INSTITUTE FOR HOMELAND SECURITY

Sam Houston State University

# ARTIFICIAL INTELLIGENCE AND SOCIAL NETWORK ANALYSIS

## FOR CRITICAL INFRASTRUCTURE RESPONSE NETWORKS AND DARK NETWORK THREAT ANALYSIS


**Institute for Homeland Security**

**Sam Houston State University**

Nathan Jones

Christian Pamfile

Juli Dutta

Oscar Contreras Velasco

Michael Aspland

Table of Contents

# Executive Summary

This project has served as a proof of concept for the use of artificial intelligence such as the large language model (LLM) ChatGPT 4o to create datasets based on large quantities of qualitative data. LLMs have thus demonstrated the capacity to synthesize large quantities of qualitative data and turn them into quantifiable structured data sets that we can then use traditional software packages for social network analysis (SNA). We have demonstrated this using a case study related to the North Houston area and critical infrastructure response networks. We began by gathering a large number of qualitative documents related to congressional testimony, government reports, newspaper articles, among many other sources related to critical infrastructure protection response networks in the North Houston area. We then engaged in a process of prompt engineering to teach LLMs such as ChatGPT how to code this into structured data for SNA. We also took subsets of the data and had humans code them. We were then able to use the human coded data as a comparison point for the quality of the artificial intelligence generated data sets.

We generally found that ChatGPT 4o could only perform the task successfully on smaller batches of text. When larger batches were entered poor data was output in terms of the number of edges and actors in comparison to the baseline. On the other hand, small batch data generated by Chat GPT 4o produced more edges and nodes than the baseline human data. It should be noted that the small batch LLM generated data suffered from the unique identifier problem in which it identifies the same actor by multiple different names.

While the human coded networks generally generated more detailed data, with additional training the artificial intelligence large language models found an increasing number of nodes and edges. It also identified different actors and did not simply find relationships based on actor mention proximity in the text. Our results demonstrated dramatic improvement in artificial intelligence over the course of the last year. Earlier research demonstrated that earlier models such as ChatGPT 3.5 were not capable of finding this relational data, indeed hallucinating incorrect answers. This project has generally shown that large language models such as ChatGPT 4o now have the capability to synthesize relational data from large quantities of qualitative data.

This finding has implications not just for the study of critical infrastructure response (blue) networks but could also be applied in the future to critical infrastructure itself, dark

networks such as terrorists or criminal organizations which threaten US critical infrastructure.

We ran some preliminary tests on small samples or qualitative data related to dark network alliance and conflict data. Artificial intelligence was generally able to identify relationships and whether those relationships were positive or negative. Additional research is needed in this area, and this will no doubt prove a fruitful area of research that scholars are quickly jumping upon. This proof of concept also has practical implications far beyond looking at critical infrastructure response networks. The ability to glean relational data from large quantities of qualitative open-source data could allow for network analysis against myriad dark and blue/bright networks, on both open-source and classified data.

With this ongoing project we ultimately hope to be able to expand this system to be able to look at many major metro areas in the United States and identify where those response networks are densely connected, based on the assumption that well-networked response actors will respond better to any major threat. This would allow for the creation of a heat map of the United States across major metro areas which would allow us to assess which areas are less likely to respond well to a major disaster or threats to critical infrastructure and might thus need or be more likely to need U.S. military troop deployments. Thus, this could be one indicator among many for strategic intelligence and the defense of the homeland.

# Introduction

This project demonstrated that artificial intelligence (AI) large language models (LLMs) have now reached a point where they can be used to glean relational data from large quantities of qualitative materials related to critical infrastructure (CI) or emergency management (EM) response networks. While we generally found that human coded data generally contained more data points, with small batch data input, we showed AI LLMs have improved dramatically in the last year, demonstrating their capability of generating quantifiable data sets for social network analysis (SNA). In some cases, the LLM generated more nodes and edges, than the baseline human coded data with the caveat of some false positive nodes due to generic terms like local government and the same actor being identified under different agency names (unique identifier problems).

This proof of concept could apply not just to critical infrastructure response network data, but to any type of network with underlying unstructured qualitative relational data. This could include critical infrastructure itself, and dark networks such as criminal or terror networks. We also, in a very small test case, demonstrated that it could be used to effectively generate relational data for dark networks.

This paper is important proof of concept because earlier iterations of large language models such as ChatGPT 3.5 were unable to create data sets from qualitative data on dark networks for social network analysis. Indeed, earlier versions of artificial intelligence or large language models were more likely to "hallucinate" inaccurate data when given these complex tasks.[1] This proof-of-concept paper demonstrates that large language models such as ChatGPT 4o (Released May 2024) now have this capability and that it improves depending upon the quality of prompt engineering and amount of text analyzed. The amount of text analyzed dramatically changed results; with smaller batches of data resulting in more actors and relationships identified. Thus, the *process* proved to be important.

## NORTHCOM Mission Set Relevance

The ability to use AI and LLMs to create data sets for SNA of CI response networks, speaks directly to NORAD/NORTHCOM mission sets such as defending the homeland. It has the long-term potential to be a tool for identifying vulnerabilities to US critical infrastructure

[1] Nathan Jones, "A Methodological Note: ChatGPT Is Not Ready for Criminal Network Analysis from Unstructured Data," *OODA Loop* (blog), May 8, 2023, https://www.oodaloop.com/archive/2023/05/08/a-methodological-note-chatgpt-is-not-ready-for-criminal-network-analysis-from-unstructured-data/.

and areas where critical infrastructure response networks may be weak necessitating DOD support.

NORTHCOM strategic principles include: "Global integration in order to achieve a globally integrated layered defense."[2] This project speaks to the homeland layer of defense: "The homeland layer consists of joint force capabilities integrated with the whole-of-government/interagency and strategic private sector partner capabilities."[3] This project speaks to cooperation and integration of "whole-of-government/interagency and strategic private sector partner capabilities"[4] and a metric (how well networked a metro area is) for its ability to respond to critical infrastructure threats from terror attacks, natural disasters, and foreign adversary attacks.

This research directly relates to enduring condition #1 as identified by NORTHCOM:

> Enduring Condition #1-Enhance National resiliency. Equally as important as defeating threats is the hardening of critical infrastructure and promoting domestic resilience in order to mitigate the consequences of attacks, both kinetic and non-kinetic. Our demonstrated ability to respond to diverse attacks with a whole-of-government response is a strong deterrent to our adversaries. Protecting our nations is a prerequisite to projecting power abroad.[5]

Whole-of-government response requires integrating with local, state, federal and private actors to respond to disasters and attacks. Rapid intelligence on the structure of those networks allows the most effective integration points such as using the principle of preferential attachment wherein new entrants get the most value from connecting to well-connected actors.[6]

## Roadmap

This paper will proceed in the following sections: (1) it will review the existing literature on the use of artificial intelligence or large language models for the creation of datasets in

---

[2] "Strategy," US NORTHCOM, accessed May 18, 2024, https://www.northcom.mil/Strategy/; "Department of Defense Instruction (3025.21)" (Department of Defense, February 27, 2013), https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/302521p.pdf.
[3] "Strategy."
[4] "Strategy."
[5] "Strategy."
[6] Albert-László Barabási and Eric Bonabeau, "Scale-Free Networks," *Scientific American*, 2003, https://www.scientificamerican.com/article/scale-free-networks/; Oscar Contreras Velasco et al., "The Use of Similarity-Based Algorithms to Predict Links in Mexican Criminal Networks," Research Paper (Houston: Rice University's Baker Institute, August 30, 2023), https://doi.org/10.25613/BQ88-X176; Nathan P Jones, Russell Lundberg, and Matthew O'Deane, "A Mixed Methods Social Network Analysis of San Diego Law Enforcement Task Forces and Agencies," *International Journal of Police Science* 1, no. 2 (2022): 58–83, https://doi.org/10.56331/487529/IJPS6.

social network analysis, (2) it will describe our methodology wherein we compared subsets of human coded data to the large language model generated data from two different LLMs; (3) it will present our case studies, focusing on the North Houston area critical infrastructure response networks and briefly touching upon a dark network subset of data; (4) it will then analyze those data sets presenting visualizations of the data and the results of various metrics to compare the human coded data two datasets created by Chat GPT and Google's open-source FLANT5; and (5) it will then provide conclusions including lessons learned, limitations, and avenues for future research.

# Literature review

We are not the first to consider the possibility of utilizing artificial intelligence on unstructured data for the purposes of social network analysis and the creation of datasets. Indeed, shortly after we submitted the proposal for this project David Bright a well-known criminal network analyst out of Australia advertised dissertation level projects and funding to use artificial intelligence to create criminal network analysis data sets. Scholars such as Salcedo have also pointed to LLMs and machine learning for SNA data, and their contributions will be discussed below.[7] To the best of our knowledge, we are the first to apply artificial intelligence (LLMs) to unstructured data for the purposes of creating systematic social network analysis critical infrastructure protection (CIP) data sets.

This literature review has two key aspects: (1) a discussion of social network analysis on dark and critical infrastructure networks; and (2) the application of artificial intelligence (AI) and large language models (LLMs) and machine learning to network analysis.

