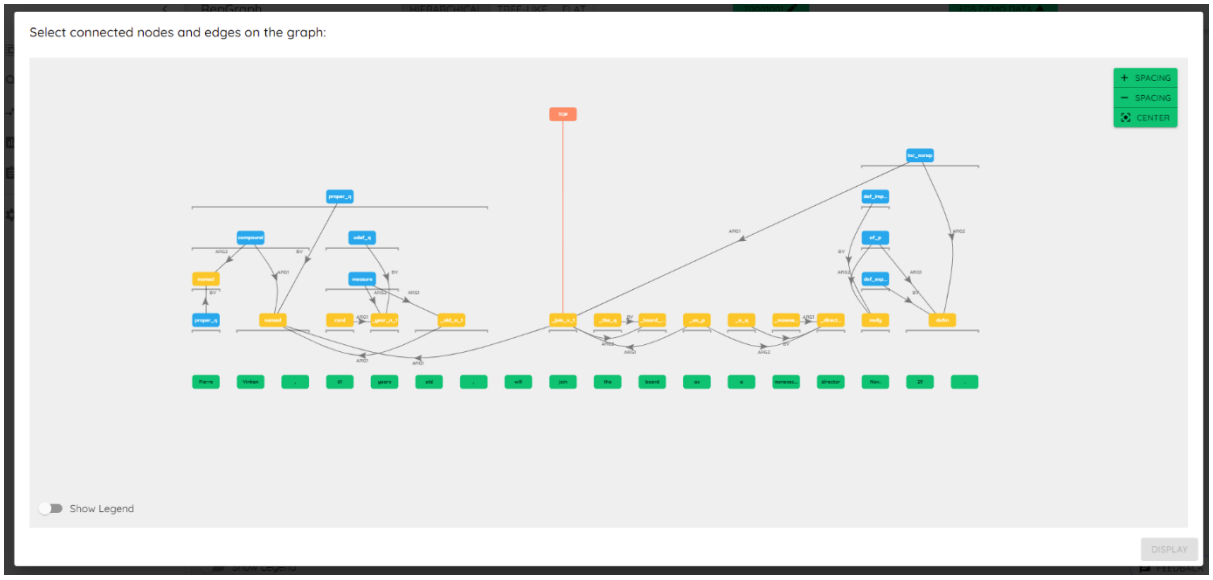


4.3. Search for Subgraph Pattern

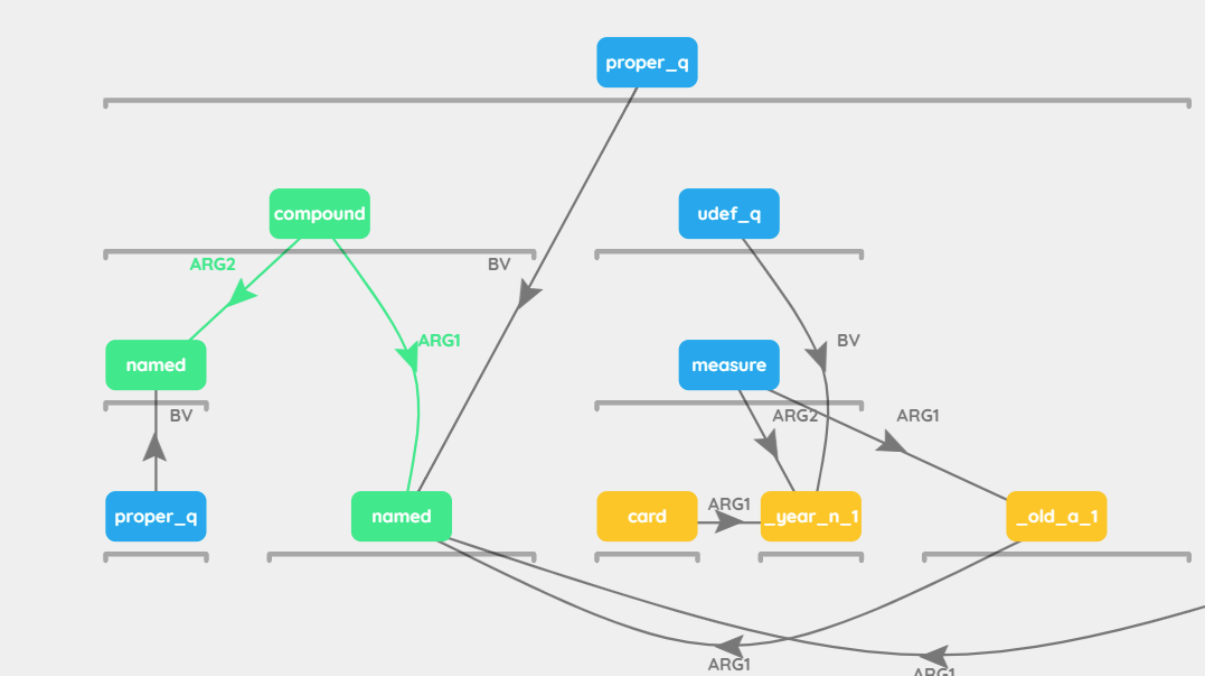
The search for subgraph pattern tool is used when the user wants a specific pattern in a graph searched for in other graphs.



