

EXPLORING INFRANODUS: A TEXT ANALYSIS TOOL

Irina Tursunkulova¹, Suzanne de Castell² and Jennifer Jenson¹

¹*Faculty of Education, University of British Columbia, Canada*

²*Faculty of Education, Simon Fraser University, Canada*

ABSTRACT

The exponential growth of scholarly publications in recent years has presented a daunting challenge for researchers to keep track of relevant articles within their research field. To address this issue, we examined the capabilities of InfraNodus, an AI-Powered text network analysis platform. InfraNodus promises to provide insights into any discourse, uncover blind spots, and enhance a scholar's perspective by representing text as a network graph with relevant topical clusters and their relations. To understand the tools' effectiveness in analyzing scholarly articles, we used a set of 15 abstracts and 15 full papers. Our findings revealed that InfraNodus could indeed create topical clusters and meaningful patterns from abstracts, but its generated questions and summaries lacked relevance and coherence with the content. A deeper understanding of how the AI operates within the tool would benefit researchers seeking to optimize their literature review processes.

KEYWORDS

InfraNodus, Text Analysis, Literature Review, AI

1. INTRODUCTION

In 2014, Smithsonian Magazine reported that in roughly 28,000 scholarly journals, 1.8 million articles were published each year (<https://shorturl.at/cLO35>). These numbers have been increasing for the past decade. According to Academia (www.academia.edu), there were 684,708 full text papers on digital games and learning in April 2023; 689,912 by August 2023. Google Scholar identified 18,200 scholarly publications on digital games and learning published in 2022. The proliferation of large numbers of articles published by scholarly journals of varying caliber creates unprecedented challenges for scholars aspiring to keep abreast of established and emerging research and theory in their fields of study. It is the expectation underlying that ubiquitous requirement imposed upon scholars and researchers for a comprehensive "review of literature" relevant to one's research question, an expectation that is, quite simply, no longer realistic and no longer realizable. Might the very tools that have created this unmanageable glut of scholarly publication be harnessed to manage and even control it?

2. EXPLORING AN AI TOOL

A positively reviewed AI-driven tool that appeared promising for academic text analysis, InfraNodus (<https://infranodus.com/>) was chosen for this exploratory study. When InfraNodus first became available, its developers positioned it as a text network analysis platform able to "generate insight and new ideas using AI and network thinking" (<https://infranodus.com/>) with capabilities to "overview... any discourse, reveal the blind spots, enhance a scholar's perspective". According to InfraNodus developers, the tool both represents texts as networks, and provides graphical representations of the most relevant topical clusters and the relations between them. It utilizes network analysis algorithms to visualize information as a graph and extract meaningful patterns from it. The InfraNodus website advises that the tool can work with "ideas, articles, books, Google search results, surveys, tweets, spreadsheets", so we were confident the platform could also analyze scholarly articles. Could that kind of tool mobilize AI to sift through a body of texts and identify core concepts and recurrent conceptual relationships within them? We designed a small exploratory study expressly to answer that question, and this paper describes that study and its findings.

Next, still using this one (split up) abstract, we tested the tool’s ability to generate questions using its GPT-3 AI, the third generation “Generative Pre-trained Transformer”. GPT4, now widely in use, is already being progressively integrated, though GPT3 is still used for InfraNodus’ lower subscription levels as of this writing.

The first question that the AI generated was “*What is the most effective approach to teaching that results in high academic outcomes?*” (see figure 3).

To assess the salience and fidelity to the abstract of the AI question generator, we tried the “more questions” feature. Figure 4 illustrates how the next question generated differed from the first.