## Social Network Analysis

Social network analysis (SNA) takes as its starting point relational data.[8] Unlike classical statistics where all data is about individual observations and their characteristics, social network analysis is interested in the relations (edges) between actors/nodes/agents which can be people, organizations, private firms, NGOs, etc.[9] SNA has boomed in the last 25

[7] Eduardo Salcedo-Albarán, Luis Garay, and José Cano Melani, "Machine Learning Models on Criminal Networks (MLMoCN): Artificial Intelligence to Disentangle Crime," 2023.
[8] Stanley Wasserman and Katherine Faust, *Social Network Analysis: Methods and Applications*, Structural Analysis in the Social Sciences (New York: Cambridge University Press, 1994).
[9] Linton C. Freeman, *The Development of Social Network Analysis: A Study in the Sociology of Science* (Vancouver: Empirical, 2004), https://www.researchgate.net/profile/Linton-Freeman-2/publication/239228599_The_Development_of_Social_Network_Analysis/links/54415c650cf2e6f0c0f616a8/The-Development-of-Social-Network-Analysis.pdf.

years with the advent of the personal computer making SNA software easily available.[10] Network analysis broadly, has been applied to critical infrastructure before by classic scholars such as Ted Lewis who looked at critical infrastructure itself through network lenses and various paradigms from physics.[11]

Albert-László Barabási famously applied scale-free network concepts to critical infrastructure networks such as airports.[12] In scale-free networks some nodes like hub airports such as Houston International have far more connections than other nodes. In this sense, the US airport network is scale-free in that it has critical hubs and is nonrandom in nature. In a random network all nodes would have roughly the same number of connections. Random networks are susceptible to terror attacks and random failures, which can result in severe disruption. On the other hand, scale-free networks are resilient to random failures and one-off terror attacks because they can easily reroute traffic. Scale free networks are, however, susceptible to intelligent adversary attacks such as multiple simultaneous terror attacks on hubs.[13]

Barabási has also shown that preferential attachment is important to understanding the scale-free nature of networks. New network actors are more likely to attach to well connected actors because it gives the new entrant more access to the resources of the network.[14] Thus, understanding the underlying network structure of a CI response network can be critical to understanding how to best integrate and support that network, which is a key piece of strategic intelligence that could prove vital to DOD support of civilian CI response networks.

Sean Everton of the CORE research lab at the Naval Postgraduate School has championed dark network analysis for the study of terror, criminal, and any network that is in opposition to the state apparatus wherever they may be.[15] The analysis of relational data allows us to glean insights into social structures which in turn can explain the behavior of actors such as terror groups.[16] Beyond predicting behaviors, social network analysis can be used for targeting dark network actors within networks. Given scarce counterinsurgent/law

---

[10] Sean F Everton, *Disrupting Dark Networks*, Structural Analysis in the Social Sciences 34 (New York, NY: Cambridge University Press, 2012).
[11] Ted G Lewis, *Critical Infrastructure Protection in Homeland Security: Defending a Networked Nation* (John Wiley & Sons, 2006).
[12] Barabási and Bonabeau, "Scale-Free Networks."
[13] Albert-László Barabási, "Scale-Free Networks: A Decade and Beyond," *Science*, July 24, 2009, https://barabasi.com/f/303.pdf; Barabási and Bonabeau, "Scale-Free Networks."
[14] Barabási, "Scale-Free Networks: A Decade and Beyond"; Jones, Lundberg, and O'Deane, "A Mixed Methods Social Network Analysis of San Diego Law Enforcement Task Forces and Agencies."
[15] Everton, *Disrupting Dark Networks*.
[16] Valdis E. Krebs, "Mapping Networks of Terrorist Cells," *Connections* 24, no. 3 (2002): 43–52.

enforcement resources, SNA can paint the way to targeting the most important actors. For example, state forces could fragment a dark network by targeting the most central actors according to various centrality metrics include degree (count of ties), betweenness (brokerage), and eigenvector (actors connected to highly connected actors).[17] Conversely, network analysis can point us toward vulnerable/critical assets to protect in critical infrastructure.[18]

There are myriad examples of dark network analysis sometimes referred to as criminal network analysis.[19] These include money laundering networks, cross-border drug trafficking networks, hydrocarbon theft networks, terrorist networks, among myriad others.[20] Recent scholars such as Contreras Velasco et al (2023) have applied machine learning and predictive algorithms to predict new dark network alliance formation.[21]

SNA can also be applied to "blue" networks that assist the state and promote legal activities.[22] This is particularly relevant here because concepts such as density of connection can help inform us about whether or not critical infrastructure response networks will respond well given their pre-existing relationships.[23] Social network analysis

---

[17] Linton C Freeman, "Centrality in Social Networks Conceptual Clarification," *Social Networks* 1, no. 3 (1978): 215–39; Daniel Cunningham, Sean Everton, and Philip Murphy, *Understanding Dark Networks: A Strategic Framework for the Use of Social Network Analysis* (London: Rowman & Littlefield, 2016).
[18] Lewis, *Critical Infrastructure Protection in Homeland Security: Defending a Networked Nation*.
[19] Luis Jorge Garay Salamanca, Eduardo Salcedo-Albarán, and Isaac de León Beltrán, *Illicit Networks, Reconfiguring States: Social Network Analysis of Colombian and Mexican Cases* (Bogotá: Metodo Foundation, 2010).
[20] Eduardo Salcedo-Albaran and Luis Jorge Garay Salamanca, "Structure of a Transnational Criminal Network: 'Los Zetas' and the Smuggling of Hydrocarbons," Working Paper (Vortex, 2014), http://www.scivortex.org/12TCNsMexUsV2.pdf; Carlo Morselli, Cynthia Giguère, and Katia Petit, "The Efficiency/Security Trade-off in Criminal Networks," *Social Networks* 29, no. 1 (January 1, 2007): 143–53, https://doi.org/10.1016/j.socnet.2006.05.001; David A Bright et al., "The Use of Actor-Level Attributes and Centrality Measures to Identify Key Actors: A Case Study of an Australian Drug Trafficking Network," *Journal of Contemporary Criminal Justice* 31, no. 3 (2015): 262–78; Jun Wu et al., "A Social Network Analysis of an MS-13 Network: Structure, Leadership Roles, and the Use of Confidential Informants," *International Criminology*, March 9, 2024, https://doi.org/10.1007/s43576-024-00113-9; Nathan P. Jones et al., "A Mixed Methods Social Network Analysis of a Cross-Border Drug Network: The Fernando Sanchez Organization (FSO)," *Trends in Organized Crime* 23, no. 2 (June 1, 2020): 154–82, https://doi.org/10.1007/s12117-018-9352-9.
[21] Contreras Velasco et al., "The Use of Similarity-Based Algorithms to Predict Links in Mexican Criminal Networks."
[22] John Arquilla and David F. Ronfeldt, eds., *Networks and Netwars: The Future of Terror, Crime, and Militancy* (Santa Monica, CA: Rand Corporation, 2001).
[23] Steve Ressler, "Social Network Analysis as an Approach to Combat Terrorism: Past, Present, and Future Research," *Homeland Security Affairs* 2, no. 2 (2006): 1–10; Arquilla and Ronfeldt, *Networks and Netwars: The Future of Terror, Crime, and Militancy*.

takes as its fundamental assumption that the structure of a network is important to its overall functioning.[24]

## Large Language Models and Network Analysis

Artificial Intelligence (AI) has recently surged in popularity heavily led by the private sector adoption of various types of AI.[25] Large language models (LLMs) have been trained on vast amounts of natural language text (NLT). ChatGPT from Open AI is one of the earliest market entrants into public facing large language models (LLMs), what is commonly referred to as artificial intelligence.[26] Google's FLANT5 from Huggingface.com is another Open Access large language model that has been trained on large quantities of natural language text (NLT). It can be downloaded to a single computer and does not rely on outside servers to analyze the data. Thus, it requires a significant graphics processing unit or GPU to run.

Scholars and technologists have identified large language models as having what is known as the black box problem.[27] Because the artificial intelligence or large language model is not thinking when it generates an answer it is not trained to tell us how it came to its conclusions. In this sense what it generates is effectively a black box and users receive answers without a reasoning for the responses. This will be discussed further in application of this project in the analytical section.

# Methods

The researchers began by gathering critical infrastructure response network relevant documents throughout the semester. We focused on the North Houston area and gathered documents related to key events demonstrating critical infrastructure response network relationships of cooperation, collaboration, or in some cases conflict. These included the response to Hurricane Harvey of 2017,[28] the Winter storm Uri of 2021,[29] among others.

---

[24] Wasserman and Faust, *Social Network Analysis: Methods and Applications*; Cunningham, Everton, and Murphy, *Understanding Dark Networks: A Strategic Framework for the Use of Social Network Analysis*; Everton, *Disrupting Dark Networks*.
[25] Thomas H. Davenport and Nitin Mittal, *All in on AI: How Smart Companies Win Big with Artificial Intelligence* (Boston: Harvard Business Review, 2023), 12.
[26] Salcedo-Albarán, Garay, and Cano Melani, "Machine Learning Models on Criminal Networks (MLMoCN): Artificial Intelligence to Disentangle Crime."
[27] Lou Blouin, "AI's Mysterious 'Black Box' Problem, Explained," University of Michigan-Dearborn, March 6, 2023, https://umdearborn.edu/news/ais-mysterious-black-box-problem-explained.
[28] Michael Kimmelman, "Lessons From Hurricane Harvey: Houston's Struggle Is America's Tale," *The New York Times*, November 11, 2017, sec. Climate, https://www.nytimes.com/interactive/2017/11/11/climate/houston-flooding-climate.html.
[29] "Texas Severe Winter Storm DR-4586," accessed May 14, 2024, https://www.tdem.texas.gov/disasters/winter-storm-uri.