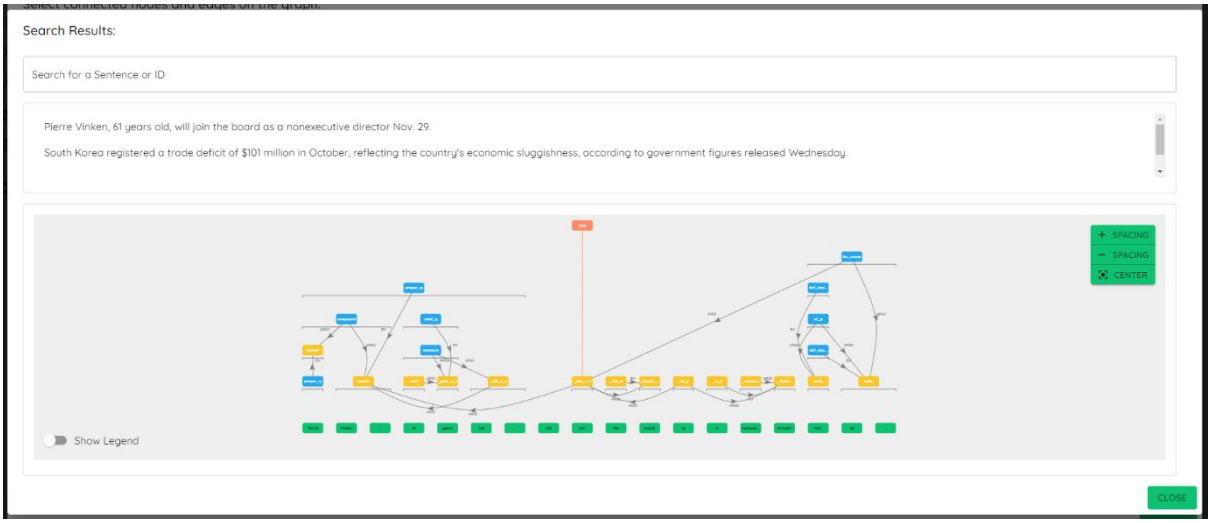
To use this tool, the user must click on the “Select Sub-Graph Pattern” button. This will open up the current selected graph in a dialog to allow the user to select the pattern visually.



Once the dialog is shown, the user must click on the nodes and edges of the pattern they want to be searched.

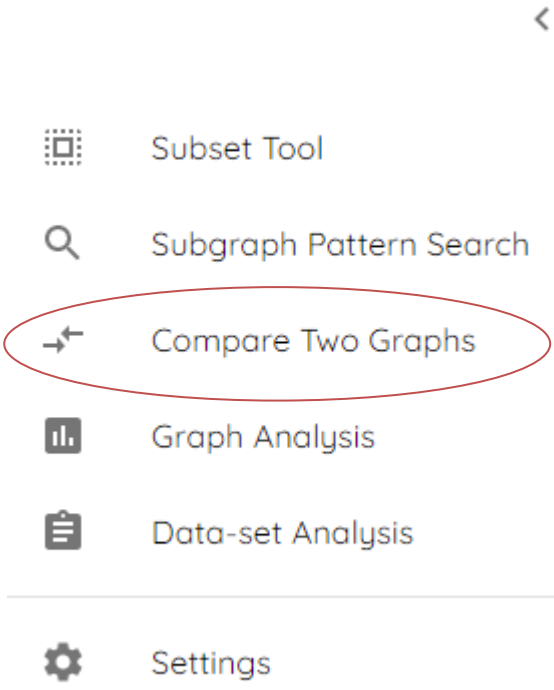


After the user selects the pattern they want, they must click the display button which will then display a dialog with the search results along with a space to visualise the results quickly.

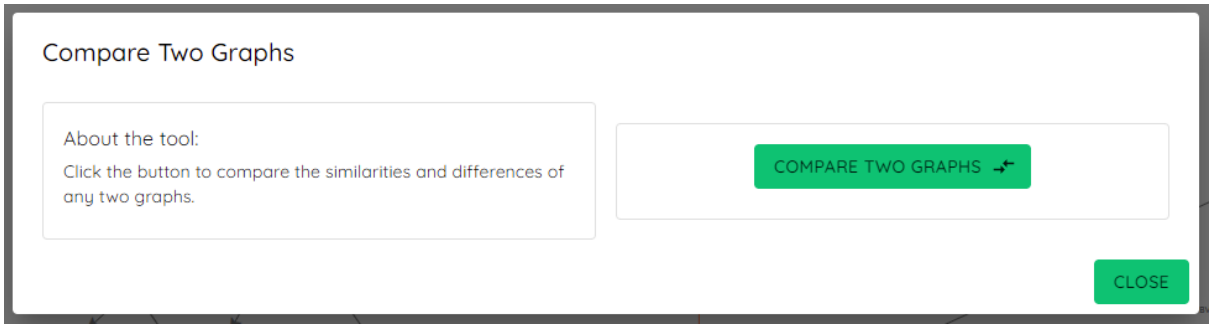


4.4. Compare Two Graphs

The compare two graphs tool is used when a user wants to compare two different graphs based on their nodes and edges.

