

Manual

for “Combining magic lenses with large language models for text corpus exploration” by Phillip Wohlfart (VIS/VISUS)

Tool description

This tool aims to facilitate the exploration of the contents of various scientific articles related to the visualization field using a large language model.

Its main functionality is to summarize selected articles in concise form and return their most salient facts. Furthermore, the tool allows the quick comparison of a selected set of articles. For these language processing functions the latest version of GPT-4o-mini is integrated into the application.

The tool employs a “magic lens” which is used to investigate a set of data points in a view. Each data point represents an abstract of a visualization article published at *IEEE VIS* from 1990-2023. The proximity between points on the screen relates to a calculated similarity of the abstract texts. This results in a clusterization of the abstracts.

You can use your mouse to drag the magic lens over the points and get insights about the abstracts behind it. The lens size can be changed by a slider.

The different visualizations and “AI functionalities” can be changed in the options.

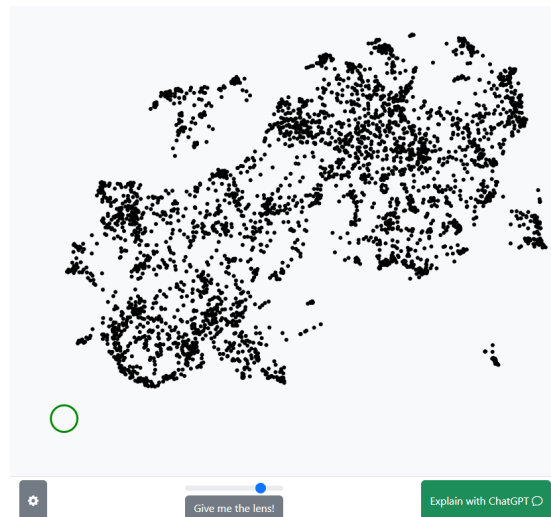


Navigating in the Graph:

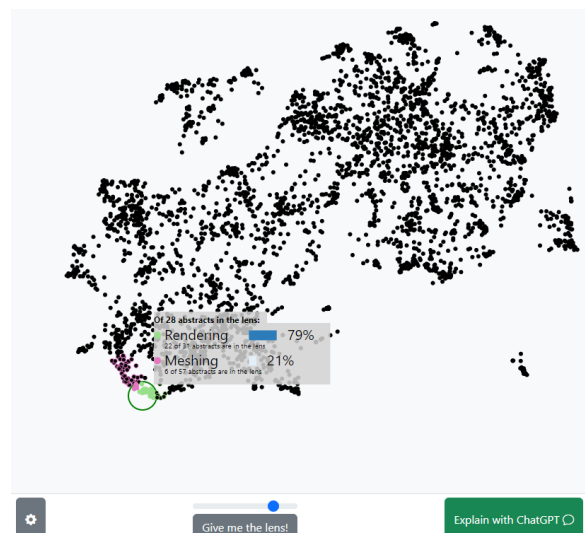
The left side of the application consists of the graph where you can explore articles and examine them.

- **Move the lens by clicking inside it and dragging it** to the desired position.
- **Double clicking on a data point** gets you more information about the related abstract
- Initially all the data points are black. As you explore and identify clusters in the graph you will be able to see their related color in the graph. When you are in the process of exploring and some data points of a cluster are inside your lens, the other points of this cluster will receive a slight highlighting with same color around them. This lets you identify related points more easily.
- When your mouse is inside the graph window you can **zoom** with the mouse wheel
- You can **pan** the whole graph in a direction by clicking inside it and dragging it in the opposite direction
- Change the lens size by using the slider at the bottom (dragging the slider or pointing at it and scrolling with the mouse wheel)
- “Give me the lens” button puts the lens in an appropriate size at the center of your graph no matter how you zoomed or shifted the graph