Documents collected included government reports, after-action reports, congressional testimony, NGO reports, newspaper articles discussing response relationships including when they were conflictual, etc. Once we had accumulated dozens of documents across thousands of pages, we identified a subset of data of approximately 100 pages for a human to code into an edge list which could be analyzed via a social network analysis software known as Gephi.[30]

Networks were coded separating people to people networks, people to organization networks, and organization to organization networks. We focus the analysis on the org-to-org networks here.

Where possible we prompt engineered to create additional columns of data beyond the edge list so that we could have and visualize data based on relationship type. In some cases, we also coded nodelists and included attributes for the data.

The task here involves identifying relationships between entities. While doing that task, the LLM must engage in name entity recognition (NER) from natural language text (NLT).[31] This has been identified as an important task within the literature on LLMs.[32] We found ChatGPT4o generally did this effectively without significant alterations and thus we focused our efforts on generating edge lists. A discussion of the problem of unique identifiers is included in the analysis.

Our methods were exploratory and adaptive. We tested results on various sizes of documents finding that splitting documents into smaller text chunks generated more edges and therefore better data.

Separate chats gave different results. Continuing with a single chat tended to be more difficult and result in worse results. Thus, for reasons of replicability and ChatGPT capacity we created new chats for each upload and relational analysis from ChatGPT.

## FLANT-5

Flan-T5 is a variant of the T5 (Text-to-Text Transfer Transformer) model developed by Google, designed for a wide range of natural language processing tasks. T5 models are pre-trained on a diverse corpus of text in a multi-task learning framework, converting all tasks to a text-to-text format, which simplifies the architecture and allows for flexible task adaptation. Flan-T5 extends this approach by incorporating instruction-based fine-tuning, enhancing

[30] Mathieu Bastian, Sebastien Heymann, and Mathieu Jacomy, "Gephi: An Open Source Software for Exploring and Manipulating Networks," vol. 3, 2009, 361–62.
[31] D. Carpintero, "Named Entity Recognition to Enrich Text," OpenAI Cookbook, October 19, 2023, https://cookbook.openai.com/examples/named_entity_recognition_to_enrich_text.
[32] Carpintero.

its ability to understand and follow complex instructions. This fine-tuning involves training the model on various tasks with detailed instructions, significantly improving its performance in generating coherent and contextually appropriate responses. The Flan-T5 model is particularly adept at tasks such as summarization, translation, question-answering, and text generation, making it a powerful tool for applications requiring advanced language comprehension and generation.

The provided code demonstrates a method to process and summarize a lengthy text document using the Flan-T5 model from the Hugging Face Transformers library. It starts by defining a function, `split_text_into_segments`, which divides the text into smaller segments based on a specified maximum length to manage long documents that exceed the model's input size limitations. The text is read from a file, split into segments, and then processed using the Flan-T5 model and tokenizer. For each segment, the text is prefixed with "summarize:", tokenized, and passed to the model for summarization. The generated output is then decoded and printed. This approach ensures that large documents can be effectively summarized in manageable parts, leveraging the capabilities of the Flan-T5 model for natural language processing tasks.

# Case Study Presentation

## The North Houston Area Critical Infrastructure Response Network

The North Houston Area is prone to major natural disasters including, but not limited to hurricanes, storms, floods, winter freezes impacting critical infrastructure. Indeed, during this project the Houston area was hit by a major wind event, cutting power to hundreds of thousands and directly impacting the researchers on this project.[33] While Houston has significant vulnerabilities, it is known for effective public and private cooperation in emergency response to these CI events. This makes it an effective pilot case for CI Response Network collaboration which can serve as a model that can be applied to other regions.

## Subset of Data for Human Coding Comparison (Baseline)

Because of the difficulty of human coding of data, we chose a subset of data for human coding and then compared it against ChatGPT 4o and FLANT-5 results. The subset of data was chosen in part for the richness of its relational data, in it included after action reports,

[33] "Storm Carves Path of Destruction Across Houston," *The New York Times*, May 17, 2024, sec. Weather, https://www.nytimes.com/2024/05/17/weather/houston-storm-photos-video.html; "Texas Severe Winter Storm DR-4586."

and congressional testimony related to Hurricane Harvey response. This subset of data was also an effective first cut at training ChatGPT and FLANT-5.

The visualization below Figure 1 shows the results of the human coding of data based on the organization-to-organization network. For ease of visualization, it only shows the main component. Nodes were sized based on betweenness centrality which is widely considered a brokerage metric. Nodes and edges in this network are colored by a community detection algorithm known as the Louvain method.[34] This method detects subgroups or communities within the overall network mathematically. For example, it is no surprise that the Coast Guard is represented in orange. In this same community, we also see Marine Corps Amphibious Unit 4, the Marine Corps, the Department of Defense, Special Forces, National Guard, and the Texas Department of Public Safety. It is no surprise that the largest node in the network based on betweenness centrality is FEMA followed by Housing and Urban Development (HUD). A separate dark green community comprised of the Army Corps of Engineers and Harris County Flood Control District among others related to water infrastructure was also detected as its own community, suggesting a coherent logic to the underlying data.

[34] Vincent D. Blondel et al., "Fast Unfolding of Communities in Large Networks," *Journal of Statistical Mechanics: Theory and Experiment* 2008, no. 10 (October 9, 2008): P10008, https://doi.org/10.1088/1742-5468/2008/10/P10008.

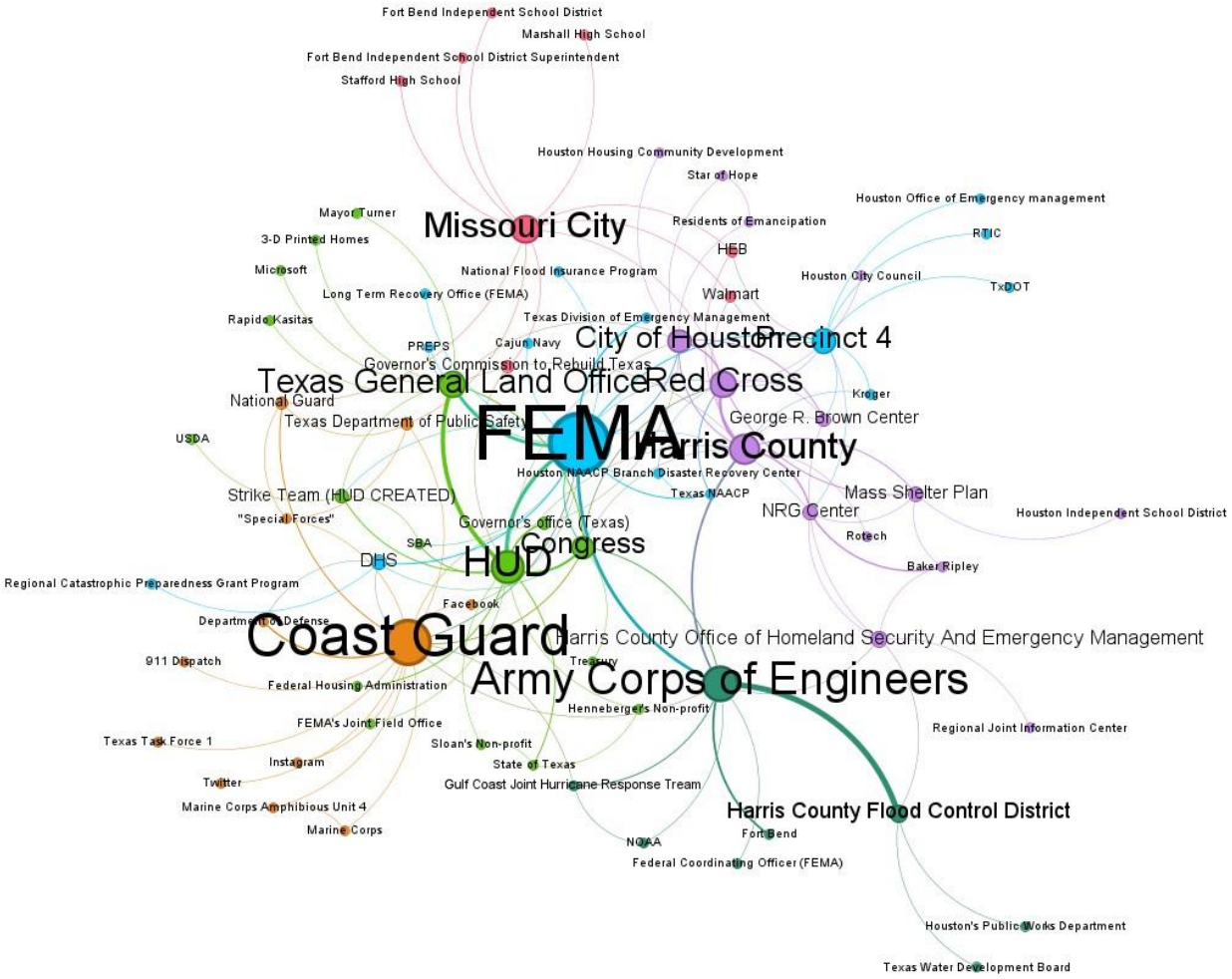



Figure 1. Human Coded Subset Main Component: nodes sized on betweenness centrality, nodes and edges colored by community detection, and Yifan Hu layout.

