UDC 004.05

Cybersecurity in AI-Driven Casual Network Formation

Dmytro Lande[1], Anatolii Feher[1], Leonard Strashnoy[2]

[1] *NTUU KPI, Kyiv, City, 01001, Ukraine*
[2] *UCLA, Los Angeles, 90001, U.S.*

Abstract

The paper describes a methodology for forming thematic causal networks using artificial intelligence and automating the processes of their visualization. The presented methodology is considered on the example of ChatGPT, as an artificial intelligence for analyzing the space of texts and building concepts of causal relationships, and their further visualization is demonstrated on the example of Gephi and CSV2Graph programs. The effectiveness of the disaggregated method in relation to traditional methods for solving such problems is shown by integrating the means of intelligent text analytics and graphical network analysis on the example of the problem of data leakage in information systems and a selection of news clippings on the selected cybersecurity topic.

Keywords: ChatGPT, Causal networks, Artificial intelligence, Subject matter models, Graph visualization, Data leakage, Cybersecurity

Introduction

Generative AI and large linguistic models like ChatGPT are increasingly prevalent, with common applications in machine translation, text annotation, and question generation for education. OpenAI's ChatGPT, a Generative Pre-trained Transformer (GPT), harnesses natural language processing for diverse user requests [1].

Rapid development of the described technologies, huge opportunities are opening up in the extraction of basic semantic concepts, noun entities, and their relations through the use of ChatGPT in factual systems, in particular, in medicine, business, economics [2], and more. Intelligent chats integrate with external systems, such as geographic information systems [2], graph and network analysis and visualization systems [3]. In particular, the authors in [4] show how to form networks of connections of characters in literary works, and networks of subject areas with public-private connections.

This paper describes a methodology for forming causal networks by repeatedly accessing the ChatGPT system, as well as visualizing and analyzing these networks using Gephi (gephi.org) and CSV2Graph, graph structure visualization programs with a free license [5], [6]. The Comma-Separated Values file (CSV) format is used to upload data to the Gephi environment, so all requests to ChatGPT will be accompanied by a requirement to support the selected format.

The creation of causal networks offers a chance to delve into the contextual and semantic analysis of chosen subjects and their associated text domains. The primary challenge in performing this type of analysis using causal networks is the development of these systems, which traditionally demands significant resources and expert input. However, the method we propose relies on artificial intelligence, mitigating the aforementioned issues.

The proposed method lays the foundation for accurate and simplified automated semantic statistical analysis of large data sets. In turn, it finds practical applications in various fields, including, but not limited to, intelligence, investigation, journalism, analysis, and scientific research. Because of this, the work is based on the study of the topic of data leakage in information systems as an applied task for a cybersecurity analyst. It is noted that the scope of data use can also extend to a variety of custom text sources, such as news, articles, posts, surveys, reports, program codes, and event logs.

1. Forming a network based on simple hierarchical access to ChatGPT

The main focus is on quantitative semantic research, which aims to develop an effective method for creating clear and stable semantic stationary relations between the most frequent and valuable entities in texts for further analysis. In this case, entities act as minimal semantic units of value, and their corresponding pairs replace the exact logic of the text.

The design of queries, or prompts, given to ChatGPT helps to identify relationships in the collected information on a topic. The corresponding prompt polishes the initial linguistic units of the text space extracts pairs of semantic entities from them and summarizes them in the form of a "virtual expert" represented by artificial intelligence [6].

Considering the problem of data leakage in information systems, form a corresponding query to ChatGPT to provide the reasons for this phenomenon known to it. Therefore, the concept of "data leakage" should become the central node of the future network of semantic entities. Successful development of such a query will determine the second level of the hierarchy - the relationship of concepts to the phenomenon of data leakage, i.e. will reveal its causality.

Based on the defined entities of the second-level hierarchy, it is possible to build the corresponding semantic relationships of the following levels, opening up the network space to find relationships between many causal processes. Such a multilevel network represents the greatest semantic correlation, but the multilevel hierarchy and the definition of its derived concepts are limited to infinity, so in the example of the method's application, a three-dimensional system was chosen to obtain a visual picture of a casual data leakage network in information systems that is not strictly hierarchical in structure. Queries of the next level in the hierarchy will be related to the concepts given in the answer and will look like the primary query format given in Attachment 1.