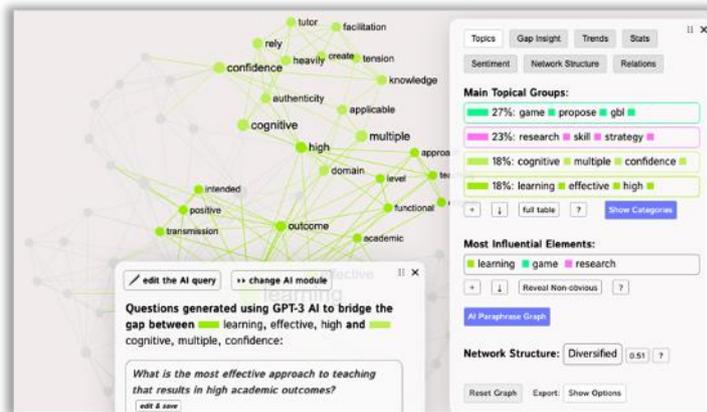


Figure 3. Question generated by InfraNodus based on Abbott's (2019) article

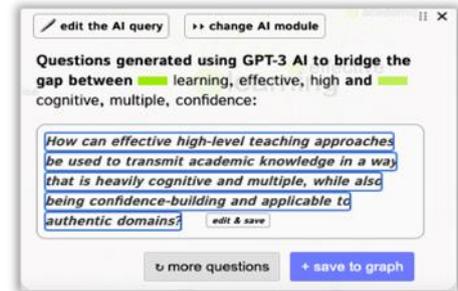


Figure 4. Additional questions generated by the AI

Questions can be regenerated until the AI starts producing the same questions; this happened after the third activation of “more questions”. The questions AI generated were, however, concerningly dissimilar in terms of meaning. For instance, while the first question asked “what teaching method is the most efficient in producing high academic results”, the second question focused on “how effective advanced teaching methods can be utilized to impart academic knowledge in a manner that is highly cognitive and diverse, while also fostering confidence and practical application in real-world contexts”.

Another feature available in InfraNodus is creating summaries of inserted text. When we activated that option, InfraNodus suggested the following summary for the Abbott (2019) abstract: “*This paper investigates constructivist and experiential strategies for effective learning and deep understanding of postgraduate research skills and proposes a game-based learning (GBL) solution.*” This is literally identical to the second sentence in the abstract, not a summary of it. In constructing the summary of the Abbott’s (2019) abstract InfraNodus focused on its first two sentences, including information about the purposes of the study, ignoring details about the study’s methods, findings and proposed solutions, all of which are described in the abstract (see figure 1).

After exploring what InfraNodus could do with a single abstract, we added fourteen more abstracts from GBL scholarly articles. To fit within the tool’s upload restrictions, n=15 articles were manually combined into one pdf, which was then converted to .txt file before being uploaded into the InfraNodus platform. The abstracts were uploaded as one text document, however the tool splits longer abstracts into smaller chunks, making the total number of text boxes 27. Finding no explanation of how many words each text box could have, we counted the maximum number of words in boxes. We found that the maximum number of words a box included at that time was n=156, however one textbox with n=146 words had been split mid-sentence into two text boxes (see figure 5).

Hoping to better understand the criteria for, and implications of, the tool’s automatically splitting longer abstracts into separate text boxes, we reached out to InfraNodus Labs to ask how the separation of an abstract impacts the analysis the AI does.

InfraNodus creator Dmitry Paranyushkin, generously responding to our questions, explained that “*There are not really word limits, but it can't process more than 3 Mb at once (and we don't recommend it, because the graph becomes too big). When the statements are split, the last word of the statement won't be connected to the first one of the next statements. This is the way the analysis is affected*” (Paranyushkin, personal communication. All subsequent italicized quotes are from this same source). However, if the textual units

input to the system have been systematically structurally skewed, preserving the conceptual integrity only of short abstracts, and dismembering (and replicating) longer ones, it stands to reason that the system' outputs will be limited accordingly.