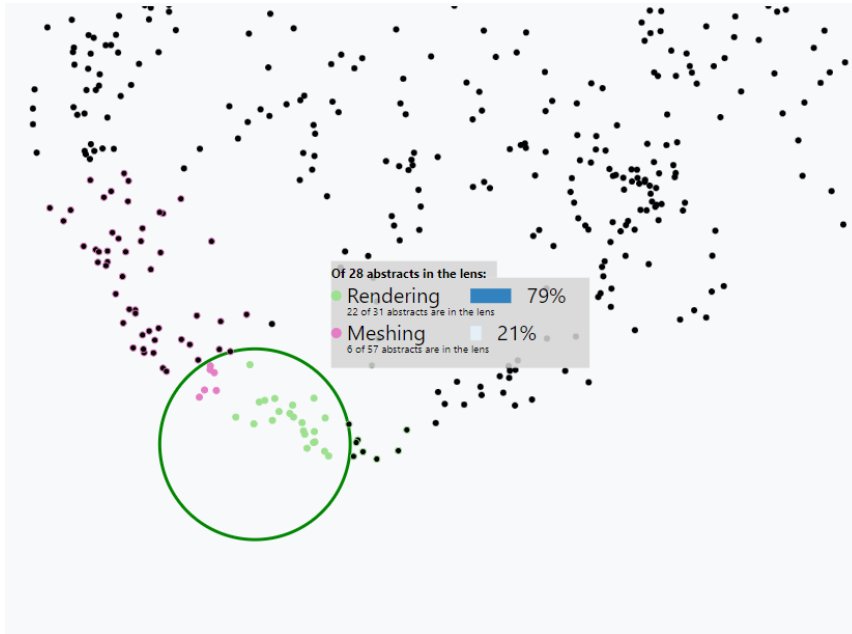
Initial graph view with lens at the bottom left:



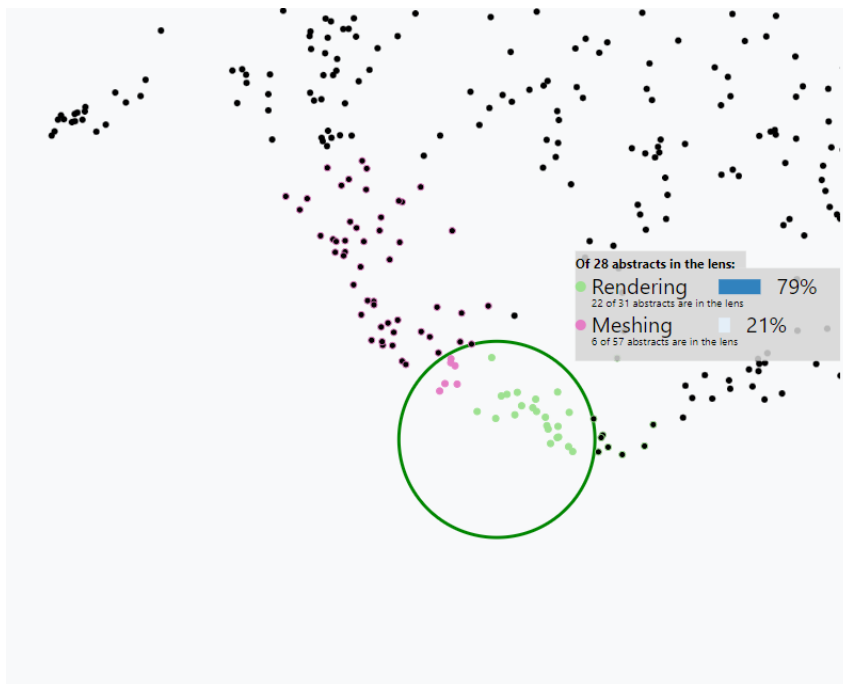
Move the lens to the right per drag to the right and drop:



Zoom in by hovering your mouse over the graph and scrolling the mouse wheel:

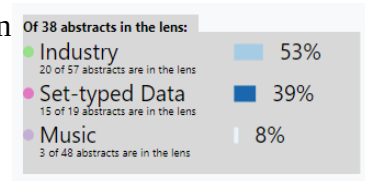


Pan the graph to the left by hovering your mouse over it and dragging it to the right:



Overview:

Info board (see p. 5): Displays how many abstracts of which topics are in your lens, in absolute numbers and relation. The saturation of the bar shows how well the current selected abstracts represent the whole topic. dark blue → selected articles represent whole topic *well*
white → selected articles represent whole topic *bad*



Quick task (see p. 7): Displays a short summary or the most important fact of the selected abstracts (change in settings). Is executed every time you drop the lens somewhere. You can get a more detailed response afterwards by clicking the corresponding button. The *lens ghost* (lens in gray) shows the old position and lets you compare the old response with the new one.