→List the causes of **data leakageage** in cyber security. The reason is to use no more than three words. The results should be presented in the format "cause;**data leakage**". Each such entry - from a new line

Attachment 1: Request to the ChatGPT system

By offering ChatGPT to work on a certain query, will get a set of reasons for the primary concept on the topic of data leakage, i.e. the first level of the hierarchy. The ChatGPT system can help with the formation of the content of a CSV file on the selected topic of the query for research, i.e., adherence to the format - a table with the corresponding names of concepts separated by semicolons shown in Attachment 2.

human error; data leakage
weak passwords; data leakage
insider threats; data leakage
misconfigured systems; data leakage
phishing attacks; data leakage
unpatched software; data leakage
malware infection; data leakage
social engineering; data leakage
third-party access; data leakage
stolen devices; data leakage
…

Attachment 2: Response to a request to the ChatGPT system

Depending on the collected set of entities, it is already possible to apply different research methods, especially the graphical approach, when the sets are expressed through automated graph visualization. Such visualization serves to better understand the defined relationships and their weight [6] between entities.

The basic approach can be characterized as building an appropriate network model adapted to pairs of semantic units. The result is the identification of interconnected causal relationships, weighted accordingly, to solve the hierarchical map of the semantic graph. In this context, links of lower strength appear as smaller components associated with clusters of objects, while stronger links are transformed into more significant network elements. The described approach also facilitates the study of semantic topology and allows combining and comparing the correlations underlying the data and their networks, which directly facilitates the work of analysts.

With the help of quantitative analysis, it is possible to evaluate the relations that are most important, weighted relations. This helps us to represent the most frequently mentioned entities and reveal the meanings behind the text, which reveals a base of semantic mean units, where more weighted pairs are claimed as more valuable and less weighted pairs as less valuable. For the selected topic of data leakage in

information systems, you can see which contexts directly correlate more with the source node and which ones do not, and accordingly which ones semantically have more influence and which ones have less.

Combined in a single CSV file, the virtual expert's answers in the form of ChatGPT are uploaded to the Gephi program for analysis and visualization of entity pair relationships. After uploading the obtained data to the Gephi system, select the size of the nodes proportional to the degree of the number of adjacent connections and, dividing the network into clusters according to the criterion of its modularity, get a visual graph of the selected semantic connections on the topic. In accordance with the methodology, the identified influential nodes of this network are shown in Figure 1 (the largest Out-Degree), they include human error (5), social engineering (4), weak passwords (3), and phishing attacks (2).

assuming that with the increase in iterations of queries to the virtual expert, the qualitative characteristics of the corresponding network will increase.

2. Forming a network based on a hierarchical appeal of a swarm of virtual experts to ChatGPT

At different points during text processing, the ChatGPT system can give different answers, both correct and, from the point of view of human logic, quite "reasonable". Each such answer can be perceived as a response from a virtual expert [3]. Assuming that by generalizing the answers of a set (swarm) of similar experts, one can get a more complete and accurate answer.

Through repetitive queries across various hierarchical levels and amalgamating responses