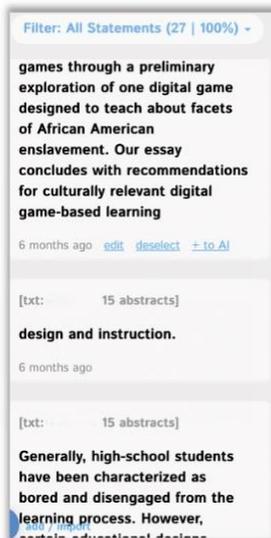


Figure 5. The tool splits articles into smaller texts



Figure 6. The topical network based on fifteen abstracts

Based on 15 abstracts, InfraNodus identified as **main topical groups**: design experience, effect activity, practice development, and data result, as seen on the right side of the image (figure 6). By default, the AI hid some **non-obvious nodes** to show **important nodes**. The user can opt for *showing* or *hiding influential nodes*, which will reconfigure the topical groups displayed. The tool allows the user to manually hide or show specific nodes, and the main topical groups change accordingly. It's worth noticing that single words are the primary semantic currency, not compound nor complex conceptual 'nodes'. And this is important for thinking about limitations, but also about how this kind of software might be most useful. We posed 3 questions to InfraNodus's developer: (1) how are "important" Nodes identified; (2) how are "important" nodes distinguished from "obvious" or "influential" ones; and (3) what criteria are used to make this differentiation.

Developer Dmitry Paranyushkin responded that InfraNodus "uses a betweenness centrality measure to identify the most important nodes. You can also switch in your settings to use the node's degree instead. The most important = influential in our case = obvious". Betweenness centrality is concerned with control over information flow and meaning. He shared an article *Betweenness Centrality: Topical Brokers* published on the InfraNodus blog website, from which we learned that central nodes in a discourse are those topics that most frequently interact with and that most frequently interrupt, other topical connections (Paranyushkin, 2023). But does the AI identify nodes in a discourse, and therefore also topical connections, conceptually or lexically? Word and concept are not homologous, concepts do not map neatly on to linguistic tokens, and not all words are names, so the frequency of a term's appearance may bear no necessary relationship to the importance, influence, or impact of the concept it points to.

By March, 2023 when the same abstracts were reviewed, again, the "Reveal Non-Obvious [nodes]" feature had been changed to "Reveal Underlying Ideas". Despite the update in the feature's name from "non-obvious nodes" to "underlying ideas", in both cases the designated function was to "remove the most influential nodes from the graph to reveal the important topics behind". If "most influential" topics are the same as "most important" topics, then on what basis does the graph change after the "reveal underlying ideas" (previously "non-obvious nodes") feature is applied?



Figure 7. Most influential concepts offered by the tool through "Reveal underlying ideas" feature. Previously this feature was called "Reveal Non-obvious"

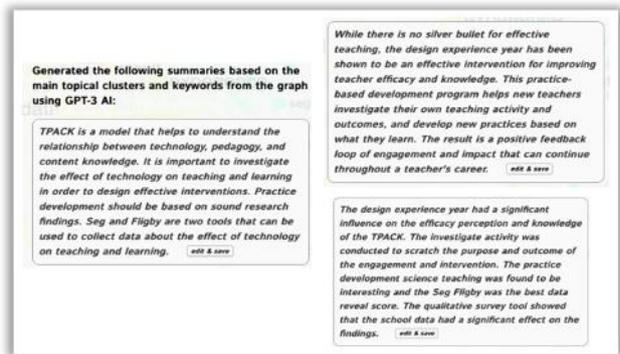


Figure 8. Summaries generated using fifteen article abstracts

As we tested the tool’s ability to create and regenerate summaries of several abstracts inserted as one document, we found at times that the regeneration of summaries led to different phrasing and significantly different meaning from the abstracts being summarized.

For example, according to the InfraNodus summary on the left, Seg and Fligby are considered to be two tools for collecting data. According to the abstract, however, “FLIGBY is a serious game initially developed with the objective of simulating the business management process and the application of Flow theory in a business context” (Almeida and Buzady, 2019). SEG, mentioned in that summary, stands for Serious Educational Games, and while SEG was the primary focus of Annetta et al. (2014), it was not mentioned in any of the other abstracts. TPACK was the framework used in Açikgöl, K (2020) article, but had otherwise no connection with the previous two abstracts. InfraNodus’ GPT-3 AI-generated summaries were looking so different from the original abstracts, we increasingly doubted the validity and reliability of its summaries, and so addressed three further questions to its developers: (1) Through what processes, and using what data sets, is the GPT-3 trained? (2) How does it learn to transform abstracts into summaries? and (3) What is meant by "transforming"? The developer proposed that these were questions properly concerned with GPT-3 AI, and suggested we ask them to OpenAI, the creators of GPT-3, not to InfraNodus.