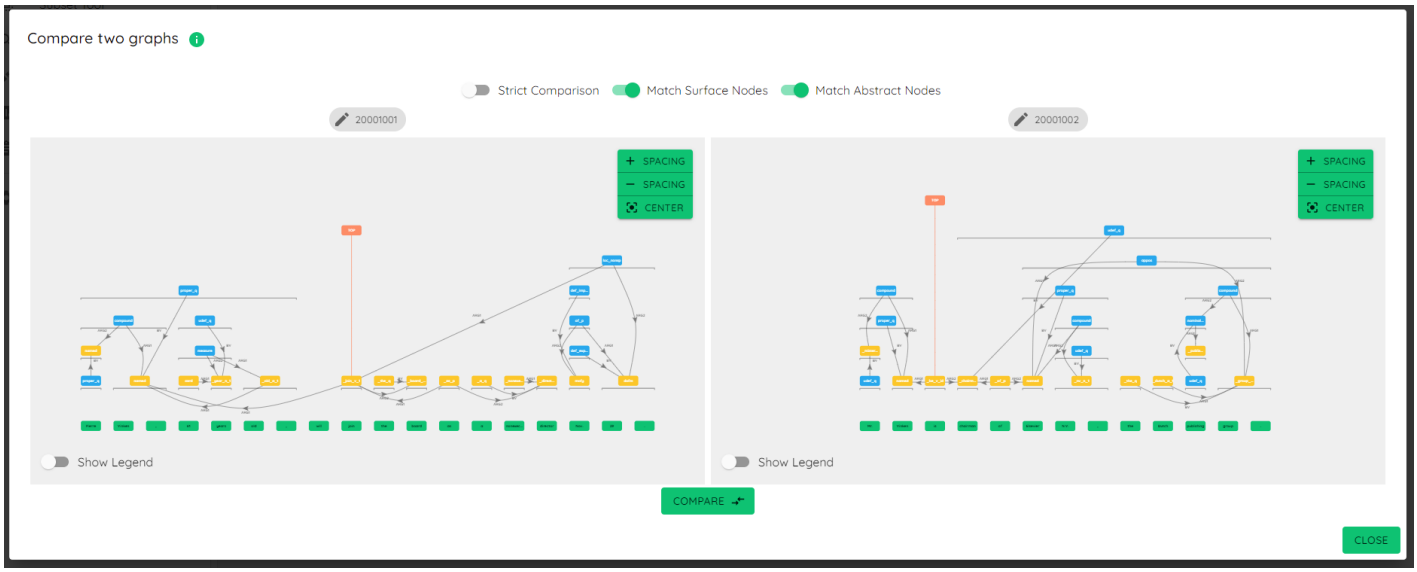


To use this tool the user must click on the “Compare Graphs” button which will open a dialog showing two spaces to select a graph.

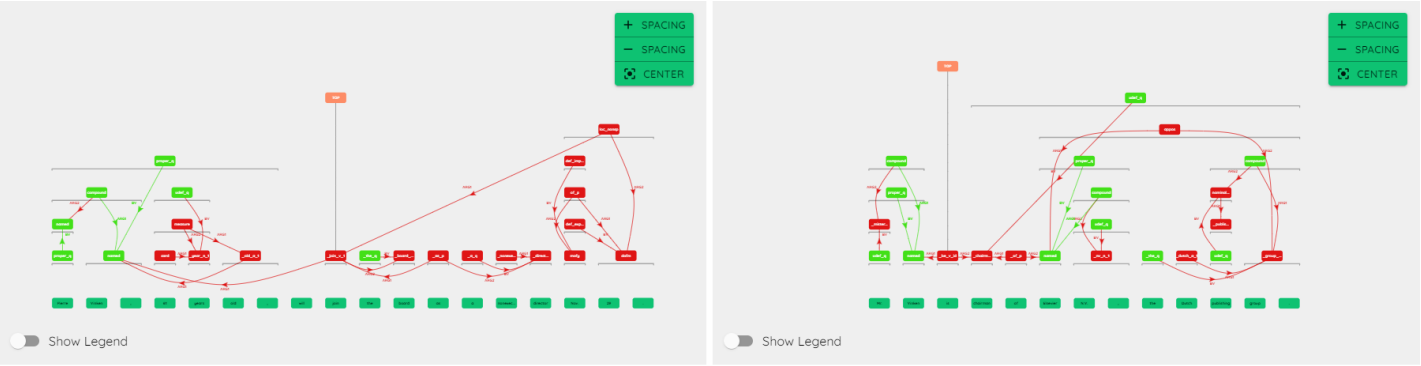


The user must then select two graphs by independently choosing them from the list by clicking the two sentence selection buttons

Once two graphs have been visualised on the dialog, the user must click the “Compare” button to compare the graphs.