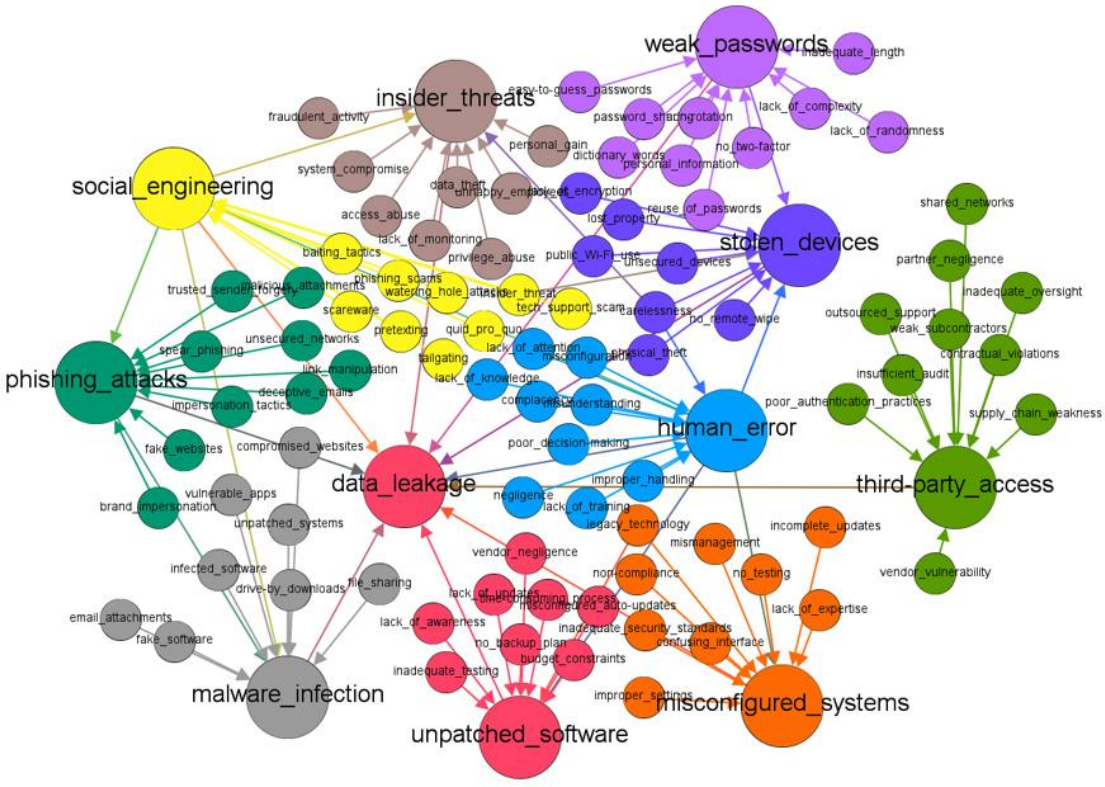

Figure 1. Directed primary causal network obtained by the simplest hierarchical access to ChatGPT

Obviously, the formed network is weakly connected, and incomplete, and the concepts presented in it may not accurately reflect the causes and consequences of data leakage in information systems due to the fact that pairs of semantic primitives are obtained as a result of one iteration of the artificial expert's query,

from artificial experts, a virtual expert team is formed. Using the Gephi program for visualization, this process enhances the graph of interconnected semantic pairs, and it can be refined iteratively with virtual expert input until it meets human expert standards.

Thanks to the chosen approach, the number of important concepts has increased compared to

the previous case, and the most influential nodes of this network are shown in Figure 2, (the largest Out-Degree), these are Human error (7), social engineering (4), weak passwords (3), phishing attacks (2), unpatched systems (2), insider threats (2).

occur less frequently than a given threshold will not be taken into account.

In accordance with the approach for a given topic, the constructed network did not consider concepts that occurred less than 2 times, the most influential nodes of this network are shown in Figure 3, (the largest Out-Degree), these are