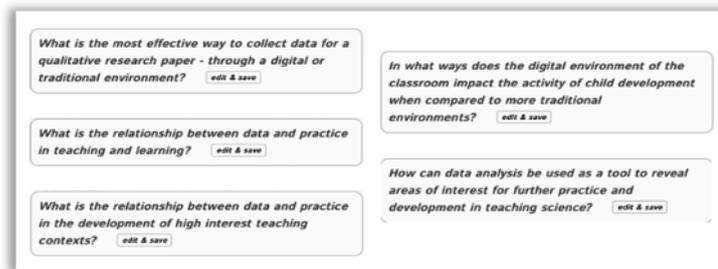


Figure 9. Questions created by InfraNodus based on the 15 abstracts

These considerations impact another feature offered by InfraNodus, that of “question generation”. This option can be used to ‘generate’ useful questions/filters, a function we supposed might inform and guide a more focused and manageable literature review. The questions generated by the AI differed until, after five consecutive regenerations, the same questions recurred. Again, the length of the abstract text and diversity of topics within one long text impacted the variety of question the AI generated: when only one abstract was used, the AI generated only three more or less similar questions before repeating itself, whereas with fifteen abstracts it generated five different questions. When creating these questions based on 15 abstracts, however, InfraNodus did not generate questions relevant to all abstracts, but rather only to (what appear to be) randomly selected abstracts, drawing wording (phrases) from only these abstracts. As with the summary-generation feature, this question-generation option raises questions about its usefulness in the analysis of research materials.

When we asked InfraNodus creator Paranyushkin what the AI bases the questions it generates (and “re-generates”) on, he responded that “It actually shows in the *AI Insight* box — on the basis of structural

gaps”. It makes sense to base questions on structural gaps, since questions mostly concern what we don’t yet know. What isn’t clear is what a “structural gap” consists in, and by what criteria it is identified.

While the Home page of the tool identifies its capabilities as including the ability to identify main topics, and generate summaries and research questions through the use of GPT-3 AI, its developer does not know, or cannot say, how those are accomplished by the GPT-3 core InfraNodus depends upon. Not knowing, or not perhaps not being able to explain to a non-specialist, the algorithms behind the topical maps, summaries and questions supplied by the GPT-3 is not the main issue. The main issue is that the questions and summaries generated appeared only superficially and literally connected to the texts purportedly being summarized and questioned.