The graphs will then change colour indicating the similar nodes and edges on each with dark green colour and non-similar nodes and edges with a red colour.

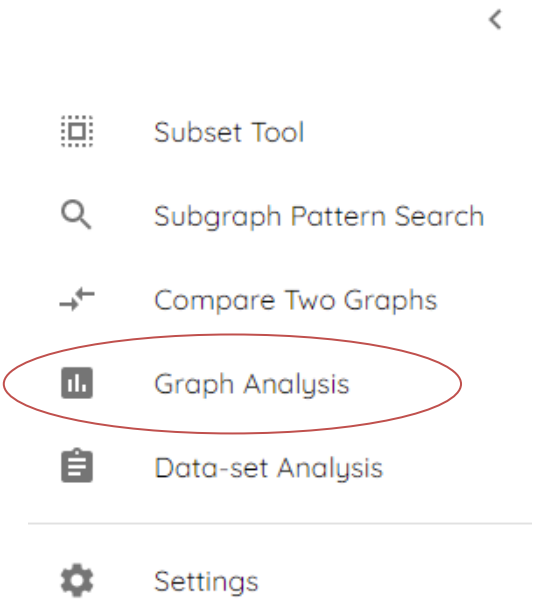


Strict mode compares the graphs in a stricter fashion, the details of which can be found in the software demonstration paper. The match surface and match abstract nodes switches are simply filters. If unselected, surface or abstract nodes will not be looked at when comparing.

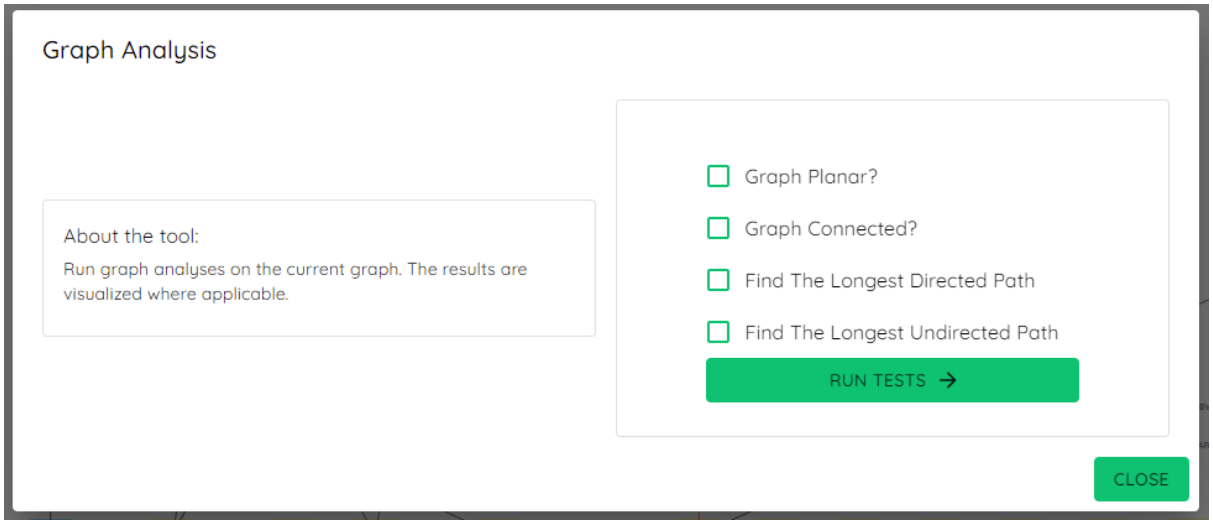
4.5. Formal Tests

The formal tests tool consists of a variety of tests the user can perform on the selected graph. The tests available are:

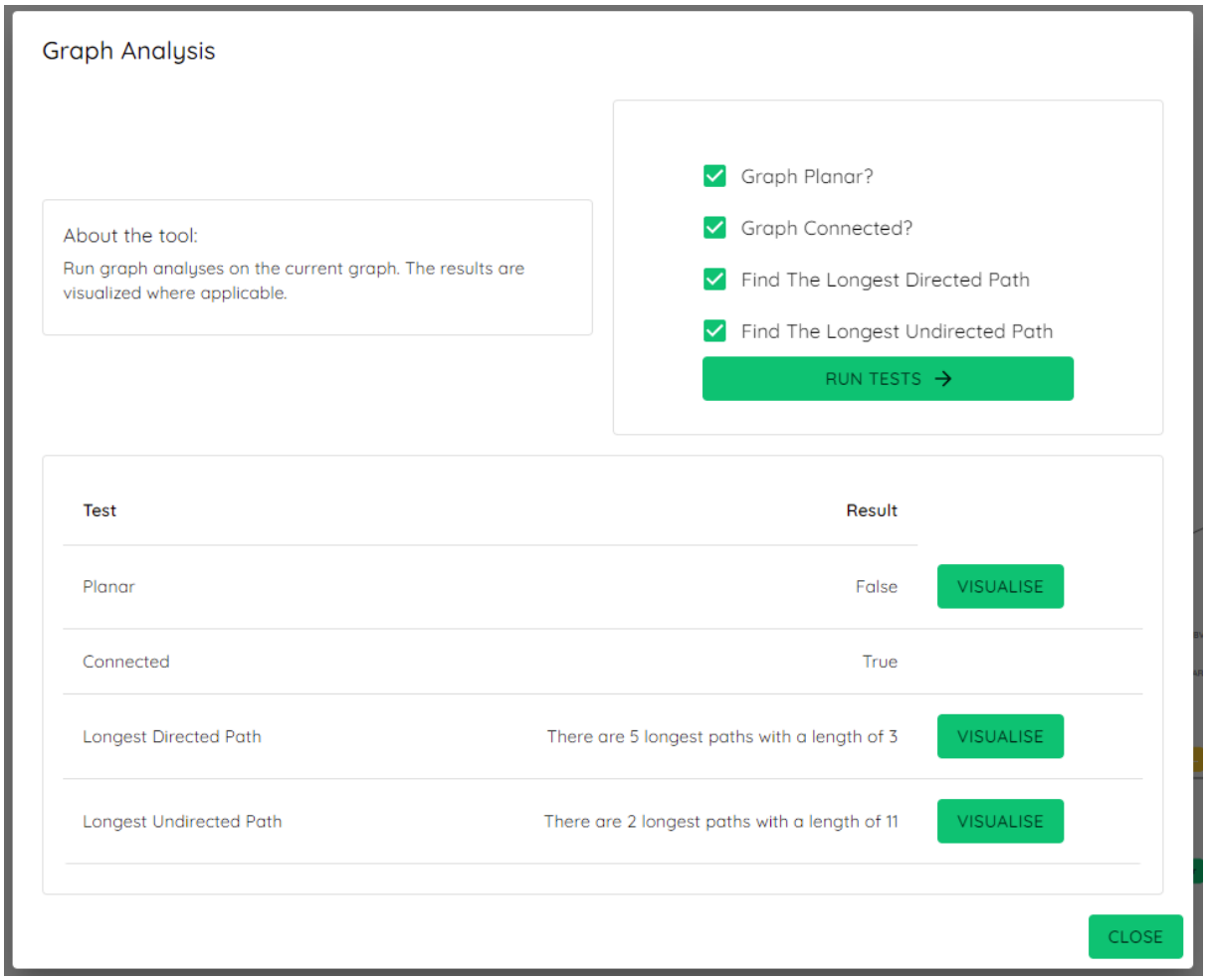
- Is the graph connected?
- Is the graph planar?
- What is the longest directed path?
- What is the longest undirected path?
- Directed vs Undirected Cycle Checker

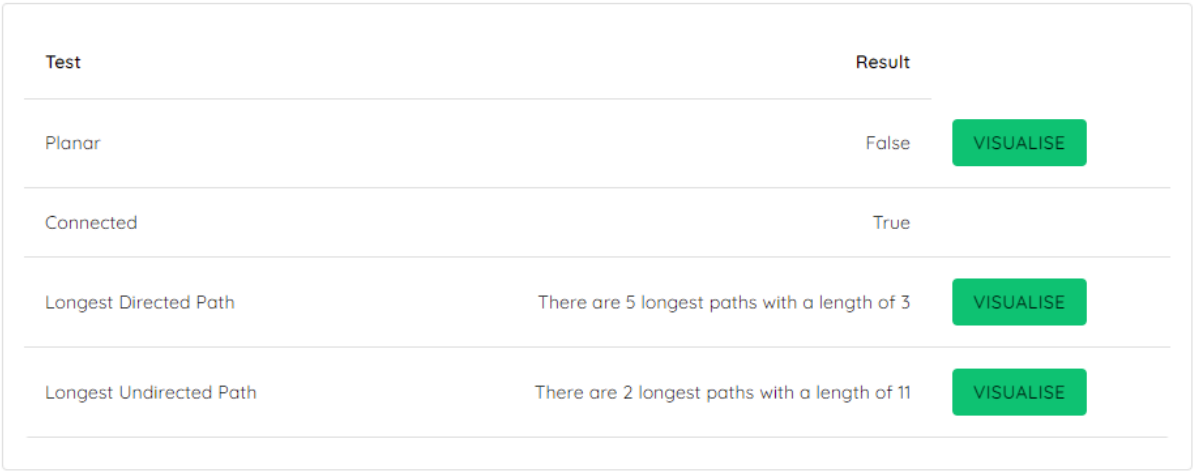


To run any of the tests, the user must select the test they want performed by ticking the checkbox and clicking run tests.