With our human coded baseline data, we moved on to creating a ChatGPT 4o data set using the prompt in Appendix 1. This followed a period of significant prompt experimentation by the researchers.

## Chat GPT 4o Large Batch Subset (Failure)

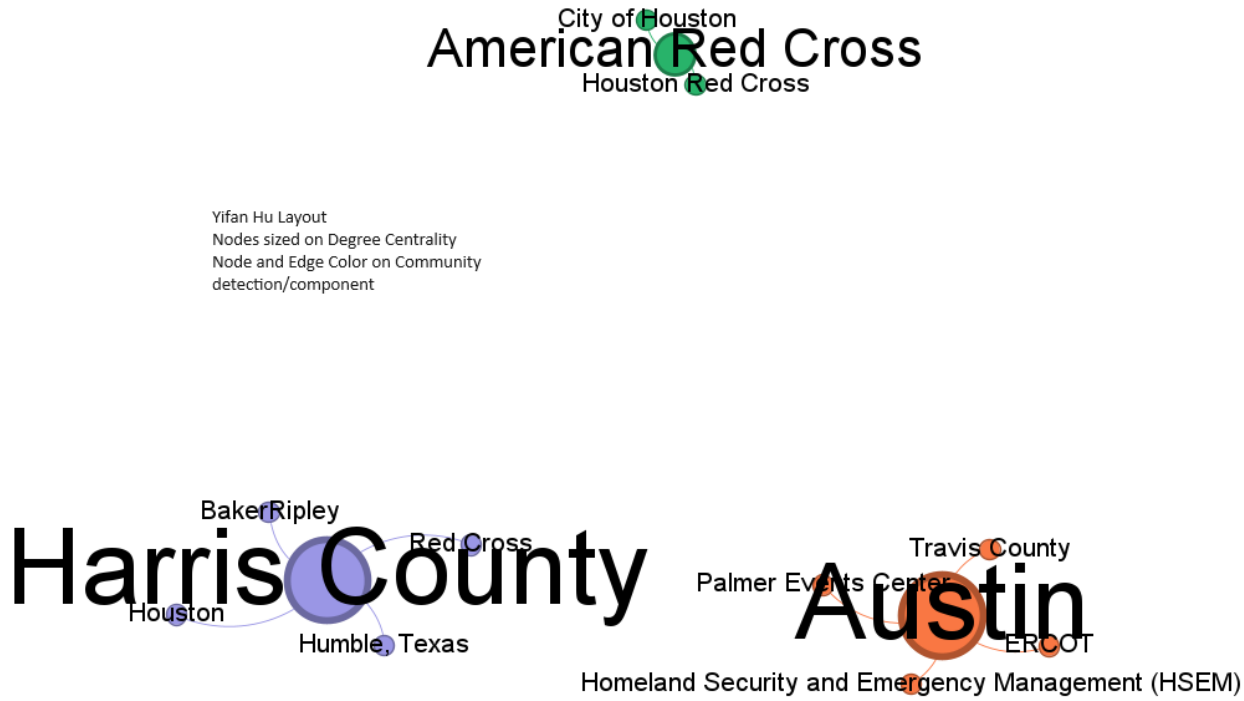



Figure 2. Large Batch Subset (Failure)

Part of the promise of large language models is the ability to upload large quantities of data and generate this relational data. We found at this time ChatGPT produced substandard results when asked to generate relational datasets from multiple large documents at a time. ChatGPT large batch document uploads simply did not perform well in terms of the number of edges and nodes generated when we uploaded all four documents at once. In the next section we will describe a successful test case in which we broke the data into smaller inputs.

## Chat GPT4o Small Batch Subset for Comparison (Success)

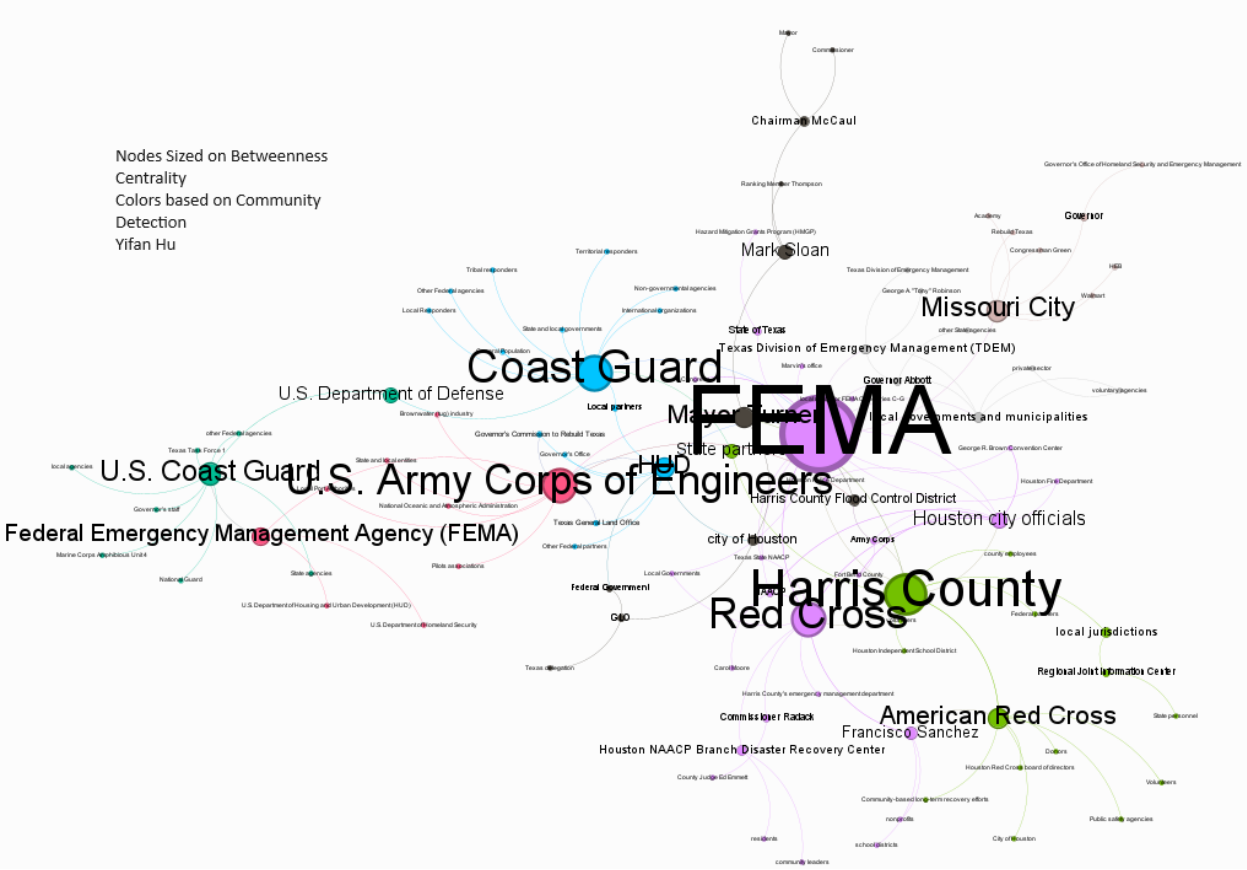



Figure 3. ChatGPT 4o Small Batch Subset (Success)