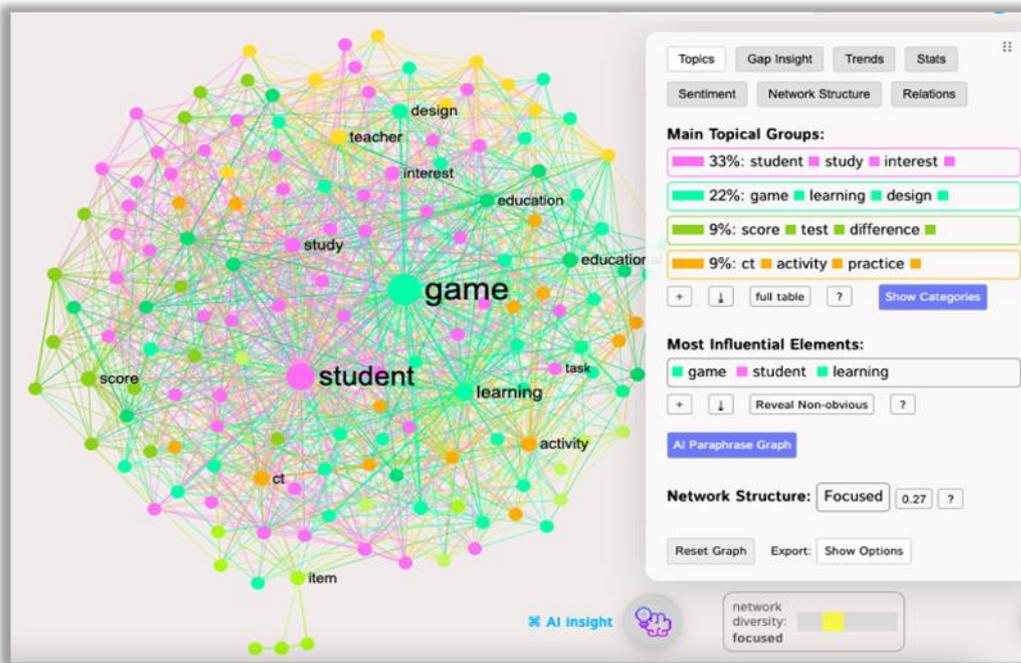


Figure 10. The topical network based on fifteen full papers. Note that ‘ct’ and ‘item’ are identified as ‘nodes’

Initial attempts to upload all 15 articles were unsuccessful due to difficulties in merging and compressing the PDFs of the articles. Visuals and signatures in the files made it challenging to compress the files to the required size. At the end, all files were copied and pasted in .txt format. Inconveniently, that method of file preparation prior to uploading the document is time consuming: over an hour to merge and convert 15 articles for upload. Not only do users need a tool to easily and quickly convert files from PDF to .txt format, it is unclear how a paper’s references, tables and picture descriptions might affect the analysis. This is something that is not mentioned anywhere on InfraNodus website.

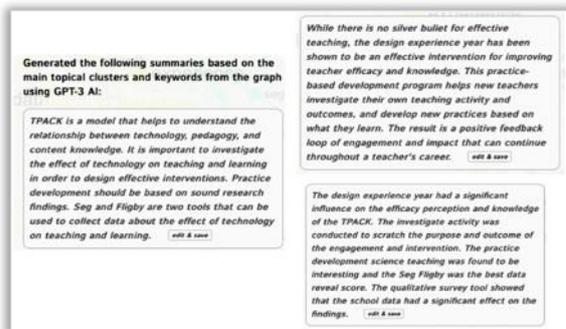


Figure 11. Summaries based on 15 abstracts

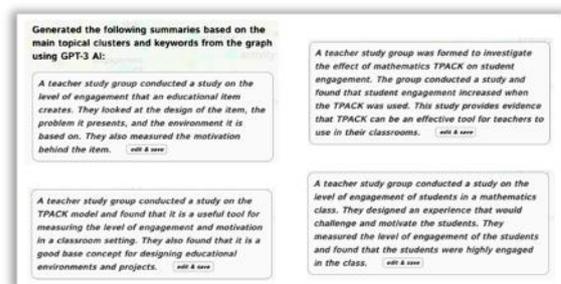


Figure 12. Summaries based on 15 full articles

The analysis of the 15 full articles showed how the inclusion of some references and all graphs, tables, and pictures description might impact the outcome. InfraNodus offered several different summaries of these 15 full articles. The four summaries presented below were randomly chosen from 8 consequent attempts at regenerating. After the 9th attempt, the summaries were repeated. Evidently the more input the tool has, the more questions and summaries it can generate prior to regenerating repetitive ones.

There is a notable difference between the summaries generated from the 15 abstracts, and those generated by the 15 full articles. The summaries generated from the full articles used more exact phrases from those articles. The summaries generated from the abstracts of these articles appeared randomly to combine terms and phrases from different articles.

Further explorations of this tool to see how differences between abstract summaries and paper summaries impact Infra Nodus' knowledge representation (and "idea generation") require comparing summaries of articles with references and without them, to identify how the names of the cited articles impact the summaries and research questions that AI creates, and looking at the differences between articles with abstracts, tables and images and graphs, and articles without these elements.

