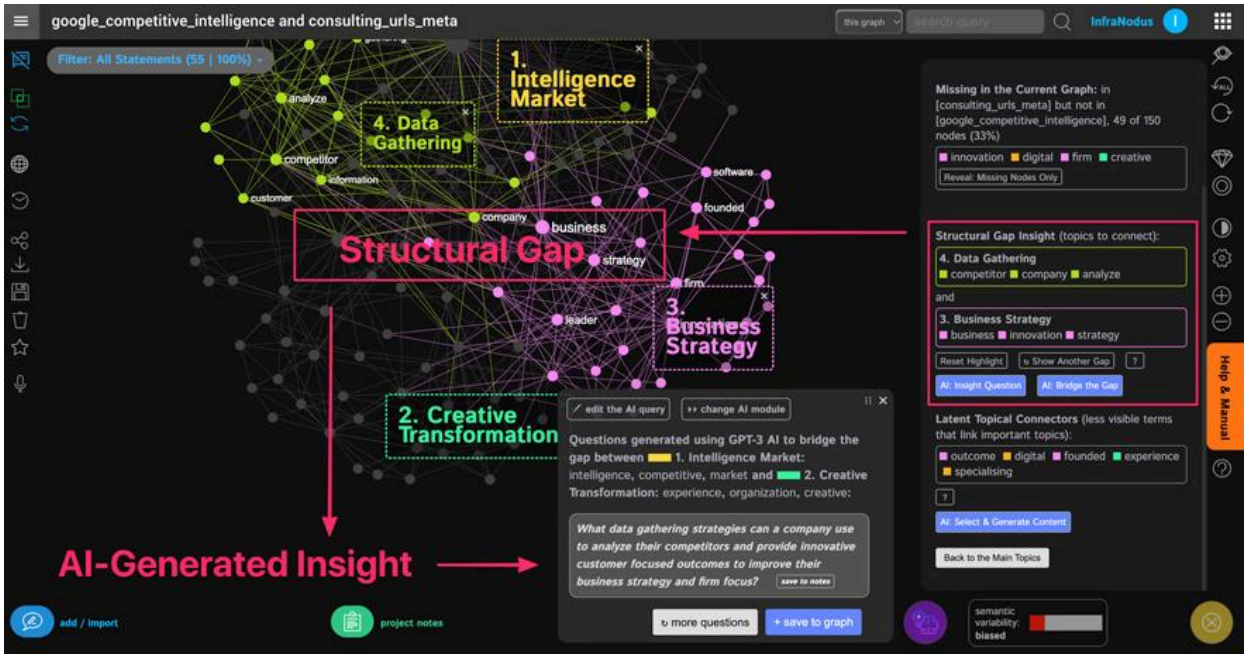
How InfraNodus Works: AI-Enhanced Text Network Analysis

InfraNodus uses a combination of text mining, network analysis, data visualization, NLP, and artificial intelligence to provide insights about any discourse and enhance a research process. Network analysis is the secret sauce that sets InfraNodus apart from other text analysis and AI-based tools. Thanks to this approach, we can design highly relevant prompts for AI systems and use the most advanced language models to derive meaning from text and see how it can be developed further.

Another advantage of InfraNodus is its interactive data visualization, which enables the users to "touch" the texts and explore any narrative in a non-linear way, remove high-level ideas, reveal underlying concepts, and explore the periphery of any discourse. You can zoom in on any group of concepts and see the context where they are used.



Finally, InfraNodus is the first tool that uses GPT AI not only for text generation but also for text analysis. We strongly believe that AI systems are not there to replace humans, but to, rather, enhance our existing processes, increasing their efficiency and making them more scalable. That's why the use of GPT AI in InfraNodus is geared towards helping derive insight through an iterative and interactive research process.

Step 1: Text Network Visualization

InfraNodus represents a text as a network, where the single concepts (or entities) are the nodes and their co-occurrences are the connections between them. Powerful algorithms from network science and graph theory are then applied to this text network to highlight the most influential concepts, topical clusters, and structural gaps between them. Advanced NLP AI algorithms are used to help users interpret the graph, derive additional insights from their data, and develop their ideas further.

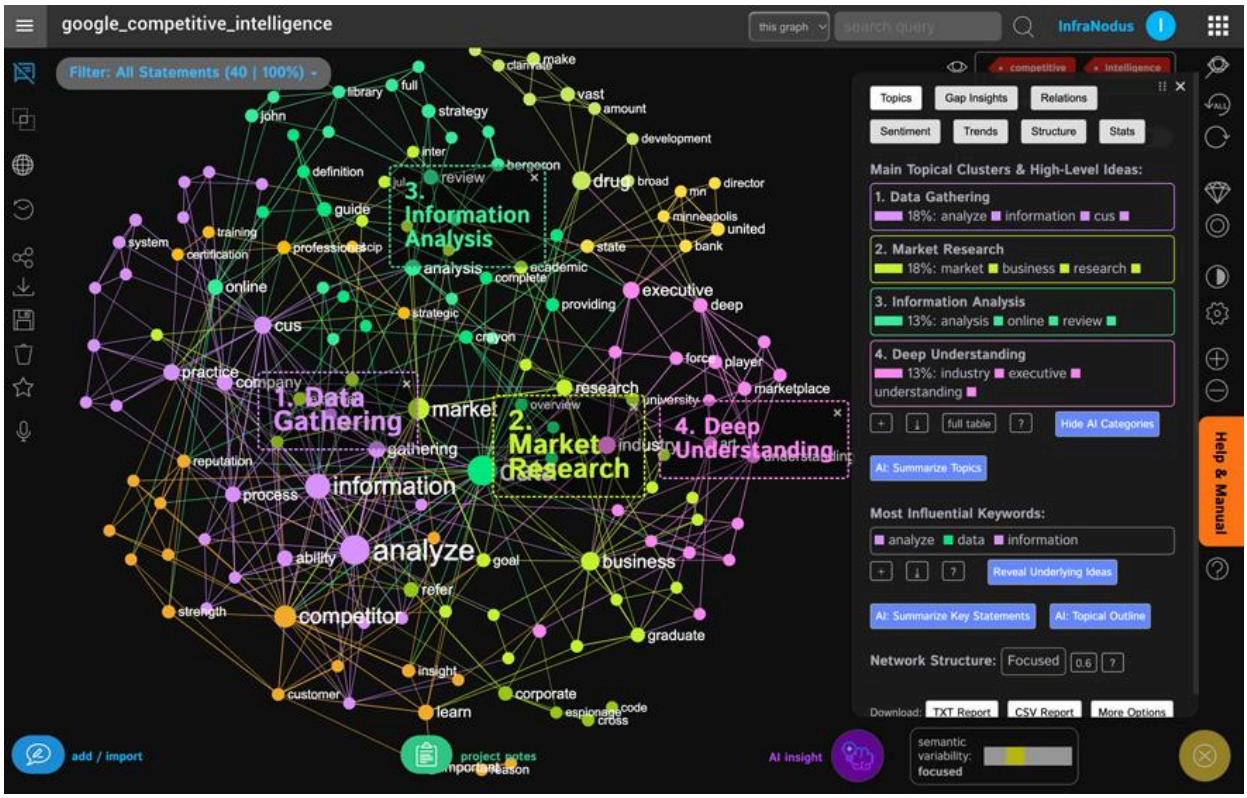
To learn more about the algorithms used in InfraNodus and for citations, please, refer to this peer-reviewed paper: Paranyushkin, D (2019). InfraNodus: Generating Insight Using Text Network Analysis, Proceedings of WWW'19 The Web Conference, www.infranodus.com (ACM library, PDF) as well as the first paper we published on the subject: Paranyushkin, D (2011). Identifying the pathways for meaning circulation using text network analysis, Nodus Labs. (Google scholar), and other Nodus Labs publications.