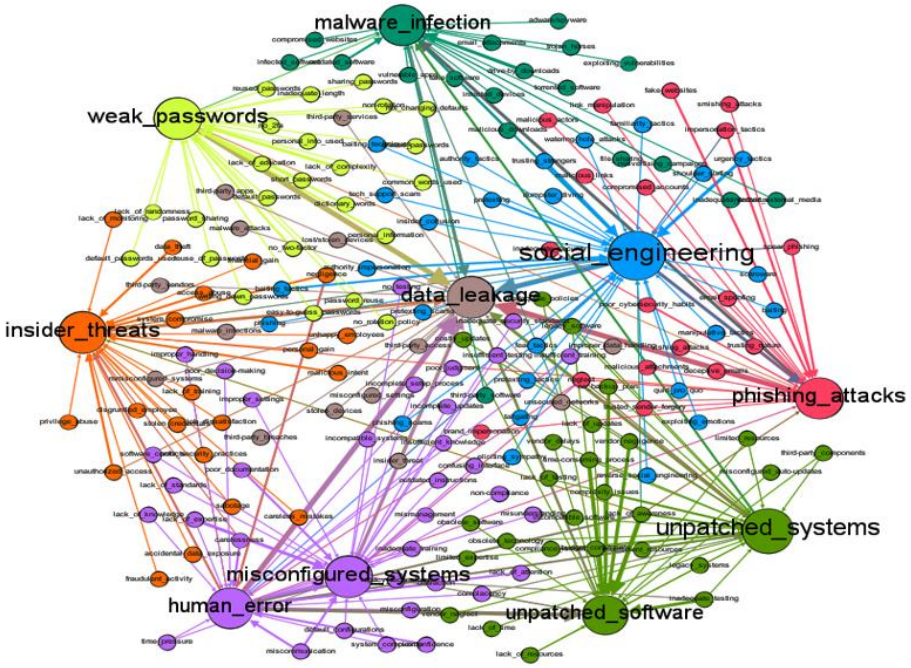

Figure 2. The directed full causal network obtained by hierarchical access of a swarm of virtual experts to ChatGPT

Human error (5), social engineering (3), phishing attacks (2), unpatched systems (2).

3. Network formation based on the generalization of hierarchical access of a swarm of virtual experts to ChatGPT

At different points during text processing, the ChatGPT system can give different answers, both correct and, from the point of view of human logic, quite "reasonable". Each such answer can be perceived as a response from a virtual expert [3]. Assuming that by generalizing the answers of a set (swarm) of similar experts, one can get a more complete and accurate answer.

The graph formed in the previous example, while having relatively high completeness of the concepts of causal relationships by topic, may contain inaccurate information or be mistakenly issued by ChatGPT when processing individual queries. Assuming that the probability of a virtual expert issuing biased semantic units is low, proposed an approach that considers building a network based on the most frequently proposed semantic pairs, where concepts that

4. Network formation based on user-defined data

The examples are currently exploring revolve around innovative approaches where semantic entities or pairs of entities are directly extracted from the well-trained Large Language Model (LLM) system. Specifically, this extraction process is focused on the topic of data leakage within information systems, with ChatGPT being the primary reference model. What sets these examples apart is their subsequent visualization through graph representation using the Gephi program. This comprehensive approach, which combines entity extraction from advanced language models with semantic graph visualization, emerges as a highly effective solution for addressing complex challenges related to data leakage topic.

The methodology employed in this specific context is widely adaptable, particularly when multiple generative systems, as described, are simultaneously engaged in extracting semantic

pairs of entities for in-depth text analysis. This collaborative approach enhances our understanding of the data leakage landscape. Nevertheless, it is crucial to acknowledge that, in certain application scenarios, a variety of limitations can manifest, depending on specific circumstances. These limitations encompass a range of factors and challenges that should be considered:

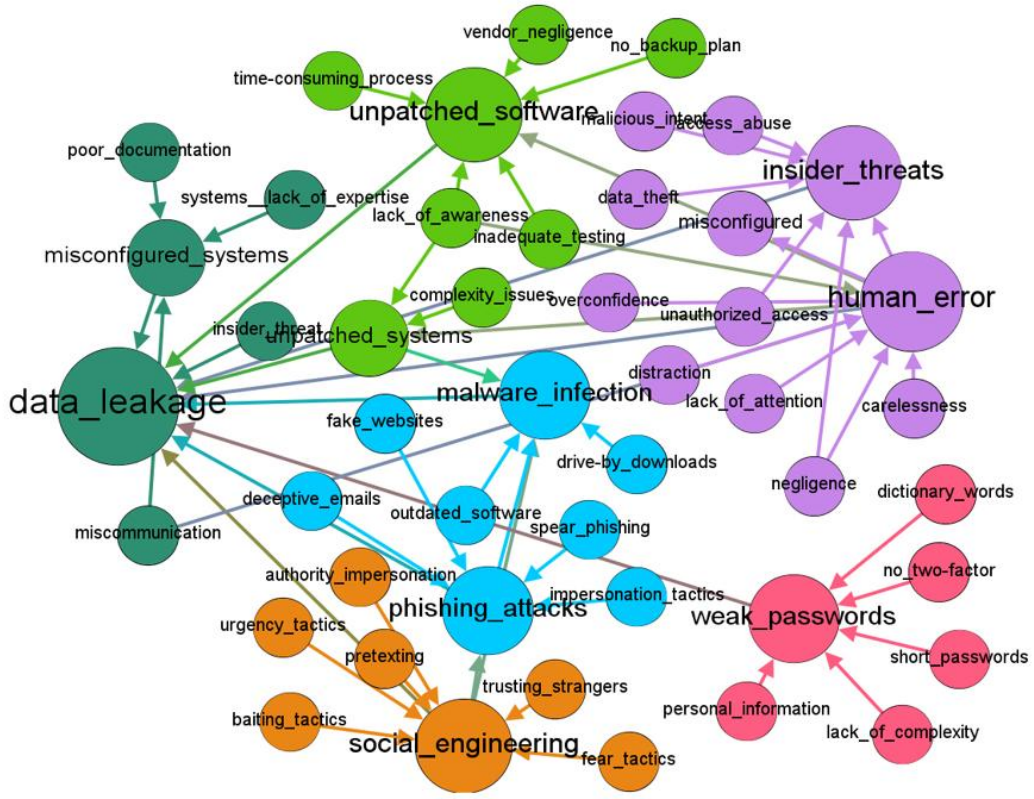

Figure 3. A directional causal network obtained by generalizing the hierarchical appeal of a swarm of virtual experts to ChatGPT

1. the relevance of generative systems, the actualization of models does not always occur in real-time, so the actual entities may not be included in the scope of consideration, and, accordingly, in the non-inertial network model;
2. information limitations of trained generative models, i.e. situations where network analysis requires the use of a set of own data that is not covered by the trained model, such as documents, reports, program codes, logs;
3. anonymity and privacy, compromise of the analyzed topics or own data sets for analysis through data leakage, not a rare phenomenon in the modern space;
4. inconvenience and limited access, the need to install selected visualization products, limited work, especially if there is a need to work with mobile devices, new operating systems, or under restrictions on the installation of third-party software;
5. difficulty in mastering, the need to delve into the theory of building connected graphs and the features of the Gephi software, mastering dozens of parameters, graph stacking modes, clustering.

Considering the aforementioned details, the process can be accomplished by submitting own datasets for analysis as text spaces to locally deployed Large Language Model (LLM) systems or secure enterprise solutions that offer API access.

In accordance with the specifications of the chosen artificial intelligence system, these datasets are fragmented, with answers sequentially linked to construct a collaborative network while adhering to token limitations. It's worth noting that the challenge of limited access to visualization tools can be addressed through online systems designed for graph analysis and visualization, such as Gephi Lite, as well as web-based platforms like WebGraphViz, which can assist in comprehending their functionalities and complexities.

The examined scenarios underscore the practical relevance of developing and deploying a dedicated software component tailored to the chosen methodology for extracting semantic units and subsequently visualizing them to analyze semantic relationships. Typically, specialized graph description formats like GML, GraphML, Pajek (NET), and GraphViz (DOT) are used. However, for subject matter analysts, a more straightforward and visually intuitive format is often required. This format entails the representation of entity names (nodes) combined in pairs, where each pair signifies a graph edge with a designated direction, typically going from the first node to the second or vice versa.

In the context of generative models producing these pairs of entities, a simplified and highly visual CSV format was chosen. Unfortunately, this format is not supported by the previously mentioned visualization systems. To address this challenge, the authors have devised a program based on the Application Programming Interface (API) library of the GraphViz system. This program forms the foundation of the CSV2Graph service, as depicted in Figure 4, which is currently accessible on the internet at https://bigsearch.space/uli.html.

This presented service has proven to be highly effective in analytical research, facilitating the visualization of expansive models across various subject areas, represented in numerous languages.