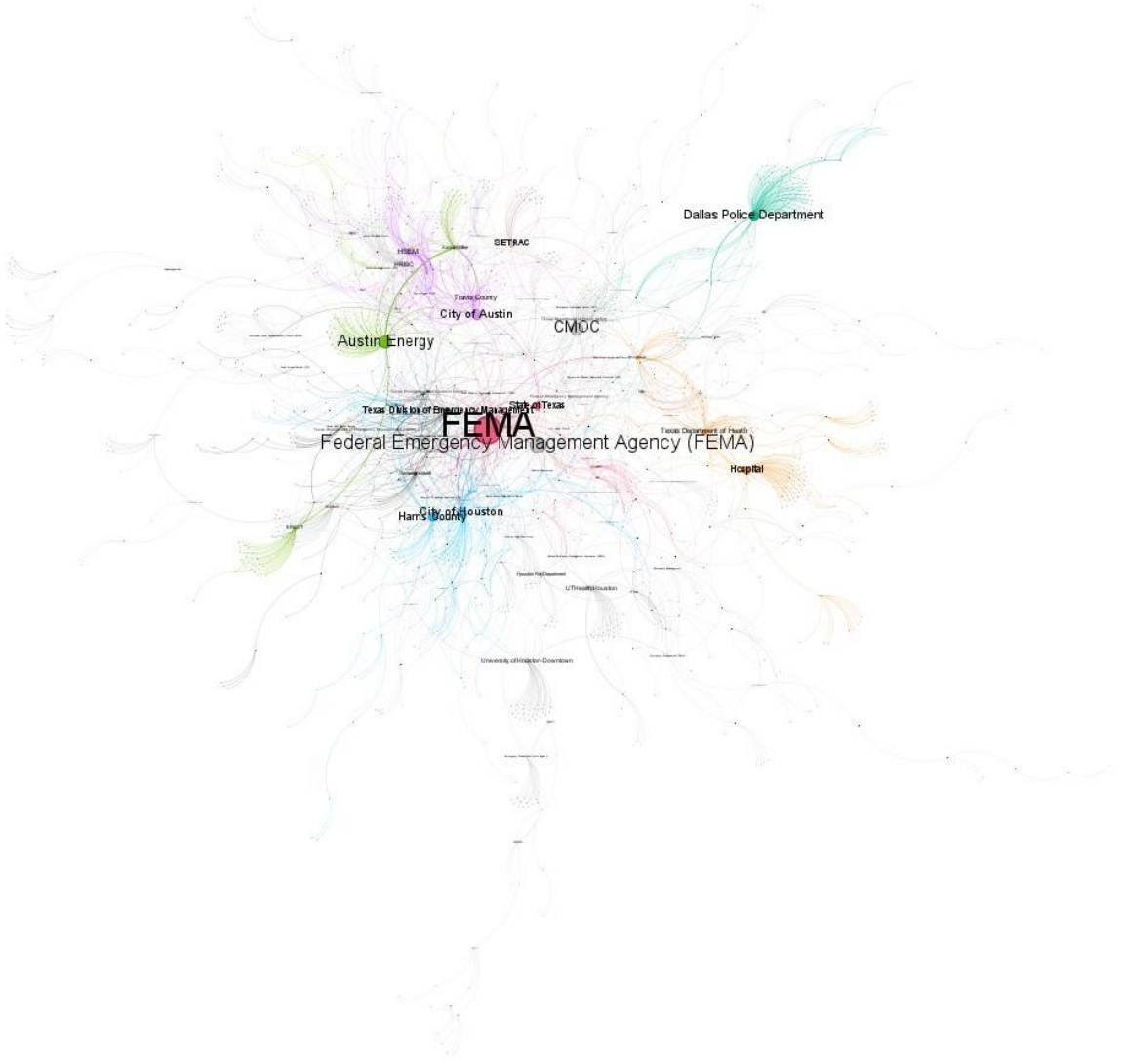
We consider the ChatGPT small batch subset data a qualified success. We base the success determination on the number of nodes and edges generated which rivaled and surpassed the human coded data. There are some caveats, however. These include broadly defined categories such as local agencies or local governments which could lead to false positive edges amongst different entities. Further, in some places, entities were identified under different names despite being the same entity, e.g. Coast Guard vs. US Coast Guard and FEMA vs. Federal Emergency Management Agency. The unique identifier problem, which will be discussed further in the analytical section. Nonetheless, the data generated was rich and specific metrics were easily derived. They will be provided in the analytical section (Table 1).

## Full Data Set Output (Full 44 Document Set added in small batch)

The full data set output with 44 documents was a qualified success insofar as it generated a large number of actors and edges. We emphasize that this is a qualified success because it also suffered from the unique identifier problem. Thus, there are a significant number of

false positive nodes. For example, Chat GPT4o identified FEMA, Federal Emergency Management Agency (FEMA), and also identified Federal Emergency Management Agency as separate actors. This goes back to the unique identifier problem and could be rectified in the future by creating a dictionary of actors and aliases.[35] Nonetheless, we did not want to run analysis on only previously identified actors because part of the value of the LLM is its ability to identify new actors in new data. This is a fertile area for future research.

Below is visualization of the largest data set in the study. Only the main largest component is shown here due to size.




[35] Jesús Espinal-Enríquez et al., "A Literature-Based Approach to a Narco-Network" (Social Informatics: SocInfo 2014 International Workshops, Barcelona, Spain, November 11, 2014, Revised Selected Papers 6, Springer, 2015), 97–101.

Figure 4. Full Data Set Output Chat GPT 4o (44 documents small batch upload)

## FLANT-5 Subset Data (Failure)

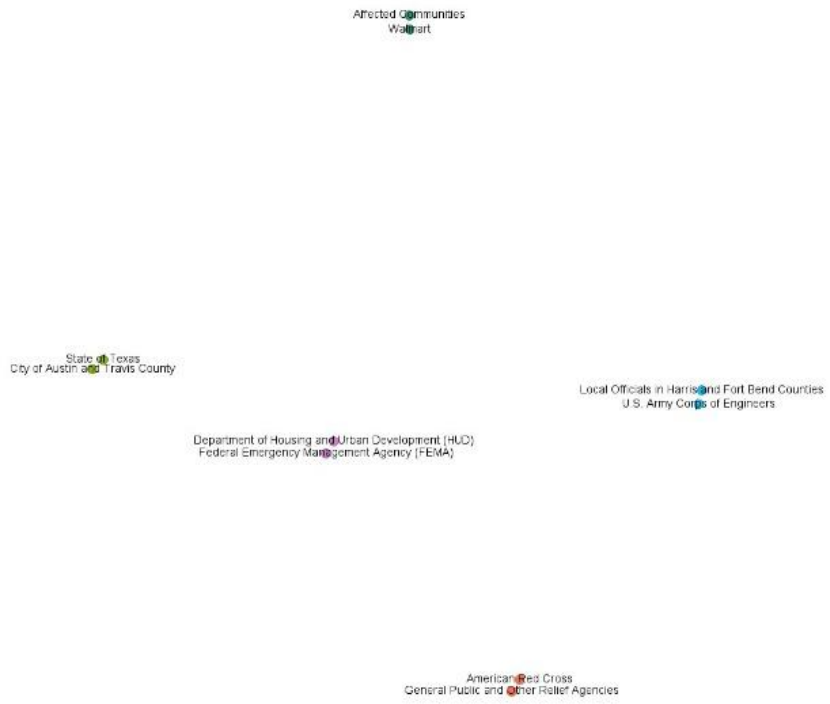



Figure 5. FLANT5 Data Output

As we can see from the visualization above, the FLANT5 data set only generated 10 nodes and five edges. The low number of five edges is why we deem this attempt at using FLANT5 for relational data a failure. 5 edges are simply too small a number of relationships to be useful for network analysis data. Because we know that the baseline number of edges based on the human coded network was 80 agencies, we know that this version of the FLANT5 data is not useful network data. We acknowledge that this could have been the result of insufficient fine tuning of data and the fact that FLANT5 is from an earlier generation of LLMs when compared to the recently released ChatGPT 4o (MAY 2024 release).

## Mini-Case Comparison of Dark Network (Mexican Cartel Conflict/Alliance)

In the course of this project, we also created a dark network mini case study utilizing Mexican cartel conflict and alliance networks. We utilized an article by Oscar Contreras Velasco which is a peer reviewed article discussing cartel alliances and conflicts.[36] Thus, this was an easy case to rapidly test the capability of artificial intelligence large language models to create dark networks from unstructured data. The visualization below demonstrates that it was able to rapidly code the data based a 30-page article which itself discussed alliances.