Step 2: Revealing High-Level Ideas

After you add or import a text, it is transformmed into a sequence of lemmas, the stopwords are removed, and each lemma is represented as a node. The nodes are connected if they appear in a 4-gram window. The closer they are, the higher is the weight of the connection.

Based on this logic, we build a network graph where the nodes are ranked by the measure of betweenness centrality (BC). The higher the measure of BC, the bigger is the node on the graph.

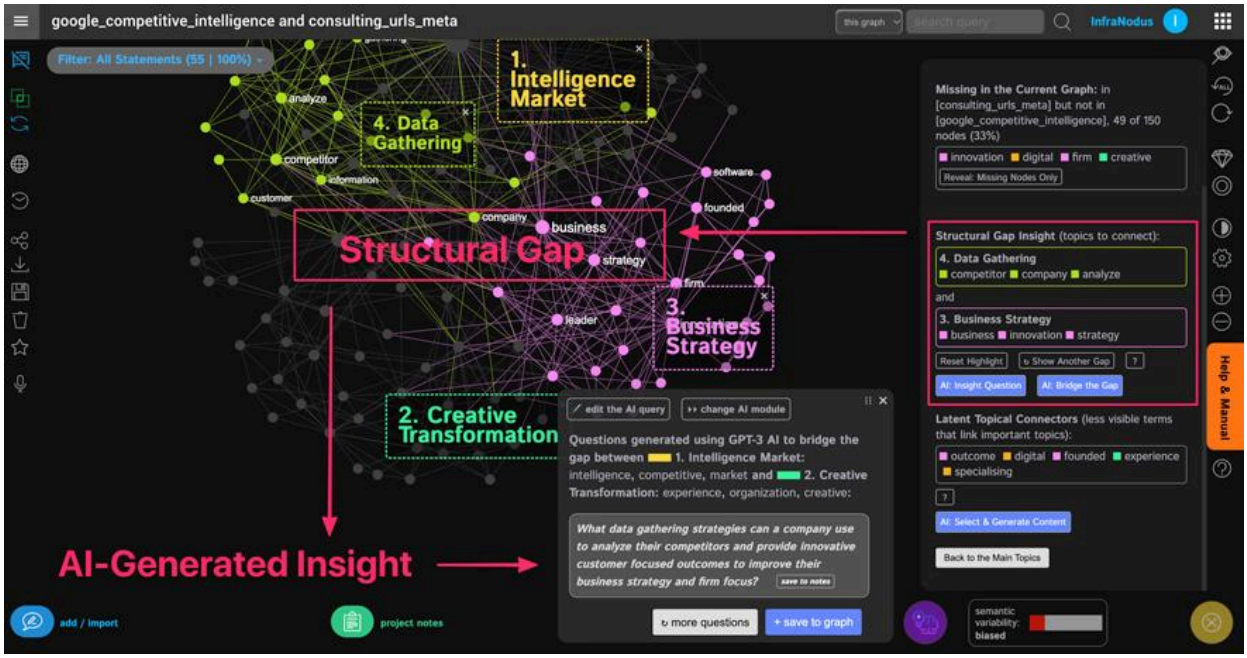
The nodes that appear in the same context more often are aligned into clusters on a two-dimensional plane and have the same color. We use community detection algorithm based on the modularity measure to cluster the nodes into topical groups. These are also shown in the Analytics panel.



At this stage, you can see an overview of any discourse: the most important concepts, topics, and pathways for meaning circulation, helping you gain a better undertanding of any text and deconstruct its dispositif (to use the term coined by Michel Foucault), revealing its rhizomatic structure (Deleuze).

Step 3: Structural Gap Detection

Once we visualize the structure of the graph, we can quickly see what is missing. InfraNodus will then apply a special algorithm to detect the structural gaps: the parts of the discourse that could be better connected. These are identified from finding the topical clusters on the graph that are distinct from one another and have a high distance. We then highlight those gaps, bringing the user's attention to the blind spots in the discourse.



While the structural gap is highlighted in the graph, we also reveal the two topical clusters that could be better connected. The users are then encouraged to think of a possible connection to bridge the structural gap and generate a new research question or an idea. Alternatively, you can also use the built-in GPT AI to generate a research question or an idea for you.

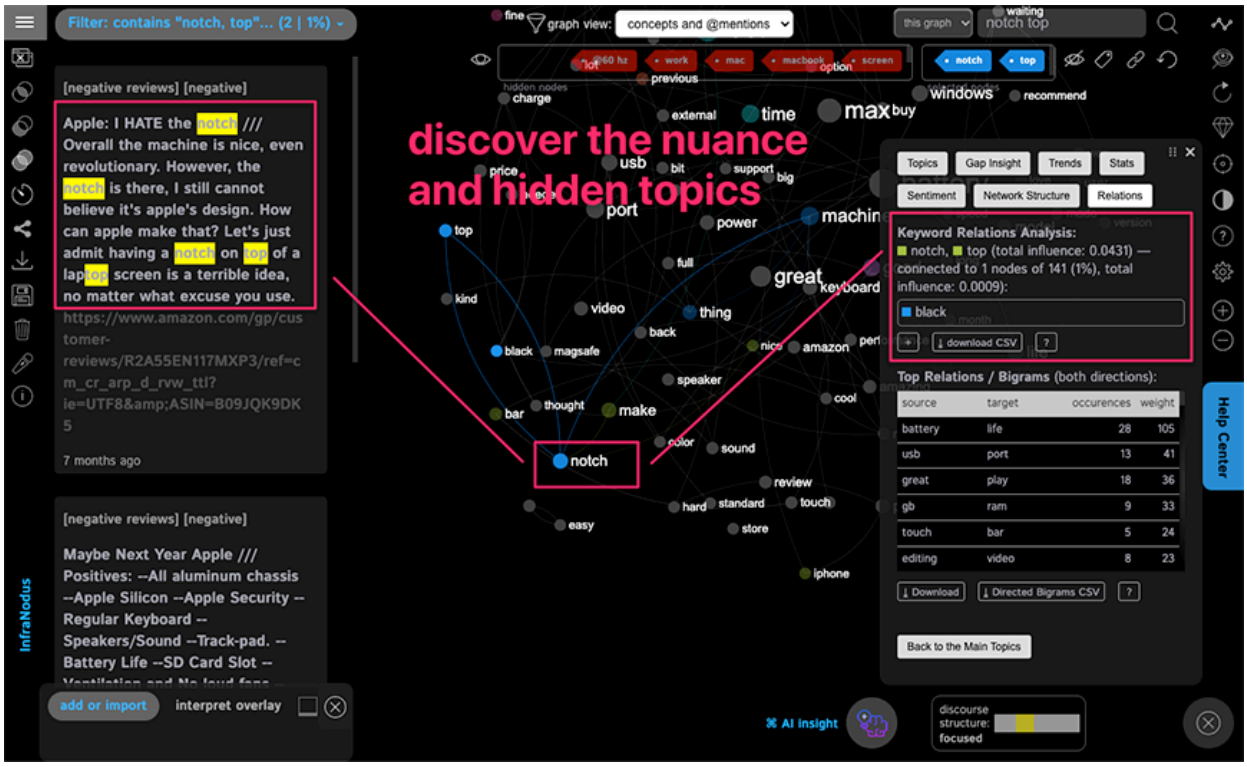
Step 4: Underlying Ideas, Entrance Points, and Periphery