These abstracts primarily discuss **visual analytics** tools and techniques designed to enhance the interpretability, debugging, and optimization of deep learning models across various applications, addressing their black-box nature and promoting human-in-the-loop interaction.

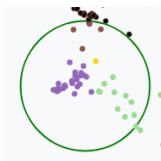
Show more details

These abstracts collectively highlight the development and application of advanced visualization and **visual analytics** tools for enhancing cybersecurity, **anomaly detection**, **network diagnostics**, and **situational awareness** in complex and large-scale systems.



Custom task (see p. 6): Choose between *summary*, *similarities*, *three most important facts* and *controversies* among the abstracts. This task can be instructed by clicking on the green “Explain with AI” button. The response will be shown in a big field that pops up on the right side.

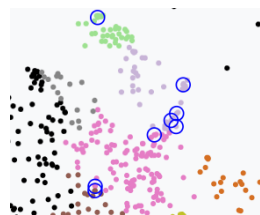
Abstract info (see p. 8): You can inspect any abstract by double clicking the data point on the graph. It will be highlighted yellow afterwards. In the yellow container on the right you get a short summary of the abstract and its corresponding topic. You can expand the text to get more information. The authors of the article are listed and the full article can be accessed.



Topological Analysis

The authors propose a novel, provably correct test for determining the parity of Parallel Vectors (PV) points on a cell face, which improves PV-extraction algorithms by ensuring closed curves in continuous fields.

Author info (see p. 8): Clicking on an author name will show information about the author. The articles of the author will be marked with a blue circle on the graph.



Keywords (see p. 10): You can click the links in the text anytime to open a new field on the side with an explanation for the keyword. There you can use this word, substrings of it and custom words to search through all the abstracts and mark them with red circles.

Photometric Calibration:

Photometric calibration refers to the process of adjusting and aligning the brightness and color output of multiple projectors in a display system to ensure uniformity and consistency across the projected images. This calibration addresses photometric non-uniformity, which can cause visual inconsistencies, by correcting disparities in brightness and color, thus enhancing the overall viewing experience.

Photometric Photometric Calibration Calibration

Remove circles Custom Search Exact Search



Article reference (see p. 9): In the texts there are links to specific articles given with reference to their authors e.g. “(Lin et al.)”. A click on this gets you to the yellow “abstract info” field and highlights the data point.

Another key finding is the focus on **user-centered design** for visualization tools that cater to specific audiences, including coaches, analysts, and fans. Systems like OBTracker (Wu et al.) and Omniculars (Lin et al.) illustrate **tailored visual analytics** that enhance the analysis of **off-ball movements** and **in-game behaviors**, respectively. This approach aims to demystify game dynamics and provide **actionable insights**, improving engagement and strategic decisions during live events.

Info Board

- Will be generated every time you move the lens and is based on the current abstracts in the lens
- Displays how many abstracts are in the lens
- Displays which themes are in the lens. The color is the same color the data points have.
- Displays how many abstracts of each theme are in the lens
- Displays the ratio of themes in the lens by a bar on the right and a percentage next to the bar.
- The significance the current articles in the lens of a theme have for the whole theme is shown by the saturation of the bar. The color reaches from white to a dark blue.

A high saturated bar means that there are many articles of the topic inside the lens compared to the maximum number of articles of the topic. This means that the topic should be well represented by the articles that are currently in your lens. This lets you better estimate how meaningful a response can be not only for your selected abstracts but for the whole topic.

- Example:

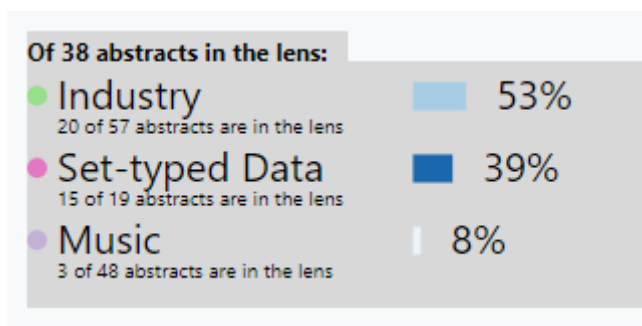
In the lens are currently 38 abstracts of three different topics.