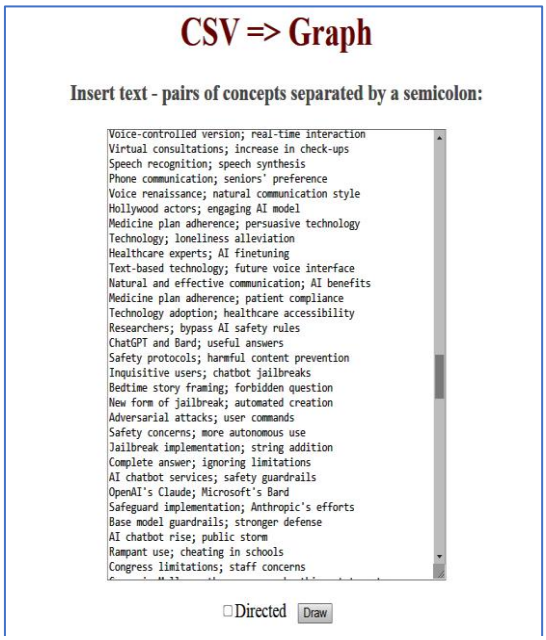

Figure 4: The interface of the CSV2Graph program

This service offers a comprehensive functionality for performing initial analysis and graphical representation of data in CSV format. In this format, each record consists of a pair of entities separated by a semicolon. To utilize the service, users input their data into a designated text field and then proceed to select the desired graph type (directed or undirected). By activating the "Draw" button, the system generates and displays the graph.

The generated graph is arranged based on the importance of elements, with nodes and edges appropriately colored to enhance visual clarity. The thickness and direction of edges are determined in accordance with the input data. The graph layout is achieved using methods integrated into the GraphViz API, and the resulting image is stored in Scalable Vector Graphics (SVG) format.

The choice of SVG format holds several advantages. Firstly, it facilitates easy integration into HyperText Markup Language (HTML) and eXtensible HyperText Markup Language (XHTML) documents, ensuring compatibility across all major web browsers. Additionally, SVG is an open standard for defining two-dimensional graphics using the Extensible Markup Language (XML), which means it is not proprietary and stores data in a text-based format. This format also supports animations and interactivity, enhancing the possibilities for presenting and exploring the generated graphs.

Moreover, the SVG format enables the implementation of hyperlinks within the image by its interactivity feature. These hyperlinks are designed to direct users to search forms of popular search engines like Google, Google News, Bing, and Bing News based on the semantic concepts contained and represented within the nodes and edges of each graph. This functionality enhances the utility of the generated graphs by providing easy access to relevant external information and utilizing required visually intuitive formatting for analysts.

The program cell and developed methodology can efficiently construct semantic networks using our proprietary data spaces, specifically tailored to the chosen topic of information systems data leakage. These networks effectively address the limitations mentioned earlier. To illustrate this, we've included an example application in Attachment 3, featuring a selection of pertinent news reports on the topic. This application demonstrates how the generated semantic network can be employed to create meaningful

insights and connections within the domain of data leakage.

***

Insurer fined $3M for exposing data of 650k clients for two years

By Bill Toulas

The Swedish Authority for Privacy Protection (IMY) has fined insurer Trygg-Hansa $3 million for exposing on its online portal sensitive data belonging to hundreds of thousands of customers.

Trygg-Hansa is an insurer for individuals, private companies, and public organizations, and also an asset management and investment consultation firm. IMY initiated an investigation on the firm after receiving a tip from a Moderna Försäkringar (now part of Trygg-Hansa) customer, who had discovered it was possible to access the insurer's backend by following links available on quotation pages sent to clients.

…

Bleeping Computer 2023.09.04 01:51

https://www.bleepingcomputer.com/news/security/insurer-fined-3m-for-exposing-data-of-650k-clients-for-two-years

***

Thousands of banking records exposed after Alberta dental benefits administrator hacked

By Adam Toy

A recent cybersecurity breach targeting a government service provider could have left thousands of Albertans' banking records and hundreds of thousands more names exposed.

On Thursday, the Alberta Dental Service Corporation (ADSC) said it had been hacked between May 7 and July 9. On the final day, the hackers communicated they had encrypted the data, were holding it ransom and demanded cryptocurrency in payment.

…

Global News 2023.08.10 07:53

https://globalnews.ca/news/9889724/alberta-dental-benefits-administrator-cybersecurity-breach

…

Attachment 3: Excerpt from the news collection for analysis

In order to construct a comprehensive network of concepts pertaining to data leakage in information systems as discussed in the news reports mentioned earlier, a predefined set of prompts was systematically employed for their analysis. These prompts are meticulously detailed in Attachment 4, and Attachment 5 provides a concrete illustration of a response generated in response to the specific queries outlined in the prompts. This structured approach facilitated the development of a robust conceptual network on the subject matter.