Most of the text analysis tools are too focused on the main concepts, topics, and entities found in text. InfraNodus, on the other hand, is a great tool for discovering the ideas that are not easily visible on the surface.

Part of the workflow that we propose is removing the most obvious concepts from the graph to see underlying ideas around them. This helps reveal hidden low-level ideas that make the discourse specific and different from the rest.

Another interesting feature is the detection of discourse entrance points: the concepts that have high influence per occurrence. These are the ideas that can provide an easy access to discourse as they function as pathways for meaning circulation but, unlike the most influential nodes, are not burdered by "meaning traffic".

Finally, network visualization also helps users reveal the periphery: these are the groups of ideas that are not central to the text itself but play an important role in connecting it to other discourses.



Ecological Thinking: Cognitive Panarchy

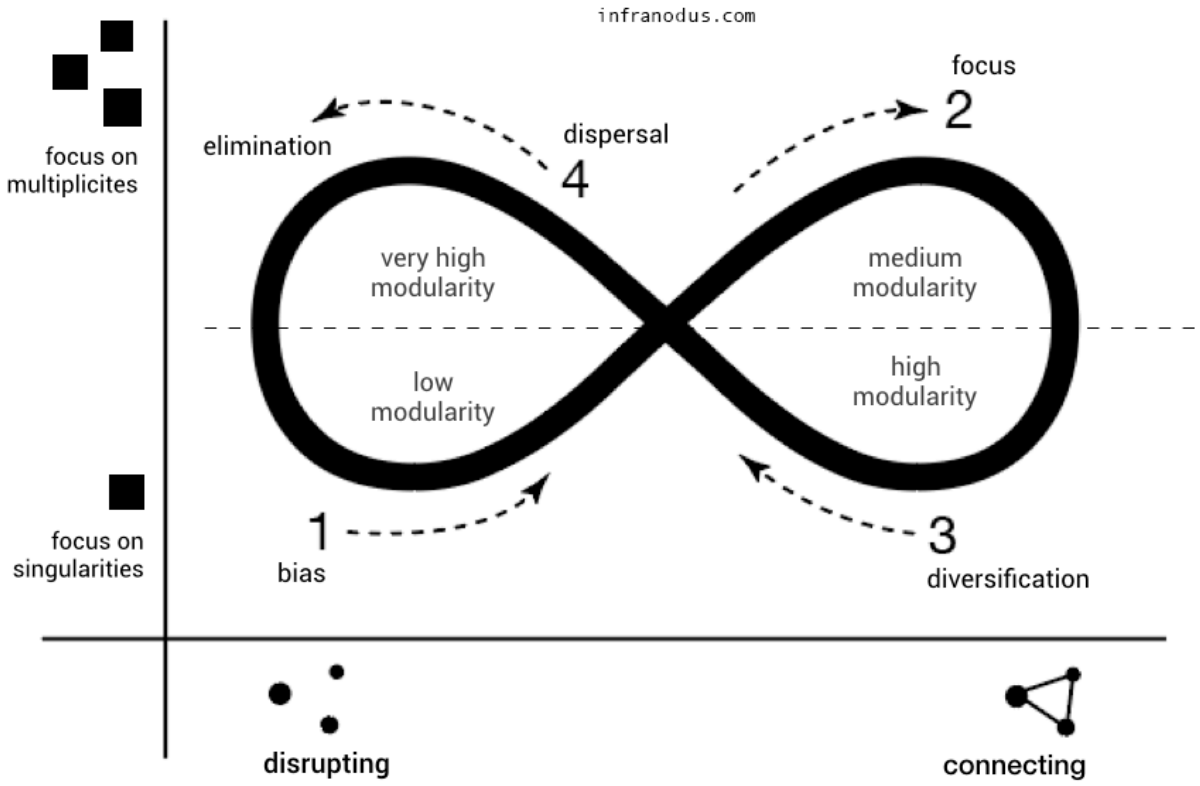
InfraNodus is a tool that promotes ecological thinking, considering any narrative as an ecosystem consisting of interrelated parts in a constant interaction with its environment. It encourages reasoning in terms of dynamic relations rather than fixed solid entities, taking into account the context around any idea and exploring how it could be developed further.

This approach helps you keep your ideas agile and your mind open, maintaining the optimal combination of coherency and disruption in your thinking, writing, and reading.

Ecology of the Mind and Panarchy

Gregory Bateson coined a beautiful term: "ecology of the mind". What is a mind that is ecological? It has the ability to have an overview, but it can also zoom into any idea. It embraces diversity, but it can also obsess over one thing when needed. It can discover the obvious, but it can also reveal the things that are hidden and ponder the gaps that have not yet been bridged. Focused and, yet, adaptive. Rational and poetic.

InfraNodus is a tool that is developed to help you think this way. It is made to promote ecological dynamics and diversity on the cognitive level. InfraNodus visualizes and analyzes ideas as a network, revealing the relations and patterns within them, so you can understand the dynamic complexity of how knowledge evolves and explore the nuances of meaning.



This approach can be visualized using the diagram shown above, which is an adaptation of the concept of Panarchy used in ecology. It conceives of a thinking process as a succession of different types of intents and scales, each suitable for a specific context, set of resources, environment, and task. If we embrace this approach, we will never get stuck and crisis will seem like an opportunity. Periods of growth are followed by periods of saturation and optimization, which help us restructure our thinking and come up with new ideas again.

When we think ecologically, we embrace multiple scales and intentions. When we conceive of a new idea, we need to focus (intent) on something very specific and singular (small scale) [Stage 1]. Gradually, we start connecting ideas (intent towards connecting) and increasing the scale (thinking in terms of topics rather than singular ideas) [Stage 2].