The topic “Industry” comprises 57 abstracts over the whole dataset in total.

Currently in the lens are only 20 of those 57 abstracts of the topic “Industry“. This is 53% of all the 38 abstracts in the lens. The length of the bar is an additional indicator of how many articles of a topic are in the lens compared to the total amount of abstracts in the lens.

However, the amount of “Industry” articles is less than half of all the “Industry” articles. Therefore the bar has a lower saturation and is only in a light blue.

The topic “Set-typed Data” is well represented, since 15 of a total of 19 abstracts about the topic are inside the lens. Therefore the bar is highly saturated. “Music” has little abstracts in the lens and is not well represented thus the bar is small and white.



Settings:

Open the settings by pressing the cogwheel button.

Visualization settings:

- The info board and quick task can be switched on or off.
- The info board can be of a static form in the corner or move dynamically as you move the lens.
- If you turn on the cluster labels you can see the labels for each cluster directly on the graph as a text.

Quick Task Settings:

- The quick task can be switched on or off
- Choose between “Summarize” and “Most important fact” as a “Quick Task” (for more information see Functionality section).

Custom Task Settings:

You can choose between four custom tasks.

You can create a **summary**, list **similarities**, show the **three most important facts** or display **controversies** among the abstracts.

The Custom Task can be started by a click on the green button “Explain with AI“. It will always be instructed based on the abstracts that are currently in the lens.

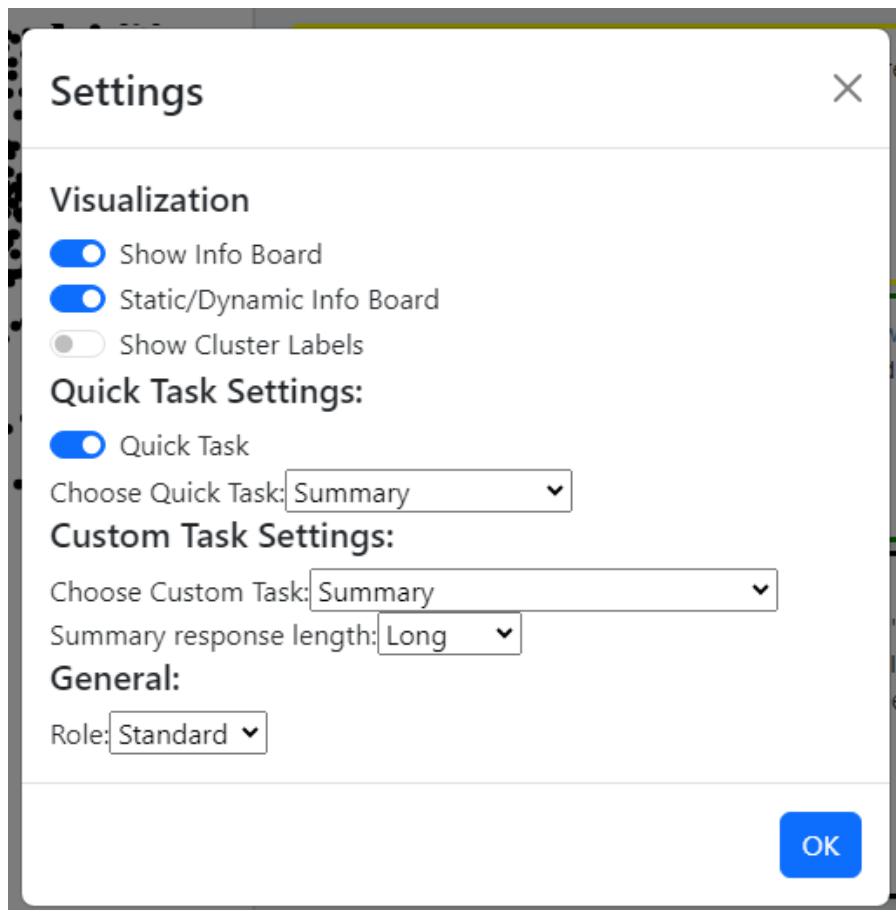
In case you chose “Summary” you have the option to choose between a medium or long response length.

General:

In the general options you can choose between three roles: “Standard“, “Expert” and “Simplistic“. This refers to the quick and user-defined AI explanations. For almost all cases “Standard” should be the role of your choice.