→ Print from the text 20 pairs of the most related entities. Each entity should be described in no more than 3 words. The output format is: 'entity 1; entity 2'. Each pair is displayed on a new line. Text: …

Attachment 4: Request to the ChatGPT system

Data Breach; university of sydney
Smartphone app; exposed
Bug fixes; security page
Data breaches; confidetial
Bank's website; safety measures
Big data companies; Swedish authority
Facebook; advertisements
encrypted data; ransom
Operating system; hackers
…

Attachment 5: Response to a request to the ChatGPT system

Based on the above-fixed prompt and the approach used, a swarm of virtual experts can be applied to increase the completeness of the network and filter biased semantic units to increase the accuracy of the network.

As a result of a similar procedure for all messages in the collection, following the described approaches, an array of message pairs was obtained, which was processed using CSV2Graph, and a network of connected pairs of different hierarchies was obtained that describes the relationships of semantic primitives from the news collection on the topic of data leakage in information systems, a fragment of which is shown in Fig. 5.

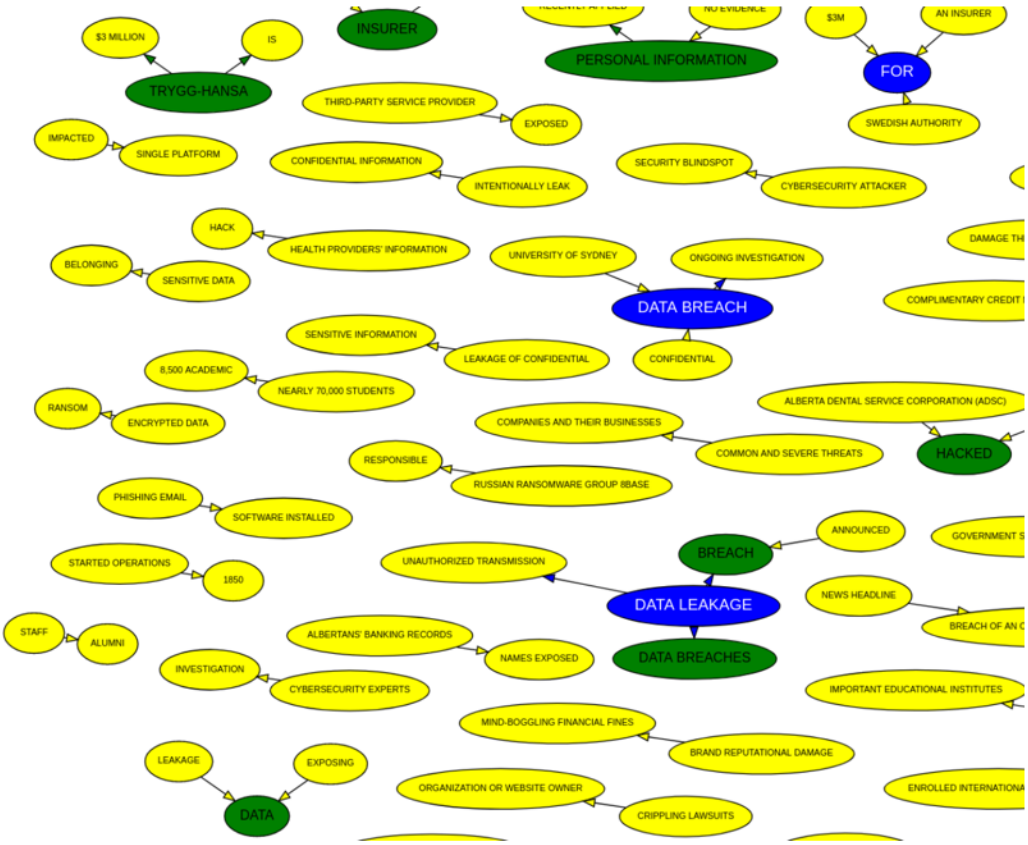

Figure 5. A fragment of the concept network corresponding to a news selection

The resulting visualization of the network using the space of own documents, where the main concepts are the most significant network nodes corresponding to the studied news selection, is a kind of information portrait of this array: DATA LEAKAGE, DATA BREACHES, FOR, HACKED, DATA, BREACH, INSURER, SOFTWARE, PERSONAL INFORMATION, MALWARE, TRYGG-HANSA, ERROR, PASSWORD, VIRUS.