At a certain point, we reach saturation, because we do not have any more resources (attention, time, data). This is the moment when we may be thinking that we are stuck, but

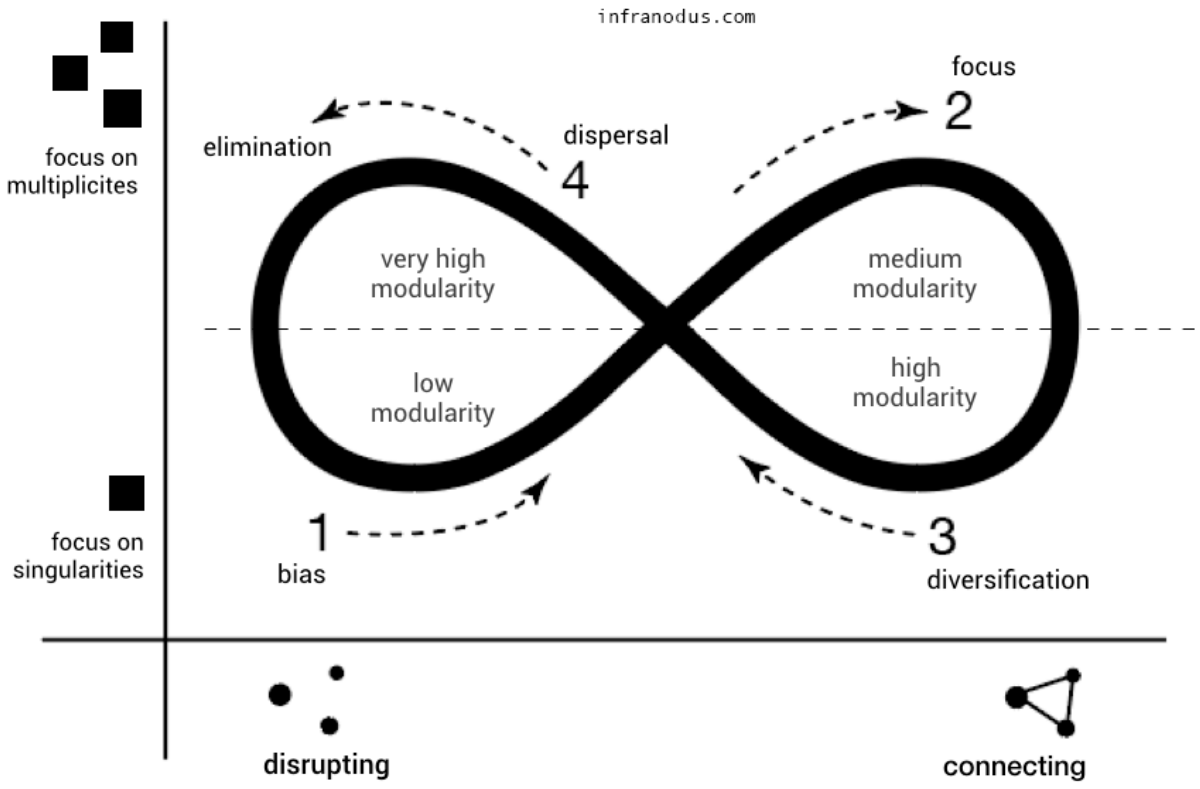
it's also an opportunity to optimize our thinking. We can shed off the old ideas, change scale again, zoom in, and reconnect some cognitive patterns, find the gaps in thinking and come up with new insights [Stage 3].

As a result, there will be space for the new ideas and we can zoom out again and see how the insights that we came up with relate to our general discourse. Alternatively, we may also disperse our old patterns and disrupt the connections to come up with something new [Stage 4].

Finally, we focus on a new idea and start growing it again [Stages 1-2].

Cognitive Variability: Towards more Diverse and Adaptable Thinking

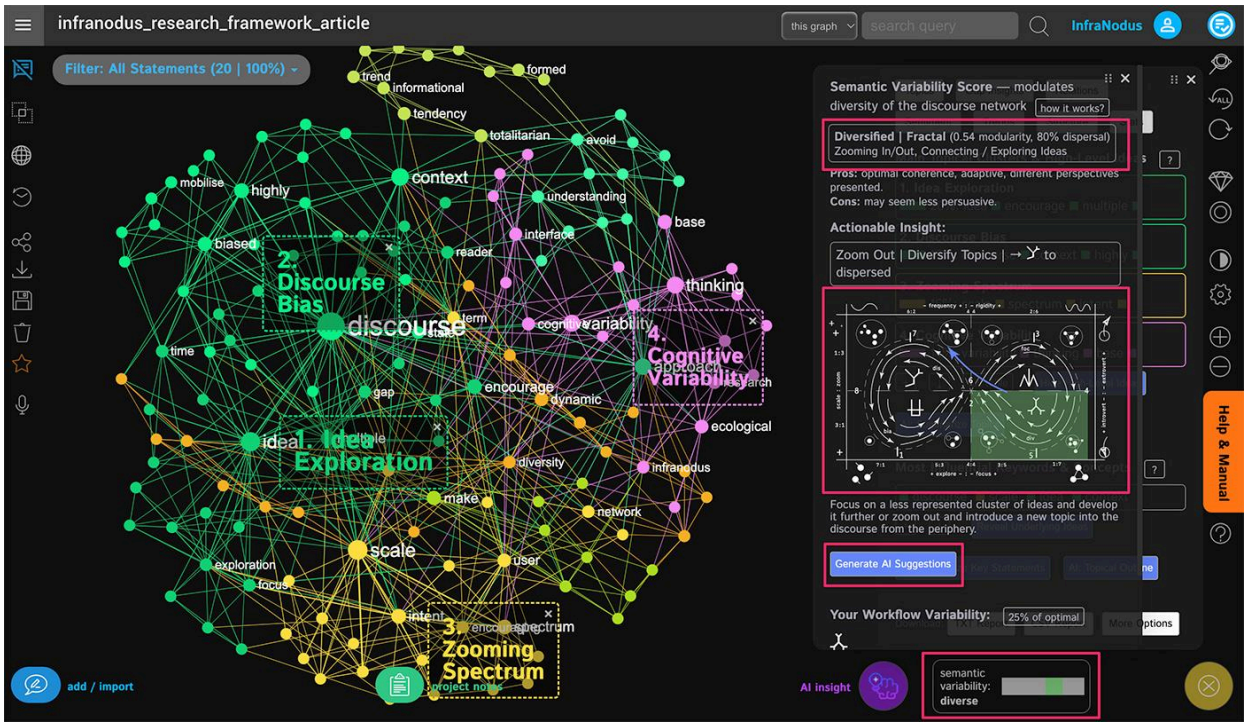
InfraNodus text analysis tool is designed to induce cognitive variability: the ability to approach an idea from multiple perspectives, to discover general patterns but also nuances, to alternate between exploring and making sense. A higher level of cognitive variability makes any discourse more diverse and open, helping avoid cognitive bias and totalitarian thinking. Below we explain how it works.



Cognitive Variability as a Mind Antivirus

Too often people get stuck with a certain idea or a pattern of thinking. InfraNodus can be used as a mind antivirus against obsessive loops — biased ideas, mundane patterns, totalitarian thinking, propaganda, and narrow-mindedness — proposing a kind of thinking that is pan~archic: spanning a range of states and modalities.