In the “Expert” mode you are identified to the tool as a computer science expert and only very specific keywords will get marked and explained to you. This is only recommended if you don’t want to receive as many links in your response texts as before.

In “Simplistic” mode, the responses are displayed in simple terms and more explanations are given. The explanations are suited for a person that has no knowledge about computer science at all.



The image shows a 'Settings' dialog box with a close button (X) in the top right corner. The settings are organized into four sections:

- Visualization**
 - ☒ Show Info Board
 - ☒ Static/Dynamic Info Board
 - ☐ Show Cluster Labels
- Quick Task Settings:**
 - ☒ Quick Task
 - Choose Quick Task:
- Custom Task Settings:**
 - Choose Custom Task:
 - Summary response length:
- General:**
 - Role:

An 'OK' button is located at the bottom right of the dialog box.

Functionality:

Quick Task:

- Returns a very short response. Based on your selection either a summary of the contents of all the articles that are in the lens or the single most important fact of all the abstracts inside the lens according to the LLM.
- The quick task (if turned on) will be executed **every time** when you drop the lens (provided the lens contains at least one abstract).
- Show more details: In the case of “Summary” a click on “Show more details” will create a longer summary. In the case of “Most important fact” it will create a response with the three most important facts of the abstracts in the lens.
- Workflow tip: It is generally advisable to search for an interesting area by moving the lens around and looking at the generated quick response. A longer summary is already being generated while you read through the short response. When you click on the “Show more details” button you have a shorter waiting time than when manually instructing a long response with the “custom task”.
- Lens ghost: The lens ghost is only active when the quick task is turned on. After you execute the quick task for the second time the lens ghost will appear. It shows the old lens position on the graph and appears as a second static lens in gray. On the text field side you can see two quick task text fields. The green one is the current quick task response and the gray one is the old quick task response. Now you can compare the current set of abstracts with the prior set of abstracts in the lens. Keywords that appear in both texts are marked in bold. Hovering over the old response text will show the old info board of the lens to make further comparisons.

These abstracts primarily discuss **visual analytics** tools and techniques designed to enhance the interpretability, debugging, and optimization of deep learning models across various applications, addressing their black-box nature and promoting human-in-the-loop interaction.

Show more details

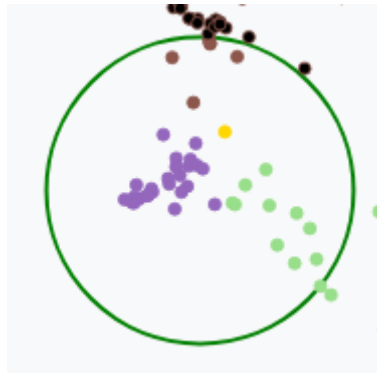


These abstracts collectively highlight the development and application of advanced visualization and **visual analytics** tools for enhancing cybersecurity, **anomaly detection**, **network diagnostics**, and **situational awareness** in complex and large-scale systems.



Abstract Info

By double clicking on a data point with your mouse you can see more information about the article behind it. It is shown in the yellow text field. The data point will be highlighted in yellow. Another double click on it will unhighlight it again.



At first a short summary of the abstract and to which topic it belongs will be shown.

Topological Analysis

The authors propose a novel, provably correct test for determining the parity of Parallel Vectors (PV) points on a cell face, which improves PV-extraction algorithms by ensuring closed curves in continuous fields.

You can extend this to read the full abstract. If the abstract strikes your interest you can click on the link to access the full article.

Underneath the abstract all authors of this article are listed. A click on their names will open another view where you can get more information about them.

curves wherever the input fields are continuous, which we exemplify in extracting ridges and valleys of scalar functions.

Authors: [Tao Ju](#), [Minxin Cheng](#), [Xu Wang](#), [Ye Duan](#)

[Full article](#)

[Back to summary](#)

Author info

This field appears when you click an author name in the abstract info field.

All the data points that refer to articles written by this author are marked with a blue circle around them.

The field displays how many articles the author contributed to and to which topic they relate in a bar chart.

Furthermore it provides a short and a detailed description about what new contributions to the field the person made.