Network visualization is a key feature of InfraNodus. Its function is to reveal a text's main topics and blind spots. When a set of abstracts encompasses several different topics, with one topic mentioned in many more abstracts than other topics, that directs the AI analysis of the set towards that more frequently mentioned topic, overlooking other less often mentioned topics that may actually be more important. However, frequency of mentions is no guarantee of significance (if we have learned anything from Xtwitter), and this is where a reasonably informed human reviewer has the edge. The tool can only at the very best re-present the text input, including all its intellectual fashion accessories and detritus, things a human reviewer would filter out of a representation of the discursive field.

2.1 How Much Text can InfraNodus Manage?

Our study's last question was concerned with InfraNodus' scalability: what could be the usefulness of this software for sifting through the enormous numbers of publications graduate students and researchers routinely encounter nowadays in undertaking a review of literature, which, in our own sub-field of digital game-based learning, for example, means at least 18,000 full papers. The tool's website states that InfraNodus can perform text analysis, but it does not specify the upper limit to the quantity of text it can analyze. The maximum input varies with level of subscription: Cloud, with 1 Mb upload limit for 9 euros monthly; Pro, with 1.2 Mb upload for 29 euro monthly, and Premium/Enterprise with 3Mb upload limit for 79 euros monthly, with Pro and Premium now using GPT4. For the purposes of analyzing academic texts, however, even the Premium plan offers very limited capability.

The developer explains, "*More is not recommended as the graph becomes too dense and unreadable. This is a tool for visual analysis. If you just want to get a list of keywords from a document you can use another tool but then of course the insight you get is quite boring*". So, is InfraNodus a primarily aesthetic visual "analysis" tool, rather than a tool that might reasonably be expected to help a user manage even a cursory or preliminary text analysis? What, in that case, do its visualizations tell us? When just 15 articles are split into 6,319 text boxes, they are well and truly "deconstructed". Recall the tool developer's explanation: "*when the statements are split, the last word of the statement won't be connected to the first one of the next statements. This is the way the analysis is affected*". Six months after this study's initial exploration of InfraNodus, the tool appears to allow more words within a box, but sentences are still split, and that splitting still influences text analysis. Even as we embrace multi-modality and applaud growing cultural and educational recognition of the value and importance of media and technologies for knowledge-representation beyond textual ones, it is difficult not to conclude from this study's exploration that such visual representations are conceptually problematic, generating kind of surface-structural plausibility at best, and complete nonsense at worst. For example, in an extension of the study reported here, we analyzed a set of self-authored publications hoping to see more clearly how the AI processed data and produced its representations, summaries and questions. The 'table of contents' it created to summarize one of those papers presented this word salad:

- II. Aspect Development Detail Educationally Argue Question Content Infuse*
- A. Overview of aspect development*
- B. Detail educationally argue question content*
- C. Infuse into education*

2.2 Literature Review as “Big Data”?

We conclude this small but in-depth look at ways the academy is being invited to engage AI ‘solutions’, for what promised benefits, and at what actual costs, by circling back to the question that led us here: how do researchers contend with a (*technologically produced*) proliferation of published scholarship that now far exceeds human capabilities to carry out what used to be standard expectation of graduate students and academics: a comprehensive and up-to-date review of literature contextualizing the research undertaken. Scoping out possible uses of AI as a technological solution to a technologically-generated problem, InfraNodus, according to its developers, offered a promising first step: “It provides a clear and actionable visual analysis of any text. Great to use as a writing assistant or to understand a new field or topic. Based on network analysis it can also indicate structural gaps in any discourse to help generate new ideas.” (<https://www.g2.com/products/infranodus/reviews>)

Based on our attempt at using InfraNodus to analyze a set of 15 texts, all game-based learning research papers, we conclude that there appear, at this point, reasonable grounds for caution about the kinds of understanding and the quality of new ideas such a tool enables, raising questions about its usefulness in generating and graphically representing the (dynamic) structure of a conceptual network analysis. What are we actually looking at, and what is its validity as a representation of the discursive network?