The results are displayed in a list below.





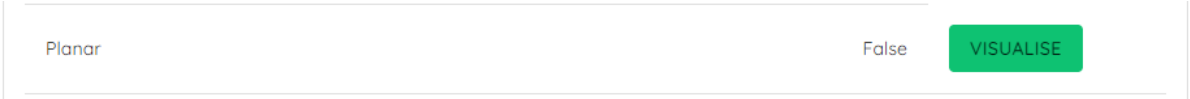
4.5.1. Connected

The connected test simply tells the user if the selected graph is connected or not. It will return true if it is connected and false if it is disconnected. The result will be displayed on the list.

4.5.2. Planar

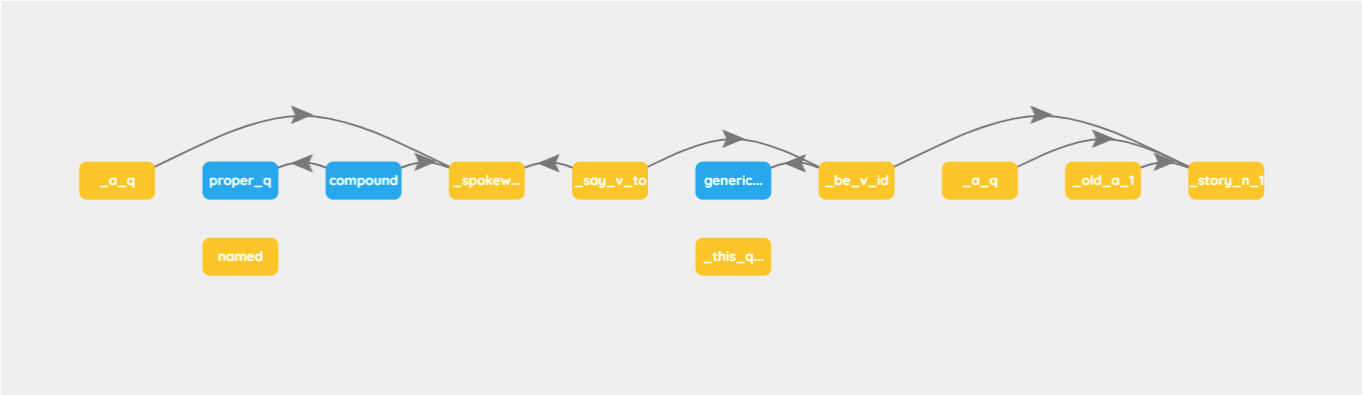
The planar test tells the user if the selected graph is planar or not. This test also provides the 4th visualisation format.

The test returns true if the selected graph is planar and false if it is not.