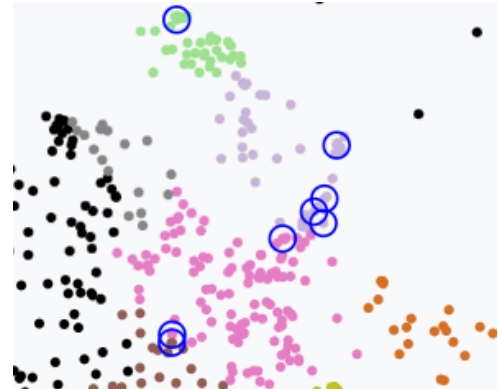
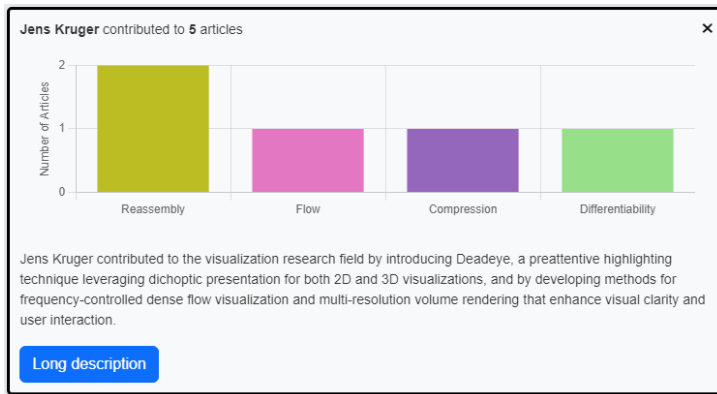
If you remove the field by pressing the “x” at the top the blue circles will disappear again.

Note: The author description has to be taken with a grain of salt. It is not meant as an accurate description of the contributions of an author but rather as a small insight to their work and what *potential* novel contributions they made. Other authors are most of the time also involved in those contributions that are not mentioned in the text.

It is produced by GPT-4o and solely based on the abstract texts of the current data set from 1990-2023 published at *IEEE VIS*.

The contributions of authors with less contributions are over represented in the texts.

This tool should not be used as a way to accurately compare the impact authors have on the visualization field.



Links in texts:

In the “quick response” and “custom response” references and keywords are marked as clickable links.

- References: An article is referenced with its leading author name and “et al.” afterwards. E.g “Wang et al.”. A click on their name lets you jump to the summary of the abstract in the “Abstract Info” field.

Furthermore it highlights the data point relating to the abstract yellow. You can unhighlight it by double clicking the data point again.

- Keywords: All technical terms are marked with a link. You can hover over a keyword to see in which other abstracts this keyword is used.

On the view all data points that relate to abstracts that contain this word are marked with a red circle around them.

A click on the link lets you jump to the keyword explanation field where you can get a definition about the word and can search for it in the graph.

Another key finding is the focus on [user-centered design](#) for [visualization tools](#) that cater to specific audiences, including coaches, analysts, and fans. Systems like OBTracker ([Wu et al.](#)) and Omniculars ([Lin et al.](#)) illustrate [tailored visual analytics](#) that enhance the analysis of [off-ball movements](#) and [in-game behaviors](#), respectively. This approach aims to demystify [game dynamics](#) and provide [actionable insights](#), improving engagement and strategic decisions during live events.

Keyword Explanations:

In the response texts keywords are marked with links. A click opens the keyword explanation field. A short explanation will be shown there. You can do an “Exact Search” for the keyword. All data points related to the abstracts that contain this exact string or the singular/plural form will be marked with a red circle.

Optionally, you can do a “Custom Search” where you can choose the parts of the word for which you want to scan the abstracts for.

Additionally, you can type in any **custom words** for which you want the scan the whole abstract set for.

Underneath the explanation you can see for which words you are scanning the data right now.

If you want to remove the red circles you can click the yellow button on the left or close the whole keyword explanation window by pressing the “x” at the top right.

Photometric Calibration: ×

Photometric calibration refers to the process of adjusting and aligning the brightness and color output of multiple projectors in a display system to ensure uniformity and consistency across the projected images. This calibration addresses photometric non-uniformity, which can cause visual inconsistencies, by correcting disparities in brightness and color, thus enhancing the overall viewing experience.

Photometric Photometric Calibration Calibration

Remove circles Custom Search Exact Search