Using text network analysis, InfraNodus identifies the structure of the discourse through the measures of network modularity, distribution of influence, and narrative variability. It then uses its algorithms and AI to steer this structure toward a more adaptive and open state.



If the discourse is too biased toward a certain idea, InfraNodus will steer it toward diversification. If the discourse is too dispersed, InfraNodus will steer it toward more focus.

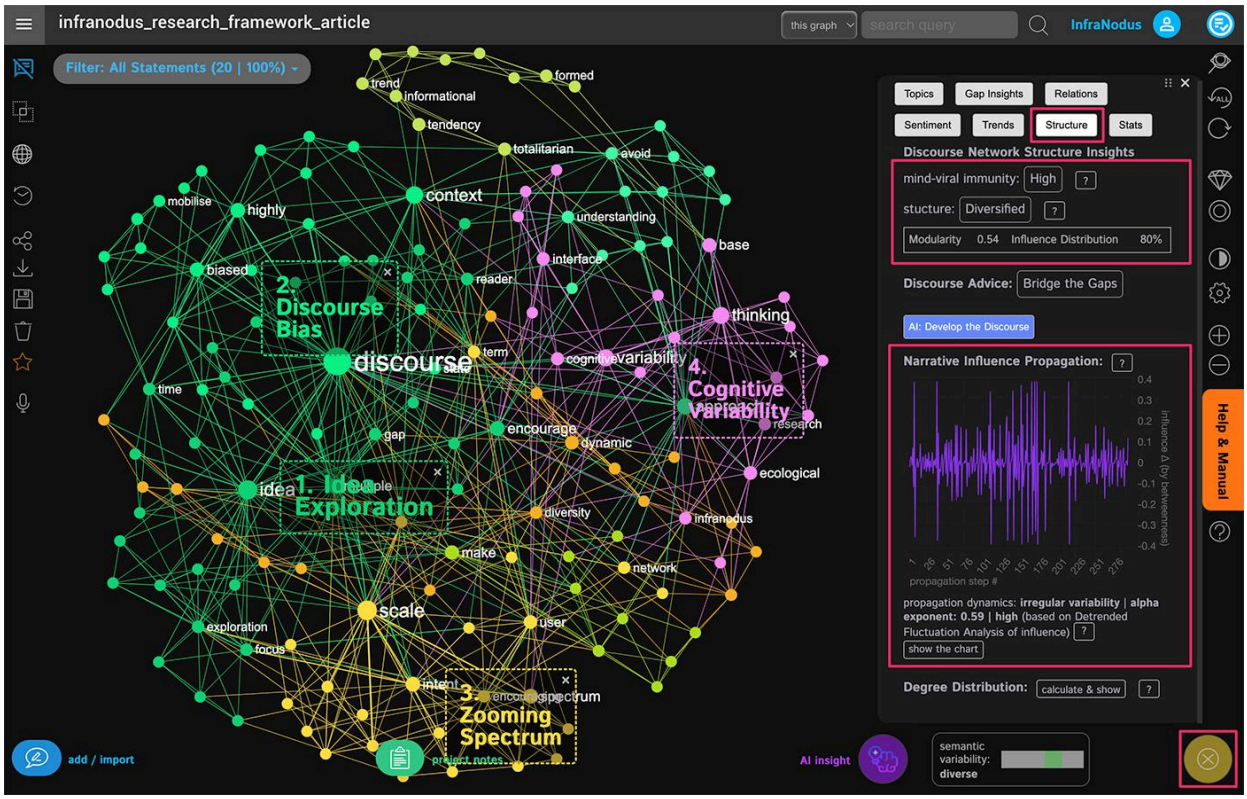
In plain terms, if InfraNodus were a conversational partner (which you can experience with the AI Chat App), it would support a conversation for some time, but then also switch the subject and go on a tangent towards a different topic. After some time, it will bring the conversation back to the central ideas, stay there for a while exploring the nuance, and then shift the subject to a new topic again.

Four States Variability: Detrended Fluctuation Analysis

When we talk about variability, there are different ways to describe it. As we are dealing with discourse, we should take into account both structural and narrative properties.

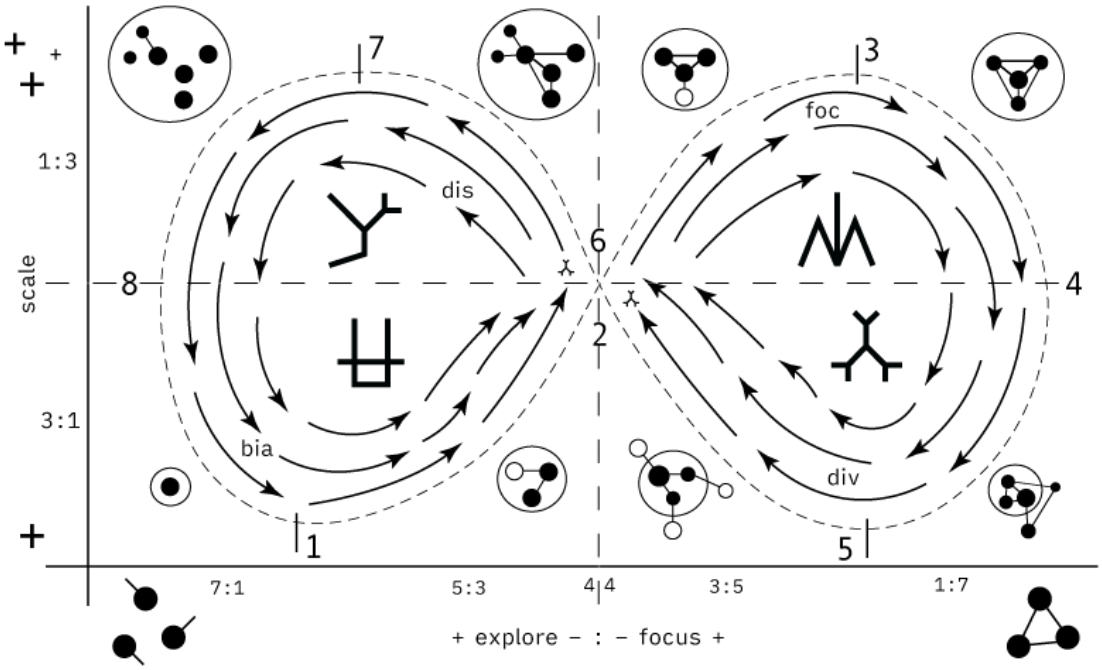
Structural variability can be identified through text network measures such as modularity and entropy. Narrative variability is less intuitive. As we shift from one concept to the next, we should somehow measure the change that occurs as we're reading a text. One possible parameter could be word vector distance in a semantic space. This distance, however, would be based on the global context (all the texts that have been written before) and doesn't take the specifics of text into account. That's why we use the measure of influence (betweenness centrality) of the network's nodes (representing the words). Betweenness centrality shows how important a word is for a specific context. High variability, in this case,

means that we're jumping between the important and peripheral ideas. Low variability means that we either remain in the periphery or reiterating the main concepts only.