It needs always to be borne in mind how rapidly technological change can happen. In this developer’s case, the expectations set out initially for the software have been both modified and extended, the tool itself has been fine-tuned, and this rate of research and development is bound to continue. The question that matters most, however, is which things can be modified (like implementing a way to ensure whole sentences remain in the same text box, or multiple textboxes accommodating a single sentence are treated as one sentence, or removing references from papers and providing a separate way to analyze these), and which problems are *technologically* insurmountable, just as the viral proliferation of academic ‘information’ is insurmountable by humans.

What can we do that a machine cannot and, much more importantly, what **MUST** we do because, otherwise, a machine can think it can? If the AI can only connect concepts that utilize the same words, then it is almost certainly missing out on connections that articulate the same or similar concepts in different words. Given the almost limitless existence of synonyms in language, this is a massive limitation. The whole point of AI is that it is supposed to be able to ‘understand’ the meanings of words and even images, but if it cannot connect corresponding ideas expressed through variable lexicon, how is it more than a glorified Google search.

According to Korab (2021), InfraNodus now readily accepts pre-configured data such as Twitter feeds, Google trends, RSS News feeds, and research papers abstracts and titles. However we have not yet been able use InfraNodus efficiently to input “new”, minimally pre-processed, data (specifically academic research papers), in a way that differentiates and manages elements standard in academic texts, such as reference lists, footnotes, charts and graphs and images, even if that’s just to separate them from the text ‘body’ prior to its analysis. And, to this point, although the tool appears to allow larger numbers of words in a text box, it still fragments sentences to fit its words into length-limited textboxes, and texts remain fragmented and disconnected. “Words” here operates more like “names”, and appear to be treated as elements conceptually disconnected from communicative context (even at a sentence level)---except when, inexplicably in our own closer look, an entire sentence is selected and quoted from among its textual relatives in service of ‘summarizing’ the text as a whole or, equally inexplicably, concepts are juxtaposed in ways that appear conceptually random or ‘off-kilter’, whether grammatically or semantically or both. The tool appears to combine together terms and phrases from different abstracts, themselves appearing randomly selected, and as was shown in Figure 10, included as topical ‘nodes’ terms like ‘ct’ and ‘item’ and ‘graph’ that surely appear in the text, but have nothing to do with its meaning. It’s not clear how “influential nodes” are identified and quantified, beyond word frequency, or whether internodal ‘relationships’ such as “betweenness centrality” (Paranyushkin, 2023) depend upon more (and other) than textual proximity and co-occurrence.

3. CONCLUSION

This exploration used InfraNodus as a beginner’s way to assess the capabilities AI tools more generally might offer us to contend with a humanly insurmountable problem. That problem resulted from, and would not have been possible without, AI, in its earlier processing/storage machine incarnations. Researchers now

must contend with a burgeoning and indeed, viral pandemic of publication, including publication of academic research and scholarship that, even with the most cursory ‘scoping out’ kind of reading, exceeds a human lifespan, let alone a doctoral fellowship. It seemed reasonable to look to AI for help, however we found InfraNodus neither accurate nor therefore useful. It proved, as this study attests, severely limited in its ability not only to identify and represent conceptual networks across a large number of texts, but as well even within a single text, leaving us with serious reservations about the kind of knowledge-representation any such GPT-enabled ‘big data’ analysis of academic research and scholarship would, or could, produce.