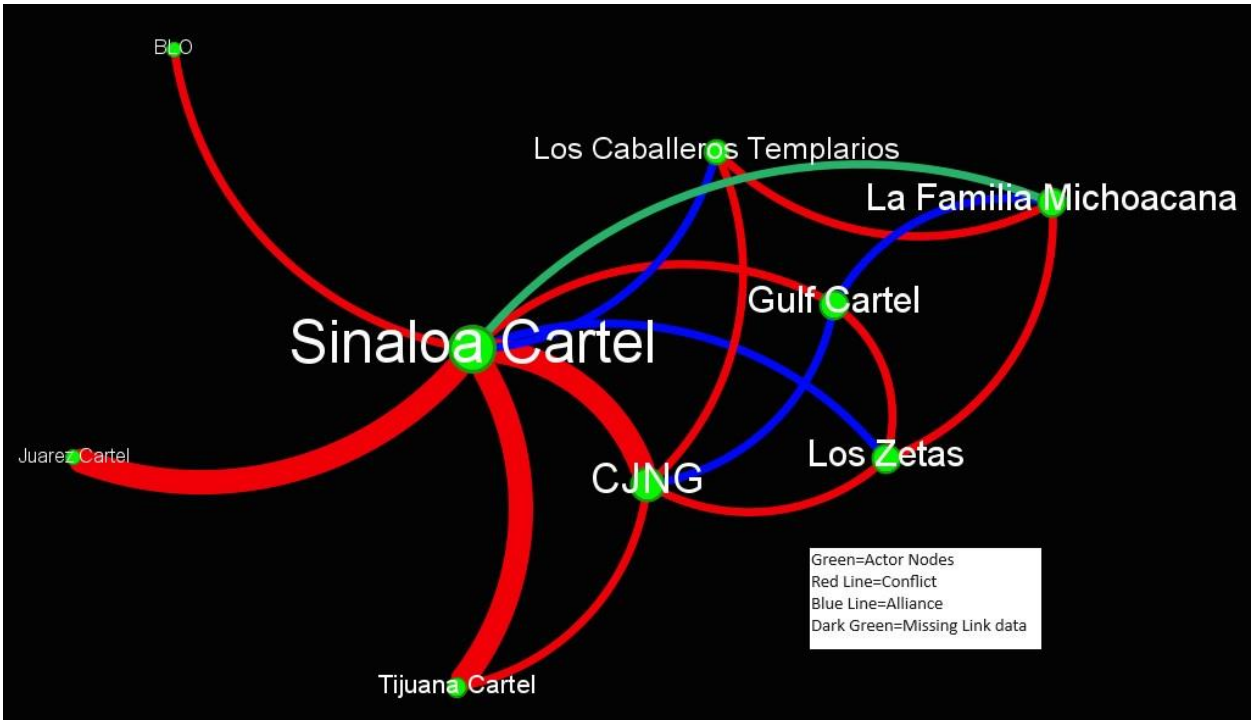



Figure 6. Mini Case Study -Mexico Cartel Conflict/Alliances

The red lines indicate conflictive relationships or wars while the blue indicate cartel alliances. Here we can see that this cursory data set identifies the Sinaloa Cartel and the Cártel de Jalisco Nueva Generación (CJNG) as the two most central in the conflict and alliance networks here. This is consistent with recent research,[37] all expert opinion, and

[36] Oscar Contreras Velasco, "Unintended Consequences of State Action: How the Kingpin Strategy Transformed the Structure of Violence in Mexico's Organized Crime," *Trends in Organized Crime*, July 10, 2023, https://doi.org/10.1007/s12117-023-09498-x.

[37] Nathan Jones et al., "A Social Network Analysis of Mexico's Dark Network Alliance Structure," *Journal of Strategic Security* 15, no. 4 (2022), https://doi.org/10.5038/1944-0472.15.4.2046.

the recently released DEA reports on transnational criminal organizations (TCOs) in Mexico.[38]

Dark network analysis from unstructured data could also have significant implications for critical infrastructure protection and homeland defense given these malevolent actors pose threats to US critical infrastructure. These threats could be from direct targeting, or the unintended consequences (second and third order effects) of their profit-seeking illicit activities.

# Analysis

Table 1 below provides network topography metrics which describe the general structure of each of the various networks analyzed here. These networks included the human coded data, ChatGPT 4o the small batch data (4 doc subset), ChatGPT 4o large batch subset of data (4 doc Subset), and the ChatGPT maximum document set of 44 documents run in small batches. We define small batch data as uploading approximately 15 pages of text. Through trial and error, we identified that more than 15 pages of text resulted in failure in the sense that few nodes and edges were identified, despite more being present. Large batch refers to analyzing entire and multiple documents in a single prompt. We found that the results of these analysis were effectively failures in terms of the number of nodes and edges generated which were unrealistically small. We treated the human coded data set as a baseline (80 actors and 136 edges) from which we could compare the other LLM results. We found that by running small batches and limiting the amount of text fed into ChatGPT at a given time, we attained far more detailed and rich results.

| Network Topography Metric (Undirected) | Human Coded Data Set | Chat-GPT 4o (Run in small batches) | Chat GPT 4o Large Batch 4 Documents | ChatGPT 4o Small Batch (Full data set 44 Docs) | FlanT-5 |
|---|---|---|---|---|---|
| Nodes | 80 | 141 | 13 | 2152 | 10 |
| Edges | 136 | 159 | 10 | 2738 | 5 |
| Average Degree | 3.4 | 2.27 | 1.538 | 2.545 | 1 |
| Average Weighted Degree | 5.375 | 2.34 | 1.538 | 2.905 | 1 |
| Graph Density | 0.043 | 0.016 | 0.128 | 0.001 | 0.11 |
| Network Diameter | 6 | 9 | 2 | 16 | 1 |
| Average Path Length | 3.029 | 4.083 | 1.56 | 5.012 | 1 |
| Avg. Clustering Coefficient | 0.503 | 0.278 | 0 | 0.26 | 0 |

[38] "National Drug Threat Assessment 2024" (Drug Enforcement Administration, 2024), https://www.dea.gov/sites/default/files/2024-05/NDTA_2024.pdf.

| | | | | | |
|---|---|---|---|---|---|
| Total Triangles | 60 | 23 | 0 | 634 | 0 |
| Connected Components | 3 | 14 | 3 | 136 | 5 |
| Girvan Newman Communities | 7 | 23 | 3 | 137 | 5 |
| Modularity | 0.482 | 0.732 | 0.64 | 0.815 | 0.8 |
| Modularity Communities | 8 | 21 | 3 | 159 | 5 |

Table 1. Network Topography Metrics[39]

As we can see from above the ChatGPT small batch data entry performed well in terms of generating the largest number of nodes and edges. We made that assessment with the caveat that the larger number of nodes could in part be the result of false positives due to the unique identifier problem (discussed further below). Nonetheless, we can see the ChatGPT generally performed better when smaller amounts of text were entered. Edgelists were then combined in spreadsheets. This led to a tedious data entry process which for a time stopped as we hit data limits. We were forced to switch between ChatGPT usernames within our team's account to continue our data entry and ultimately had to stop at 44 documents instead of our full roughly 60 documents.

## Unique Identification Problems

In numerous places ChatGPT generated data with problems of unique identification. These problems stem from non-unique identifiers for the same entity. For example, in one place ChatGPT coded “Coast Guard”, but in another coded “U.S. Coast guard”. These are clearly the same overarching entity coded twice under different identifiers. This will result in an inflated number of nodes, but also a reduced degree for those nodes which are improperly identified multiple times. This can also be an issue with human coding, but in general we had more faith that human coding would do this less. We experimented with prompt engineering that would reduce the number of non-unique identifiers or agencies referred to by generic terms like “local agency.” But we also found that the quality of the edges went down when we did this. We base this on a human qualitative assessment of the data based on our previous experience with human coding and our reviews of the documents. This

[39] Michelle Girvan and Mark EJ Newman, “Community Structure in Social and Biological Networks,” *Proceedings of the National Academy of Sciences* 99, no. 12 (2002): 7821–26, https://doi.org/10.1073/pnas.122653799; Matthieu Latapy, “Main-Memory Triangle Computations for Very Large (Sparse (Power-Law)) Graphs,” *Theoretical Computer Science* 407, no. 1–3 (2008): 458–73; Blondel et al., “Fast Unfolding of Communities in Large Networks.”

result stems from the failure to create a dictionary or appendix of all named entities and their aliases.[40]

## Black Box Test

Part of this project's goal was to overcome the black box problem by comparing artificial intelligence large language model generated networks to human coded networks to see where the key differences might lie. We did this in a fairly simple way. First, we human-coded the data on a small subset of data on congressional testimony from Houston Mayor Turner on Hurricane Harvey response. This smaller amount of data allowed us to do a textual analysis by hand and it was a rich subset of data. Second, we took this same congressional testimony and had ChatGPT4o code it. Third, we then coded it by hand but not using human understanding, but typical text script analytic methods that a simple algorithm would use.[41] For example, in one we coded any two entities mentioned in the same paragraph as connected. By comparing across these networks, we were able to see whether ChatGPT and the LLM was simply coding relationships based on proximity. Figure 6 below visualizes the results of those networks.

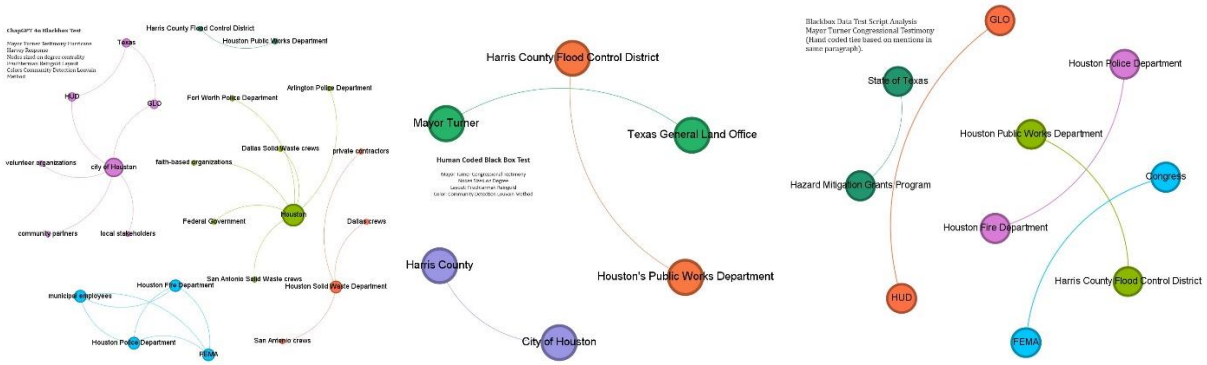



Figure 7. Black Box Test- ChatGPT 4o, Human Coded, and Script Analysis Mimic

One issue we identified in human coding was repeated data. Our human coder, through context, identified that in congressional testimony there was the oral statement (transcript) followed by the written statement submitted for the record. This was often repeated material, with minor differences. The human coder did not code twice unless new material was present. In a small data test, we compared a sample section. The human coder

[40] The establishment of dictionaries or appendices of actors and aliases is discussed in Michael Kenney and Stephen Coulthart, "The Methodological Challenges of Extracting Dark Networks: Minimizing False Positives through Ethnography," in *Illuminating Dark Networks: The Study of Clandestine Groups and Organizations*, ed. Luke Gerdes (Cambridge: Cambridge University Press, 2015); Espinal-Enríquez et al., "A Literature-Based Approach to a Narco-Network."