Using the approach above, identify 4 states of variability, which are comprised of 8 stages. Each state and every stage correponds to a certain combination of discourse structure parameters as measured through network science algorithms that we use. For example, how densely connected the discourse is, how modular it is, concentration of the main ideas in the main topical cluster, distribution of influence across the discoursive network, degree distribution, dynamics of influence in the narrative.

To represent these 4 states and 8 stages we plot a two-dimensional graph. On the X axis is the spectrum of intent, on the Y axis is the spectrum of scale. The intent can alternate between exploration and focus. The scale can alternate between small scale and big scale. As a result, we get 4 possible combinations of intent and scale. Depending on our position in the graph, we can distinguish more states. Some of the intent / scale parameters are also interchangeable. For example, across the intent scale we can also think in terms of construction and deconstruction.



In order to study variability, we borrowed the terminology used in DFA methodology (detrended fluctuation analysis), which is commonly used for identifying HRV (heart-rate variability) in medicine. This is a robust approach to providing quantifiable data on variability and it also distinguishes four different states (uniform, regular, fractal, complex), which correspond to four different types of variability (repetitive, stationary, self-similar across scale, and non-stationary).

Cognitive Variability Stages

Now that we presented the theory behind the approach, here is how it is implemented into InfraNodus, which identifies 8 stages of discourse evolution (indicated with numbers on the scheme above).

1→2 #bias #growth #vector #exponent:
If the discourse is too biased toward a certain group of concepts (stage 1 to 2 on the graph), InfraNodus will highlight the less represented ideas and propose you to explore them and to connect them (50%/50% explore / focus ratio and 50%/50% zoom in / out ratio at the stage 2).

2→3 #focus #saturation #plateau:
This will bring the discourse toward a focused state (stage 2 to 3). Higher connectedness, focus on multiplicities, increasing the scale, zooming out, linking ideas more than exploring (80% to 20%).

3→4 #focus #conservation #intensification:
Now we can start zooming in again, decreasing the scale, connecting ideas within a smaller area, deepening the focus inside the existing structure (stage 3 to 4 on the graph).

4→5 #diversification #assimilation #release:
Once the discourse has become even more "focused" (stage 4 to 5), InfraNodus will also suggest to develop the specificities, zoom in further, go bigger scale, and, thus, reduce (global) focus and increase the proportion of exploration again (to 20%).

5→6 #diversification #redirection #fractalization:
It will then propose to diversify the discourse by zooming out and exploring more clusters, developing each of them further (stage 5-6: moving between locally related ideas and also jumping across the structural gaps in the graph).

6→7 #dispersion #reorganization:
At some point, we shift toward dispersion by giving more weight to the exploration process, rather than focus (state 6-7: dispersion).

7→8 #dispersion #reset:
Once the scale is large enough, we zoom into the small scale again while exploring until we find a new concept we'd like to develop (stage 7-8).

8→1 #bias #genesis:
We then focus on it and develop it further again (stages 8 to 1).

InfraNodus Research Framework and Methodology

InfraNodus proposes a framework for research, which is based on ecological thinking where it considers any discourse as a dynamic ecosystem of interdependent parts. The objective is to bring those parts into play by inducing variability across two spectrums: scale and intent.

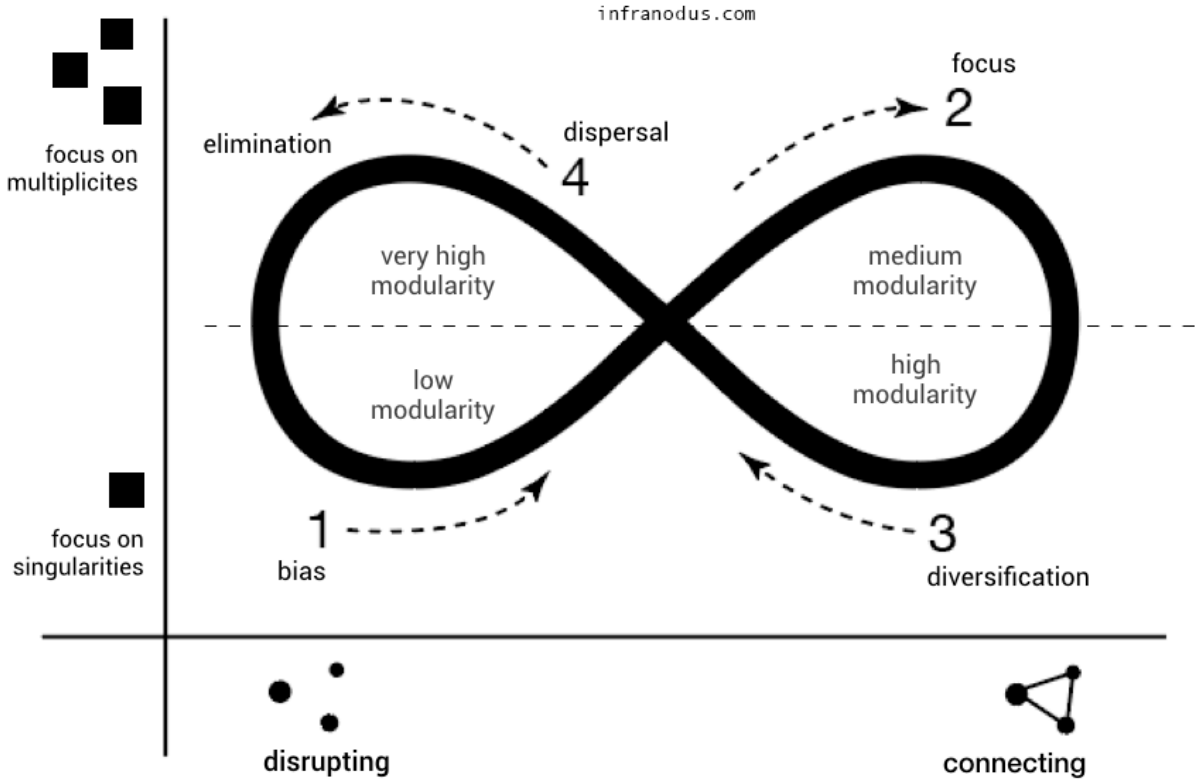
A researcher using InfraNodus framework will be encouraged to become aware both of high-level ideas (big scale) and underlying nuances (small scale). They will also be encouraged to alternate between connecting ideas (focused intent) as well as deconstructing the discourse and looking for new ideas beyond its immediate limits (exploratory intent).

Such approach ensures that a research process remains coherent and, yet, open to external influence. It is inspired by dynamical systems theory, network science, and deconstrutivism. It helps discover patterns, relations, interesting perspectives, new insights, blind spots, and knowledge gaps rather than fixed ideas and conclusive findings.