The creation of domain-specific models using Language Model (LLM) technologies like GPT and BERT has emerged as a powerful approach for tailoring models to precise tasks and contexts. This approach grants the flexibility to adapt models according to specific needs. However, it's important to acknowledge the limitations imposed by text size and time constraints. LLM systems are restricted by the number of tokens they can process, which can pose challenges when dealing with lengthy or intricate texts. Nonetheless, there is a promising trend toward gradually increasing token limits, thereby broadening the horizons for analysts and developers. This progress enables the sequential assembly of responses from artificial intelligence systems without the need to segment the selected text spaces.

Looking ahead, it's plausible to consider the sustained utilization and automation of analytical procedures. This includes streamlining all manual processes involving interactions with artificial intelligence. Such efforts hold the potential for advancing the development and enhancement of the programmed components designed for various professions and specialties. These components can be instrumental in constructing networks of interconnected semantic concepts, fostering more efficient and insightful analysis across diverse domains. This continuous refinement and automation are pivotal in driving the evolution of AI-driven analytics.

Conclusions

The study underscores its exceptional proficiency in managing extensive datasets, streamlining the preparation of meta-data for various analytical purposes. It illuminates intricate semantic connections within the written content, equipping researchers with essential linguistic building blocks to advance their investigations. The integration of graphical

representation not only enhances semantic pattern analysis but also makes these insights more accessible to a broader audience. As artificial intelligence evolves, it promises to revolutionize research further, offering new avenues of exploration and deeper insights into the complexities of knowledge representation.

In more detail, based on expert opinions, it can be concluded that the primary causal network obtained by the simplest hierarchical access to ChatGPT covers the largest number of concepts that are relatively weakly connected, but can nevertheless serve as a good "raw material for further analytical processing". The second network, a casual network obtained by hierarchical access to ChatGPT by a swarm of virtual experts, shows a more complete range of concepts in the semantic network. The third network obtained by generalizing the hierarchical access of a swarm of virtual experts to ChatGPT, which has the highest average, represents the most accurate network excluding biased semantic units. This indicates that such a network is the most appropriate for further application of scenario analysis.

It should be noted that the approach to extracting entities from models based on generative systems is universal, but has its limitations. The main limitations include delays in updating models, the need to work with proprietary documentation, and restrictions on Internet access. Along with restrictions on the choice of products for visualization, this can make it difficult to conduct such analytical studies. To solve these problems can be recommended the process of fragmenting documentation into industrial products, combining model responses to create a common network of selected text spaces, or you can install your own LLM-type system on the corporate network and train it based on your own documentation for more accurate entity extraction. The developed CSV2Graph suite helps to solve the difficulties of limited access and difficulty in mastering.

Despite the significant gains in resource consumption (both time and human), it is important to note that both the process of building causal networks and the analysis of the results require the data scientist to have experience in the subject area under study, and human supervision is still necessary to ensure the reliability, accuracy of the results and their interpretation.

References

[1] St. Wolfram. "What Is ChatGPT Doing ... and Why Does it Work?". – Wolfram Media, Inc. March 9, 2023. 112 p.

[2] Brady D. Lund, Ting Wang, Nishith Reddy Mannuru, Bing Nie, Somipam Shimray, Ziang Wang. ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. JASIST, 2023. / Volume74, Issue5. Pages 570-581. DOI: https://doi.org/10.1002/asi.24750

[3] Tamilla Triantoro. Graph Viz: Exploring, Analyzing, and Visualizing Graphs and Networks with Gephi and ChatGPT (March 30, 2023). ODSC Community.

[4] Lande, Dmitry and Strashnoy, Leonard, Concept Networking Methods Based on ChatGPT & Gephi (April 17, 2023). SSRN. Available at http://dx.doi.org/10.2139/ssrn.4420452

[5] Ken Cherven. “Mastering Gephi Network Visualization”. – Packt Publishing, 2015. 378 p.

[6] Dmytro Lande, Leonard Strashnoy. GPT Semantic Networking: A Dream of the Semantic Web – The Time is Now. – Kyiv: Engineering, 2023. – 168 p. ISBN 978-966-2344-94-3