REFERENCES

- Abbott, D., 2019. Game-based learning for postgraduates: An empirical study of an educational game to teach research skills. *Higher Education Pedagogies*, vol. 4, no. 1, pp.80-104.
- Abdusselam, M.S., 2020. The Effect of Serious Games on the Awareness of Energy Sources Consumption. *International Technology and Education Journal*, vol. 4, no. 1, pp.36-42.
- Açikgöl, K., 2020. The Effect of Technological Pedagogical Content Knowledge Game Activities Supported Micro-Teaching Practices on Preservice Mathematics Teachers' Self-Efficacy Perception Levels. *Acta Didactica Napocensia*, vol. 13, no. 2, pp.157-173.
- Acosta, M. M., & Denham, A. R., 2018. Simulating Oppression: Digital Gaming, Race and the Education of African American Children. *Urban Review: Issues and Ideas in Public Education*, vol. 50, no. 3, 345–362.
- Admiraal, W., Huizenga, J., Akkerman, S. and Ten Dam, G., 2011. The concept of flow in collaborative game-based learning. *Computers in human behavior*, vol. 27, no. 3, pp.1185-1194.
- Ak, O., & Kutlu, B., 2017. Comparing 2D and 3D game-based learning environments in terms of learning gains and student perceptions. *British Journal of Educational Technology*, vol. 48 no. 1, pp. 129–144.
- Akpınar, Y., & Aslan, Ü. (2015). Supporting Children’s Learning of Probability Through Video Game Programming. *Journal of Educational Computing Research*, vol. 53, no. 2, pp. 228–259.
- Aleven, V., Kay, J., Rowe, J. P., Shores, L. R., Mott, B. W., & Lester, J. C., 2011. Integrating Learning, Problem Solving, and Engagement in Narrative-Centered Learning Environments. *International Journal of Artificial Intelligence in Education (IOS Press)*, vol. 21, no. 1, pp. 115–133.
- Almeida, F., & Buzady, Z. 2019. Assessment of Entrepreneurship Competencies through the Use of FLIGBY. *Digital Education Review*, vol. 35, pp. 151–169.
- Altanis, I., Retalis, S., & Petropoulou, O., 2018. Systematic Design and Rapid Development of Motion-Based Touchless Games for Enhancing Students’ Thinking Skills. *Education Sciences*, vol. 8, no. 1, p. 18.
- An, Y.-J., & Cao, L., 2017. The Effects of Game Design Experience on Teachers’ Attitudes and Perceptions regarding the Use of Digital Games in the Classroom. *TechTrends: Linking Research & Practice to Improve Learning*, vol. 61, no. 2, pp. 162–170.
- Annetta, L., Vallett, D., Fusarelli, B., Lamb, R., Cheng, M.-T., Holmes, S., Folta, E., & Thurmond, B., 2014. Investigating Science Interest in a Game-Based Learning Project. *Journal of Computers in Mathematics & Science Teaching*, vol. 33 no. 4, pp. 381–407.
- Anyaeibu, R., Ting, W., & Li, Y., 2012. Serious Game Motivation in an EFL Classroom in Chinese Primary School. *Turkish Online Journal of Educational Technology - TOJET*, vol. 11, no. 1, pp. 154–164.
- Arnold, U., Söbke, H., & Reichelt, M., 2019. SimCity in Infrastructure Management Education. *Education Sciences*, vol. 9, no. 3, p. 209.
- Asbell-Clarke, J., Rowe, E., Almeda, V., Edwards, T., Bardar, E., Gasca, S., Baker, R. S., & Scruggs, R., 2021. The development of students’ computational thinking practices in elementary- and middle-school classes using the learning game, Zoombinis. *Computers in Human Behavior*, vol. 115.
- James, K., Horst, R., Takeda, Y. P., & Morales, E., 2020. The Patch: An artful syn (aes) thetic mapping of linguistic data through collaborative digital/analogue literacy processes. *McGill Journal of Education*, vol. 55, no. 3, pp. 641-665.
- Korab, Peter (2021) InfraNodus: An Excellent Tool for Textual Data Analysis <https://towardsdatascience.com/infranodus-excellent-tool-for-textual-data-analysis-2b4839e6cd10>
- Paranyushkin, D. (2018). Measuring discourse bias using text network analysis. *Towards Data Science*. <https://towardsdatascience.com/measuring-discourse-bias-using-text-network-analysis-9f251be5f6f3>
- Paranyushkin, D. (2023). *Betweenness centrality: Topical brokers - Nodus Labs Support Center*. Nodus Labs Support