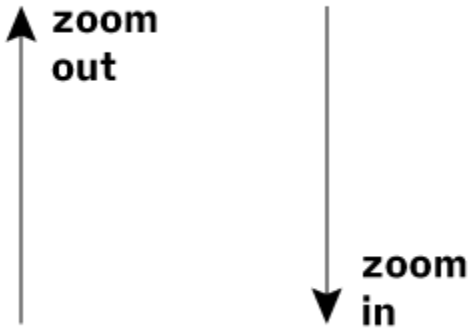
1. Scale and Intent Variability

One of the main tenets of stable ecological systems are diversity and adaptability. These qualities make them robust and open to change. InfraNodus represents a discourse as a network of ideas and then encourages the user to explore it in a way that opens up multiple diverse perspectives.

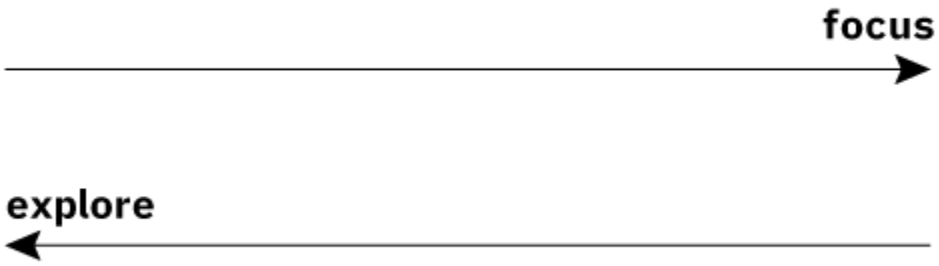
This is done through instilling cognitive variability across mainly two spectrums: scale and intent.



The first spectrum, that of the scale (vertical Y axis), is encouraging the user to shift between different scales: from small scale (zooming in onto specific concepts) to big scale (zooming out to see the whole picture and topical patterns).



The second spectrum, that of intent (horizontal X axis), is encouraging the user to shift between different types of intent: from exploration (exploring periphery and deconstructing ideas) to focus (bridging the gaps, connecting ideas).



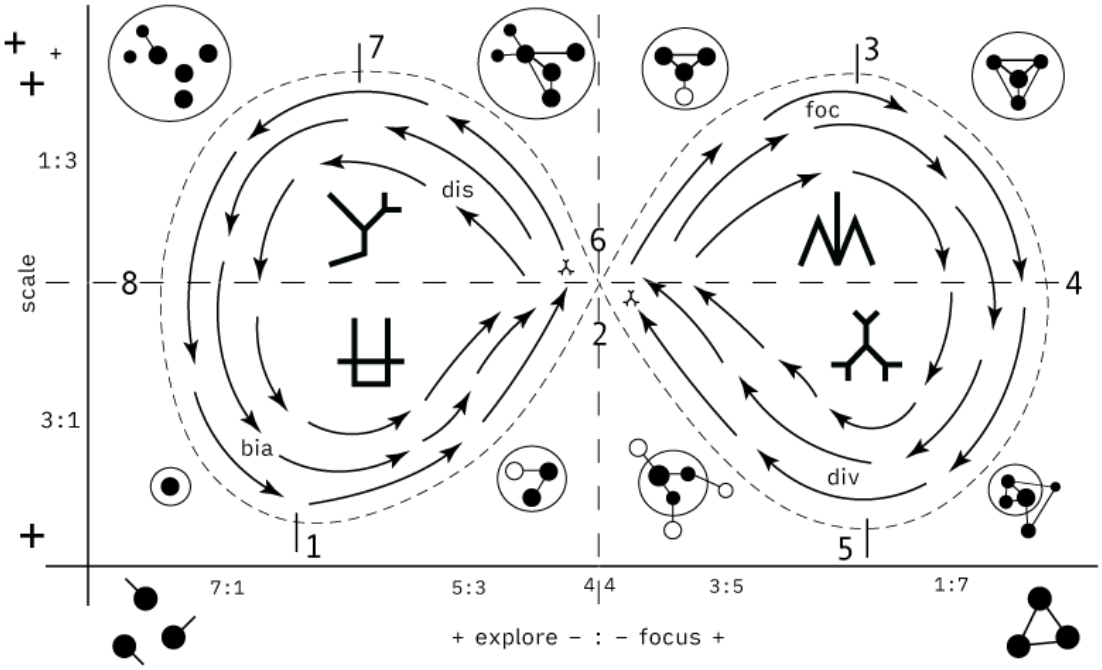
This kind of approach has several advantages that makes a research process, perception, and reading more diverse and adaptable.

2. Advantage: Open-Minded Coherency

First, constant shifting across different scale makes sure that the discourse or our reading of it stays coherent, but not too biased or too dispersed. The ideas are organised into topical clusters that are distinct from one another (diversity) yet, connect on the global scale (coherency).

Second, constant shifting between exploration and focus ensures that there is a constant influx of new ideas into the discourse (exploration) and, at the same time, those ideas also get connected into conceptual clusters that make sense. This ensures that we stay open to new ideas but without losing the track of the key concepts.

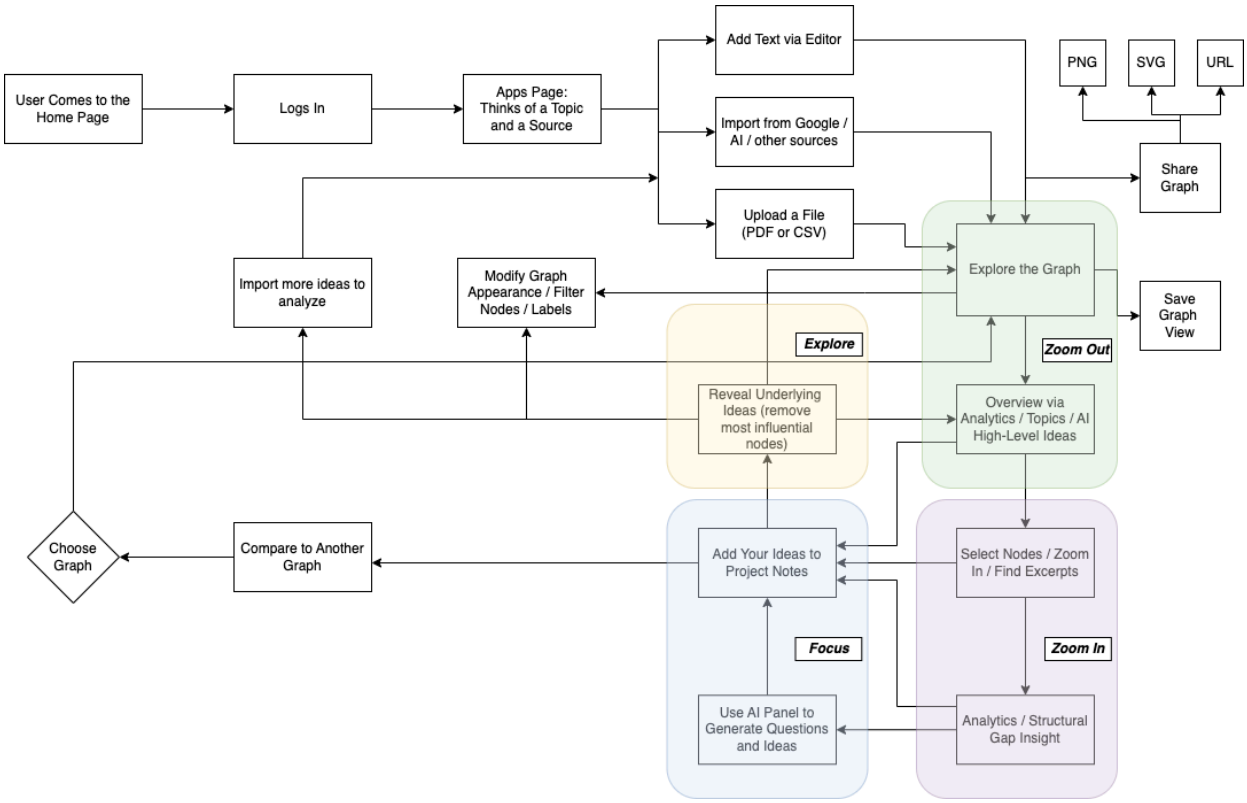
Obviously, this does not mean that a discourse has an "ideal" state. We approach cognitive variability as a dynamic process that is also context-specific. That is why distinguish 4 different states for each discourse, which has a different combinations of the scale / intent parameters.