[41] Espinal-Enríquez et al., "A Literature-Based Approach to a Narco-Network."

generated 3 edges whereas the LLM generated 23. The increase was in part from non-unique identifiers such as "Dallas crews" volunteers from Dallas not part of a formal organization that assisted with waste disposal. The LLM coded this as a tie, the human coder did not. No evidence of hallucinating was found in this carefully checked subset.

Broadly speaking the ChatGPT LLM data coded far more actors in part because it was willing to code generic actors without unique identifiers that humans were not. This has value, but calls into question the underlying value of the social network analysis data generated. Non unique identifiers are a classic problem in data analysis for social network analysis. It can lead to our actors being artificially inflated in importance which are not unique actors such as in the Armed Conflict Event Data Project (ACLED) in which if we were to input the data into a SNA, civilians would be one of the primary actors because cartels target civilians so often. While this has an element of truth to it and value, it is not the same civilians being targeted every single time.

The human coding of data was by far the most conservative method on this subset of data. That was followed by the script method which was also coded by a human but according to typical algorithmic methods such as being mentioned in the same paragraph.[42] The LLM was much more liberal in its assigning of relationships in part because it was willing to use non unique identifiers. Interestingly, CHATGPT did not produce the exact same edges as the script analysis did. For example, the script analysis which we mimicked using a human coder, coded a connection between FEMA and Congress. This tie did not exist within the CHATGPT coded data. It coded a significant number of other ties between FEMA and other agencies but not to Congress in this subset of data. This is one example of many, demonstrating that ChatGPT 4o was doing something more than a straight algorithmic assigning of ties based on proximity. This is a fertile area for future research.

# Conclusions

Artificial intelligence in the form of large language models (LLMs), can now generate relational data from unstructured sources. This is an important finding given as recently as a year ago it would simply hallucinate actors (2023 Chat GPT 3.5).[43] Despite its recent success, it does, however, have significant shortcomings which can be mitigated by better prompting or "small batching" the data AI is asked to analyze. The use of non-unique identifiers is a double-edged sword. On the one hand the LLM is identifying new trends and actors, on the other hand non unique identifiers can weaken the value of other metrics.

[42] Espinal-Enríquez et al.
[43] Jones, "A Methodological Note."

Following Moore's law improvement in AI capacity and quantum computing, data limitations are likely to erode over time allowing the entry of larger amounts of material for analysis. These unstructured sources could include Open-Source Intelligence (OSINT), such as those analyzed here, social media, or other internal data sources. Thus, these LLMs with some finetuning that will be discussed below could provide rapid quantifiable and manipulable datasets, generating both strategic and actionable tactical intelligence on both bright and dark networks. This could have immense value in efforts to protect US critical infrastructure.

## Lessons Learned

ChatGPT 4o responded differently to smaller amounts of data and the number of documents entered. One document was optimal and no more than 15 pages. This is an interesting finding because ChatGPT itself recommends no more than 500-1000 words for optimal results. In the test cases, more edges were generated when a single document was uploaded versus four documents.

Prompting Chat GPT to be thorough and analyze every page generated more edges. Consolidated prompts helped with replicability and often resulted in more edges generated but fewer nodes or actors than when a separate prompt for a node list was requested. There seems to be a tradeoff between identifying more nodes versus edges.

The black box test demonstrated that ChatGPT 4o was doing more than just script analysis by ignoring ties that the script analysis generated. In other words, ChatGPT was not simply assigning ties based on actors mentioned in close proximity. This suggests some level of contextualization for the establishment of ties.

## Limitations

This project had numerous limitations. First and foremost, was the time constraint. It should be noted that this project will be ongoing, and this research will continue at the Institute for Homeland Security (IHS) at Sam Houston State University (SHSU). Second, this project focused on a limited amount of data when comparing across human, ChatGPT, FLANT-5 generated networks. This is because of limited human labor capacity and the tedious nature of converting unstructured data into structured data.

Aspects of this project were limited to open-source data from unstructured sources such as government reports, legal documents, congressional testimony, newspaper articles, etc. Thus, any social network analysis based upon these documents will suffer from the same underlying biases of these open-source documents, such as an unwillingness to reveal negative interactions amongst bright network actors. We tried to overcome these by

including newspaper articles which specifically described conflictive relationships, though for most of the data analyzed here we focused on positive relationships.

## Avenues for Future Research

Future research could expand the size of the data sets comparing human and artificial intelligence generated data. With a larger sample size, researchers would likely be able to glean greater insights and garner more robust results. This project used exploratory data methods and descriptive statistics. Future iterations of this research will include more hypothesis testing and robustness checks.

Future researchers could expand the number of large language models used to create networks from unstructured data. This would likely require increased computing capacity and more infrastructure, training, and models which evolve over time. These models are likely to be able to accept larger amounts of data more rapidly. Thus, what was once a tedious human process can be rapidly automated by large language models allowing us to derive clean structured datasets from vast amounts of unstructured qualitative data. This will no doubt generate new insights and the ease of application will lead to innovative questions and research areas.

Future research could apply these methodologies to internal datasets or non-open-source data. This was one of the key aspects of this project in the sense that is a proof of concept designed to be applied to more than just the specific topic here. Further, this is why the researchers chose the FLANT-5 model because it does not use outside servers in its largest language model. Thus, it provides a model for the analysis of internal non-open-source data.

Future iterations of this research could automate this process by web-scraping CI relevant documents, generating rapid data sets across all major municipalities in real time. Future research could move beyond CI Response/EM networks to utilizing these methodologies to analyze the networks embedded within the infrastructure itself. Additionally, future research could apply this to dark network threat actors such as terror networks and move beyond the criminal network analysis (CNA) provided here.

## Recommendations:

1. Following the principle of preferential attachment this project has generated data sets that show most central actors for a potential military deployment to interface with to reach the most other actors in the CI response Network. In the Houston area for past events these included (Based on ChatGPT small batch full data set results):

a. Degree Centrality (Most raw connections): FEMA, Dallas PD, Catastrophic Medical Operations Center (CMOC), City of Houston, Harris County Homeland Security and Emergency Management (HSEM), Austin Energy, City of Austin, Texas Department of Emergency Management (TDEM), Austin Water, Harris County, UT Health Houston, Texas Health Department, ERCOT, and Southeast Texas Regional Advisory Center (SETRAC).
b. Betweenness Centrality (Brokers): FEMA, Austin Energy, CMOC, Dallas PD, City of Austin, Harris County TDEM, SETRAC, Texas Department of Health, UT Health Houston, HRISC, Red Cross, Texas Department of Public Safety (DPS).
c. Eigenvector Centrality (Connected to other Important actors): FEMA, City of Austin, Austin Energy, City of Houston, HSEM, TDEM, Harris County, Dallas PD, Governor Abbott, CMOC, Texas A&M University System, Red Cross, Texas General Land Office, Harris County Flood Control District.

2. Target Subgroups for Coverage- Identify interface partners in subgroups or communities:
    a. In the Houston Area Examples include FEMA, Catastrophic Medical Operations Center (CMOC), Texas Department of Public Health, Harris County, Dallas PD, and Southeast Texas Regional Advisory Center (SETRAC).
3. Integrate artificial intelligence gathering of critical infrastructure protection datasets for social network analysis (SNA) into civil support planning.

## Acknowledgements

The researchers would like to thank Natasha Jimenez of HII, the Homeland Defense Institute, and the support of the Institute for Homeland Security (IHS) at Sam Houston State University for their funding and support during this project.

## Funding Statement

This publication was sponsored by the Homeland Defense Institute.
The views expressed in this publication do not necessarily represent the views of the United States Air Force Academy, North American Aerospace Defense Command and United States Northern Command, the Department of Defense, or the United States Government.

# Appendices

## Appendix 1: ChatGPT 4o Prompt Text

"I need you to create a CSV file with column 1 labeled source and column 2 labeled target. I then need you identify when organizations interact with each other. Organization being defined as: local, state, and federal government organization, non-governmental organizations, government programs, government units, and people who speak on behalf of organizations. Connection being defined as: working together, collaborating, communicating, and exchanging funding. Based on the definitions I gave, list instances of organization collaboration in the CSV file. To do this, write the first organization in column 1 and the organization you identified working with that organization in column 2."

# References


Arquilla, John, and David F. Ronfeldt, eds. *Networks and Netwars: The Future of Terror, Crime, and Militancy*. Santa Monica, CA: Rand Corporation, 2001.

Barabási, Albert-László. "Scale-Free Networks: A Decade and Beyond." *Science*, July 24, 2009. https://barabasi.com/f/303.pdf.