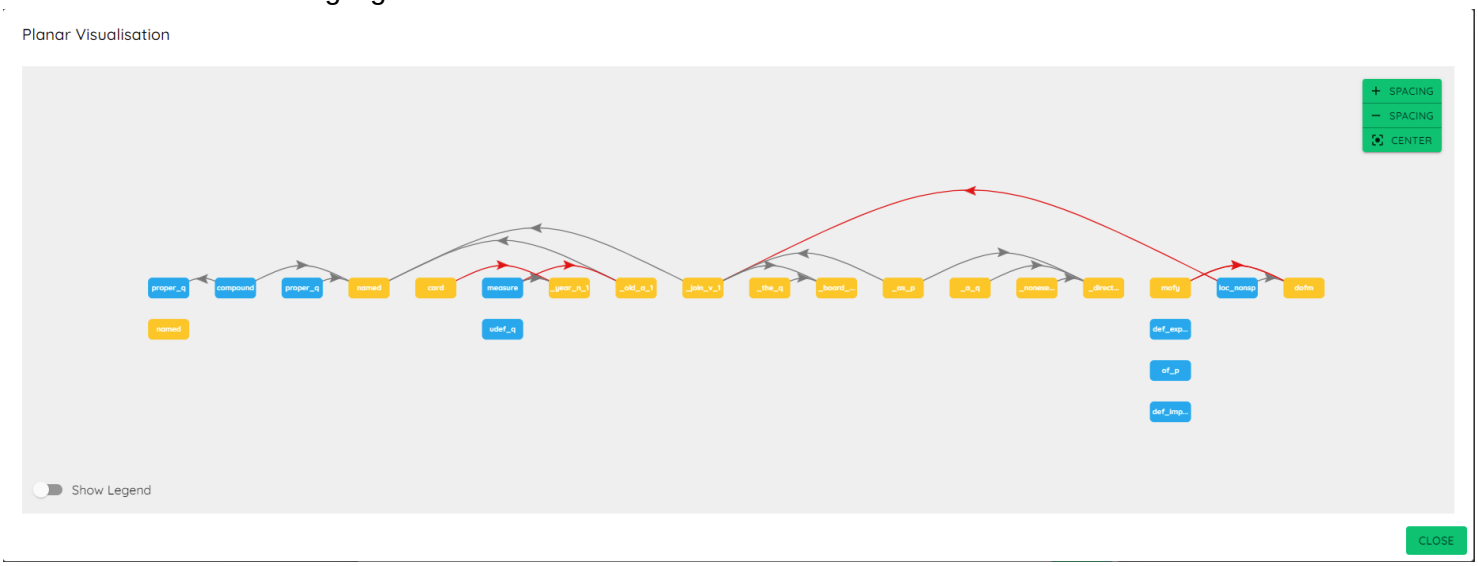


To see the 4th visualisation illustrating the planarity of the graph. The user must select the visualise button.

Below is an example of the planar visualisation for a graph that is planar. As can be seen below, all edges that obey the planarity requirements and are shown in black.



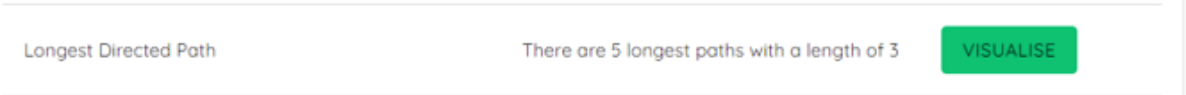
Below is an example of the planar visualisation for a graph that is not planar. As can be seen below, the edges that break the planarity property are highlighted in red.



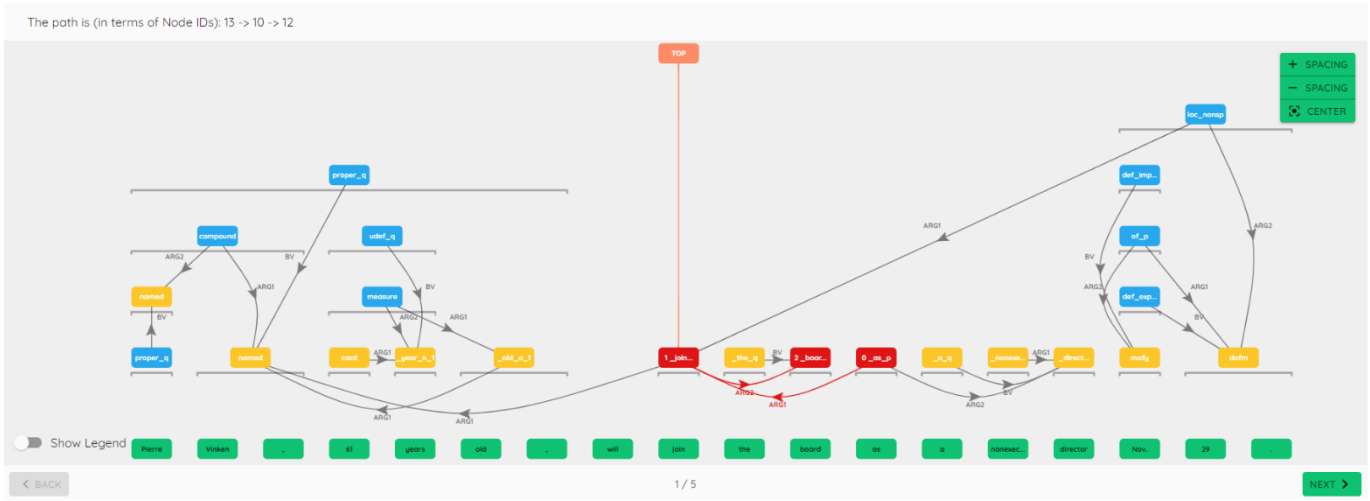
4.5.3. Longest Path

4.5.3.1. Directed

The directed longest path test shows the user the longest directed path either by text which has the node id's or by visually highlighting the nodes and edges of the path in a dialog when the user clicks the "visualise" button.

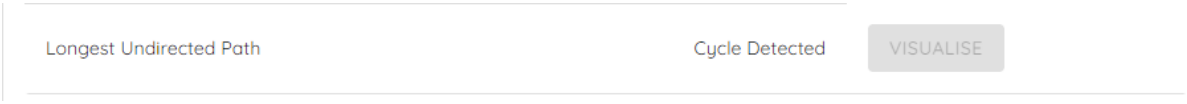


Once the button is clicked, a dialog will appear with the graph having it's nodes and edges highlighted to illustrate the longest directed path.