A discourse is always a part of an ecosystem. There are some environments and situations that require discourses to be biased and some other contexts that require for discourses to be highly dispersed.

3. Implementation

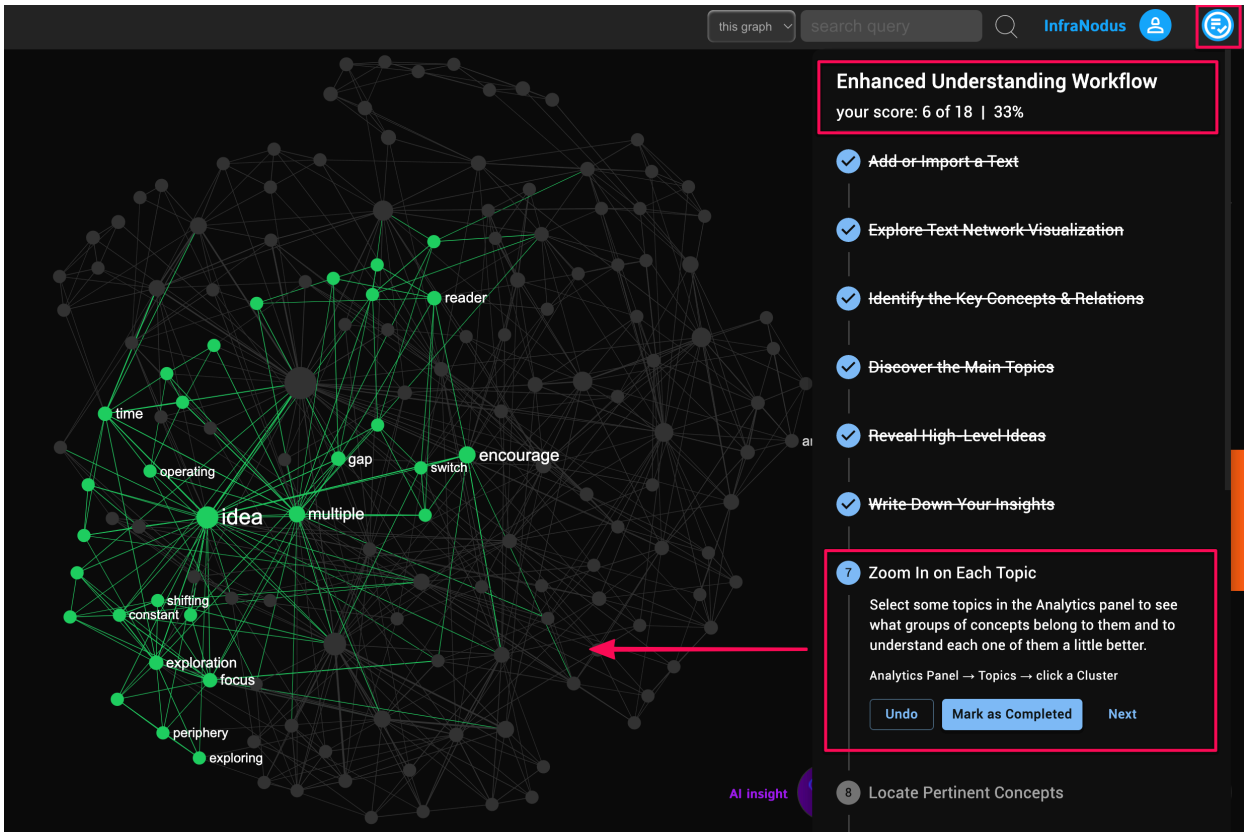
In InfraNodus, the research framework outlined above is promoted through modulating cognitive variability using the interface elements, workflow scenarios, and interface items that encourage the user to switch attention, intention, and scale.



We propose four different workflows through our interface that can be used to:

1. Enhance Understanding of a Discourse
2. Get an Overview of an Idea
3. Refresh One's Memory of a Text
4. Develop a Discourse from an Idea

An example of one such workflow as implemented in InfraNodus is shown below:



Each of the workflows proposes several actions that alternate between scales (zooming in and zooming out) as well as the intent (exploration vs. focus, deconstruction vs. construction).

Our internal studies (link) have shown that such approach leads to a more comprehensive understanding of a topic or text as the reader is encouraged to switch between different perspectives and to look at the ideas presented from multiple viewpoints, thus, enriching their understanding of context and avoiding their built-in bias.

A research framework based on induced cognitive variability helps us escape echo chambers and avoid totalitarian thinking, because it promotes diversity and openness as a dynamic practice in writing and reading.

4. Bias and Poetry

By default, our framework helps to induce an optimal level of coherency. However, this approach can also be used to shift towards higher bias or, on the contrary, towards a high level of dispersion and even incoherency. There are several reasons why this may be useful and interesting.

For instance, when a discourse needs to mobilise people for a collective action it will work better if it is biased. At the same time, a highly biased discourse may also be highly ideological and may have negative effect in the long-term, especially in the contexts that have totalitarian tendencies (e.g. informational trends, media-fuelled frenzy, uniform opinions formed around fear-based narratives).

On the other side, if we talk about poetry, it may be beneficial for it to be highly dispersed as it will leave multiple structural gaps in the meaning, which the reader has the pleasure to fill with their own thoughts and emotions.

In general, however, most situations will benefit from a discourse that combines multiple properties over a period of time: operating at multiple scales (granular concepts that align into big themes) and encouraging multiple intents (from exploration to focus).

5. More Information

To read more about this approach, please, see our research article on Ecological Thinking Framework on Nodus Labs and also the article on Cognitve Variability of Thinking on our Support Portal.

The basic workflow outline is available on our support portal and directly in InfraNodus' interface