Barabási, Albert-László, and Eric Bonabeau. "Scale-Free Networks." *Scientific American*, 2003. https://www.scientificamerican.com/article/scale-free-networks/.

Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy. "Gephi: An Open Source Software for Exploring and Manipulating Networks," 3:361–62, 2009.

Blondel, Vincent D., Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. "Fast Unfolding of Communities in Large Networks." *Journal of Statistical Mechanics: Theory and Experiment* 2008, no. 10 (October 9, 2008): P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008.

Blouin, Lou. "AI's Mysterious 'Black Box' Problem, Explained." University of Michigan-Dearborn, March 6, 2023. https://umdearborn.edu/news/ais-mysterious-black-box-problem-explained.

Bright, David A, Catherine Greenhill, Michael Reynolds, Alison Ritter, and Carlo Morselli. "The Use of Actor-Level Attributes and Centrality Measures to Identify Key Actors: A Case Study of an Australian Drug Trafficking Network." *Journal of Contemporary Criminal Justice* 31, no. 3 (2015): 262–78.

Carpintero, D. "Named Entity Recognition to Enrich Text." OpenAI Cookbook, October 19, 2023. https://cookbook.openai.com/examples/named_entity_recognition_to_enrich_text.

Contreras Velasco, Oscar, Nathan P. Jones, Daniel Weisz Argomedo, John P. Sullivan, and Chris Callaghan. "The Use of Similarity-Based Algorithms to Predict Links in Mexican Criminal Networks." Research Paper. Houston: Rice University's Baker Institute, August 30, 2023. https://doi.org/10.25613/BQ88-X176.

Cunningham, Daniel, Sean Everton, and Philip Murphy. *Understanding Dark Networks: A Strategic Framework for the Use of Social Network Analysis*. London: Rowman & Littlefield, 2016.

Davenport, Thomas H., and Nitin Mittal. *All in on AI: How Smart Companies Win Big with Artificial Intelligence*. Boston: Harvard Business Review, 2023.

"Department of Defense Instruction (3025.21)." Department of Defense, February 27, 2013. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/302521p.pdf.

Espinal-Enríquez, Jesús, J Mario Siqueiros-García, Rodrigo García-Herrera, and Sergio Antonio Alcalá-Corona. "A Literature-Based Approach to a Narco-Network," 97–101. Springer, 2015.

Everton, Sean F. *Disrupting Dark Networks*. Structural Analysis in the Social Sciences 34. New York, NY: Cambridge University Press, 2012.

Freeman, Linton C. "Centrality in Social Networks Conceptual Clarification." *Social Networks* 1, no. 3 (1978): 215–39.

Freeman, Linton C. *The Development of Social Network Analysis: A Study in the Sociology of Science*. Vancouver: Empirical, 2004. https://www.researchgate.net/profile/Linton-Freeman-2/publication/239228599_The_Development_of_Social_Network_Analysis/links/54415c650cf2e6f0c0f616a8/The-Development-of-Social-Network-Analysis.pdf.

Garay Salamanca, Luis Jorge, Eduardo Salcedo-Albarán, and Isaac de León Beltrán. *Illicit Networks, Reconfiguring States: Social Network Analysis of Colombian and Mexican Cases*. Bogotá: Metodo Foundation, 2010.

Girvan, Michelle, and Mark EJ Newman. "Community Structure in Social and Biological Networks." *Proceedings of the National Academy of Sciences* 99, no. 12 (2002): 7821–26. https://doi.org/10.1073/pnas.122653799.

Jones, Nathan. "A Methodological Note: ChatGPT Is Not Ready for Criminal Network Analysis from Unstructured Data." *OODA Loop* (blog), May 8, 2023. https://www.oodaloop.com/archive/2023/05/08/a-methodological-note-chatgpt-is-not-ready-for-criminal-network-analysis-from-unstructured-data/.

Jones, Nathan, Irina Chindea, Daniel Weisz Argomedo, and John Sullivan. "A Social Network Analysis of Mexico's Dark Network Alliance Structure." *Journal of Strategic Security* 15, no. 4 (2022). https://doi.org/10.5038/1944-0472.15.4.2046.

Jones, Nathan P., W. Layne Dittmann, Jun Wu, and Tyler Reese. "A Mixed Methods Social Network Analysis of a Cross-Border Drug Network: The Fernando Sanchez Organization (FSO)." *Trends in Organized Crime* 23, no. 2 (June 1, 2020): 154–82. https://doi.org/10.1007/s12117-018-9352-9.

Jones, Nathan P, Russell Lundberg, and Matthew O'Deane. "A Mixed Methods Social Network Analysis of San Diego Law Enforcement Task Forces and Agencies." *International Journal of Police Science* 1, no. 2 (2022): 58–83. https://doi.org/10.56331/487529/IJPS6.

Kenney, Michael, and Stephen Coulthart. "The Methodological Challenges of Extracting Dark Networks: Minimizing False Positives through Ethnography." In *Illuminating Dark Networks: The Study of Clandestine Groups and Organizations*, edited by Luke Gerdes. Cambridge: Cambridge University Press, 2015.

Kimmelman, Michael. “Lessons From Hurricane Harvey: Houston’s Struggle Is America’s Tale.” *The New York Times*, November 11, 2017, sec. Climate. https://www.nytimes.com/interactive/2017/11/11/climate/houston-flooding-climate.html.

Krebs, Valdis E. “Mapping Networks of Terrorist Cells.” *Connections* 24, no. 3 (2002): 43–52.

Latapy, Matthieu. “Main-Memory Triangle Computations for Very Large (Sparse (Power-Law)) Graphs.” *Theoretical Computer Science* 407, no. 1–3 (2008): 458–73.

Lewis, Ted G. *Critical Infrastructure Protection in Homeland Security: Defending a Networked Nation*. John Wiley & Sons, 2006.

Morselli, Carlo, Cynthia Giguère, and Katia Petit. “The Efficiency/Security Trade-off in Criminal Networks.” *Social Networks* 29, no. 1 (January 1, 2007): 143–53. https://doi.org/10.1016/j.socnet.2006.05.001.

“National Drug Threat Assessment 2024.” Drug Enforcement Administration, 2024. https://www.dea.gov/sites/default/files/2024-05/NDTA_2024.pdf.

Ressler, Steve. “Social Network Analysis as an Approach to Combat Terrorism: Past, Present, and Future Research.” *Homeland Security Affairs* 2, no. 2 (2006): 1–10.

Salcedo-Albarán, Eduardo, Luis Garay, and José Cano Melani. “Machine Learning Models on Criminal Networks (MLMoCN): Artificial Intelligence to Disentangle Crime,” 2023.

Salcedo-Albaran, Eduardo, and Luis Jorge Garay Salamanca. “Structure of a Transnational Criminal Network: ‘Los Zetas’ and the Smuggling of Hydrocarbons.” Working Paper. Vortex, 2014. http://www.scivortex.org/12TCNsMexUsV2.pdf.

“Texas Severe Winter Storm DR-4586.” Accessed May 14, 2024. https://www.tdem.texas.gov/disasters/winter-storm-uri.

*The New York Times*. “Storm Carves Path of Destruction Across Houston.” May 17, 2024, sec. Weather. https://www.nytimes.com/2024/05/17/weather/houston-storm-photos-video.html.

US NORTHCOM. “Strategy.” Accessed May 18, 2024. https://www.northcom.mil/Strategy/.

Velasco, Oscar Contreras. “Unintended Consequences of State Action: How the Kingpin Strategy Transformed the Structure of Violence in Mexico’s Organized Crime.” *Trends in Organized Crime*, July 10, 2023. https://doi.org/10.1007/s12117-023-09498-x.

Wasserman, Stanley, and Katherine Faust. *Social Network Analysis: Methods and Applications*. Structural Analysis in the Social Sciences. New York: Cambridge University Press, 1994.

Wu, Jun, William Layne Dittmann, Nathan P. Jones, and John P. Sullivan. “A Social Network Analysis of an MS-13 Network: Structure, Leadership Roles, and the Use of Confidential Informants.” *International Criminology*, March 9, 2024. https://doi.org/10.1007/s43576-024-00113-9.

# INSTITUTE FOR HOMELAND SECURITY

Sam Houston State University

The Institute for Homeland Security at Sam Houston State University is focused on building strategic partnerships between public and private organizations through education and applied research ventures in the critical infrastructure sectors of Transportation, Energy, Chemical, Healthcare, and Public Health.

The Institute is a center for strategic thought with the goal of contributing to the security, resilience, and business continuity of these sectors from a Texas Homeland Security perspective. This is accomplished by facilitating collaboration activities, offering education programs, and conducting research to enhance the skills of practitioners specific to natural and human caused Homeland Security events.

Institute for Homeland Security
Sam Houston State University



Jones, Nathan, Pamfile, Christian, Dutta, Juli, Contreras Velasco, Oscar, & Aspland, Michael (2024) Artificial Intelligence and Social Network Analysis for Critical Infrastructure Response Networks and Dark Network Threat Analysis. (Report No. IHS/CR-2024-1022). The Sam Houston State University Institute for Homeland Security.
https://doi.org/10.17605/OSF.IO/9P27U