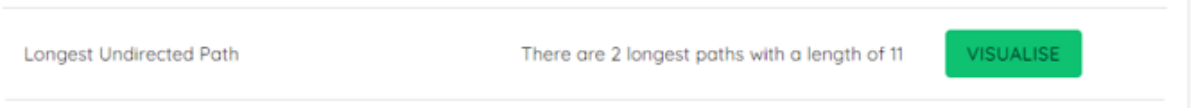
The user can also go through the multiple longest paths by making use of the "Next" and "Back" buttons

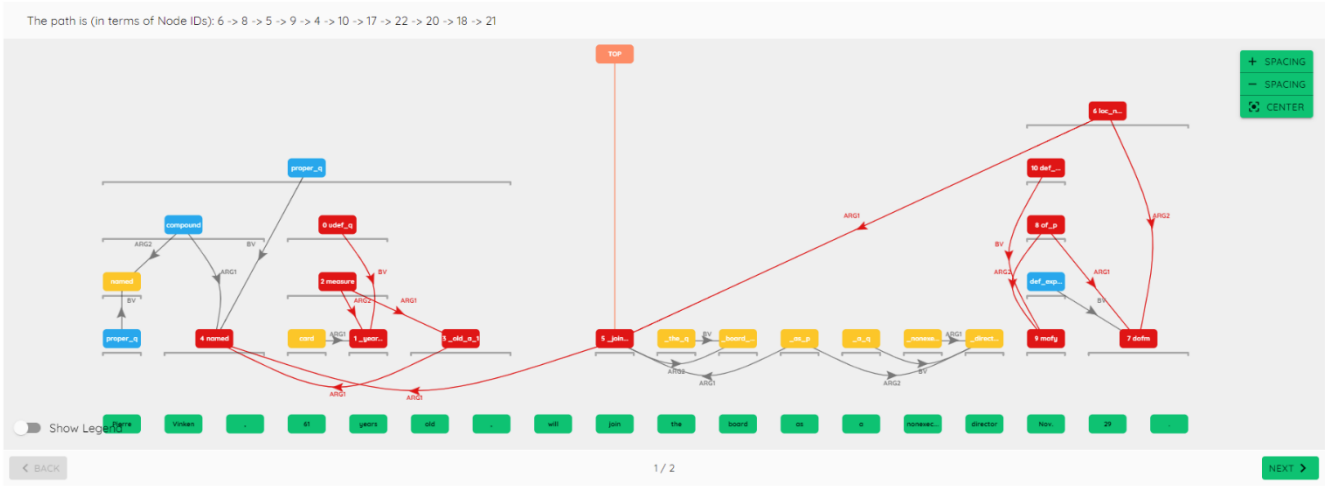
This test also acts as a directed cycle checker as if there is a cycle the longest path cannot be viewed and instead of a result the longest path value will say "Cycle Detected".



4.5.3.2. Undirected

The undirected longest path test acts in the same way as the directed path test except the edges directions do not matter.



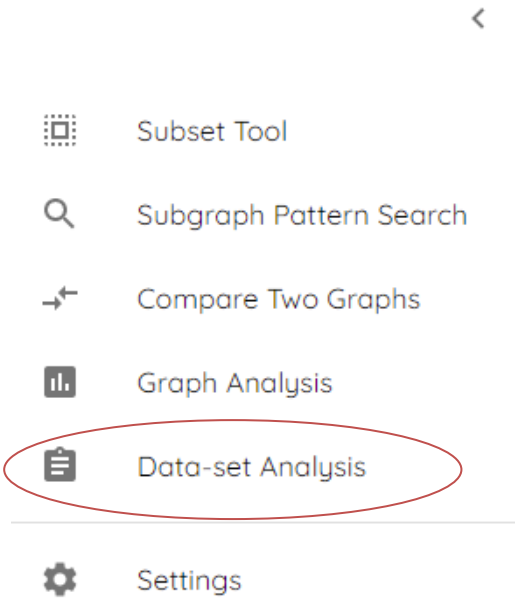


The user can also go through the multiple longest paths by making use of the “Next” and “Back” buttons

This test also acts as an undirected cycle checker and will inform the user in the same way as the directed test.

4.6. Data-set Analysis

The user can get a set of statistics about the uploaded dataset.



Once clicked, a dialog will appear to confirm the action.

When the user clicks the “run analysis” button, the following list will appear:

Test	Result
Average Number of Edges	23.9
Average Number of Nodes	24.8
Average Number of Tokens	20.8
Average Span of Node	1.7
Percentage of Directed Cyclic Graphs	0.0%
Percentage of Disconnected Graphs	0.0%
Percentage of Planar Graphs	30.0%
Percentage of Undirected Cyclic Graphs	10.0%
Total Number of Edges	239
Total Number of Graphs	10
Total Number of Nodes	248
Total Number of Tokens	208

Giving the user a general idea about the graphs present in the dataset.