Erschienen in: 2025 IEEE Visual Analytics Science and Technology Challenge (VAST Challenge). Piscataway, NJ: IEEE, 2025, S. 15-16. ISBN 979-8-3315-9090-1
https://doi.org/10.1109/vast-challenge69463.2025.00012

OVA – Oceanus Visual Analytics

Benedict da Silva Araújo Finsterwalder* Pascal Le Hong† Daniel A. Keim‡
Julius Rauscher§ Lucas Joos¶

University of Konstanz

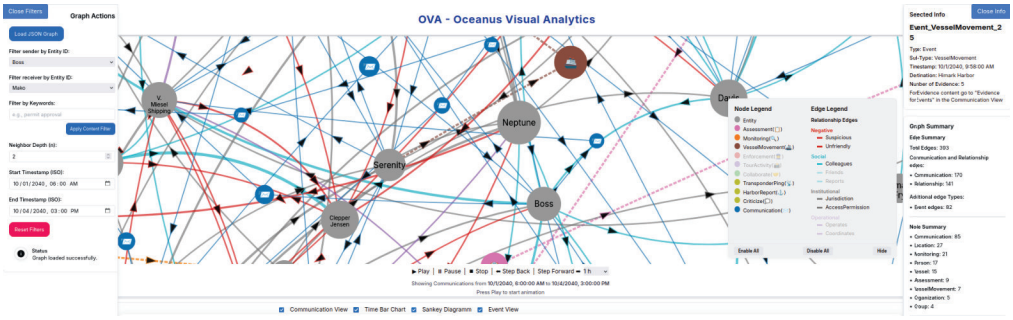

Figure 1: OVA – Oceanus Visual Analytics tool to analyze Clepper's transcribed communications.

Abstract

Mini-Challenge 3 of the VAST Challenge 2025 centers around the investigation of a covert operation taking place in and around Nemo Reef on the island of Oceanus. To facilitate the analysis, we developed a visual analytics application emphasizing transparency, objectivity, and interactivity. Our tool allows analysts to examine the provided knowledge graph through multiple coordinated visualizations. By avoiding any assumptions or pre-computations, we ensure that the information displayed remains grounded in the provided evidence. The application enables users to identify key events, relationships, pseudonyms, and influence structures, as well as to reconstruct the narrative behind the operation. Major findings include the timeline of the covert activities, the identification of pseudonyms used by key individuals, and the involvement of Nadia Conti (alias "Boss") in illegal operations.

Index Terms: VAST Challenge 2025, Mini-Challenge 3, Visual Analytics, Knowledge Graphs, Oceanus.

1 Introduction

Mini-Challenge 3 (MC3) of the VAST Challenge 2025 presents participants with the task of analyzing a complex knowledge graph constructed from intercepted radio communications. The objective is to uncover significant events, reveal influence patterns among individuals, identify groups and topical associations, detect the use of pseudonyms, and determine the extent of Nadia Conti's involvement in suspicious activities. The intercepted communications center around events occurring on and around Nemo Reef, located on the island of Oceanus.

This scenario focuses on the ongoing investigation by journalist Clepper Jensen, a former analyst at FishEye, who has been monitoring developments tied to the controversial closure of Nemo Reef. Jensen's work highlights irregularities in the permitting process, potential political interference, and the suspected participation of various actors, including the pop icon Sailor Shift, in an illicit operation. To tackle this challenge, we developed a visual analytics tool tailored to support investigative workflows. Implemented using React, D3, Neo4j, FastAPI, and Docker, our system emphasizes usability and objectivity. The tool relies exclusively on the ground truth provided in the communication content and does not perform pre-computations or incorporate assumptions that could bias interpretation. Our analysis pursues several key goals as posed by MC3: (1) identifying temporal communication patterns and changes over time; (2) detecting groups of closely connected entities and their associated topics; (3) uncovering and resolving pseudonym usage; and (4) evaluating the suspected role of Nadia Conti in the covert operation. Through this structured, evidence-driven approach, we aim to reconstruct the underlying narrative and provide actionable insights into the dynamics of the situation at Nemo Reef.

*e-mail: benedict.da-silva-araujo-finsterwalder@uni-konstanz.de
†e-mail: pascal.le-hong@uni-konstanz.de
‡e-mail: keim@uni-konstanz.de
§e-mail: julius.rauscher@uni-konstanz.de
¶e-mail: lucas.joos@uni-konstanz.de

2 Data

The dataset for MC3 includes two JSON files and a supplementary PDF document. The first JSON file contains the knowledge graph, consisting of nodes (Entities, Events, and Relationships) and their interconnecting edges. The second JSON file provides the schema defining the structure and semantics of node and edge types.

The accompanying PDF details the methodology behind the graph's construction and offers interpretation guidance. The graph was manually created by journalist Clepper Jensen and his intern based on intercepted radio communications from October 1st to October 14th, 2040. These capture interactions among individuals and vessels around Nemo Reef and other locations on Oceanus.

Although constructed with care, the documentation notes potential inaccuracies. Our validation revealed that approximately 10% of Communication Events contained incorrect sender or receiver assignments, which we addressed during our preprocessing phase.

3 Analysis Approach

Our visual analytics tool is designed to be interactive, interconnected, and user-friendly. It is intentionally developed to present only objective information. We avoid any assumptions or pre-computations that could bias interpretation.

The tool comprises several coordinated views: a node-link diagram, a Sankey diagram, a bar plot, a communication view with different tabs, and an event view. Supplementary components include a filter panel, an information panel, and a legend. The node-link diagram features an animation that helps visualize entity interactions over time. The communication view emphasizes key elements of

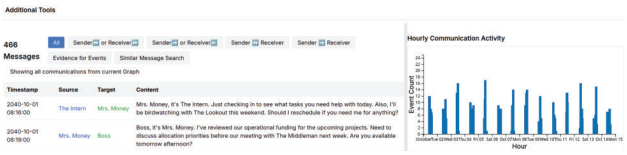

Figure 2: Left: Communication View displaying message content. Right: Hourly communication activity represented in a Time Bar Plot.

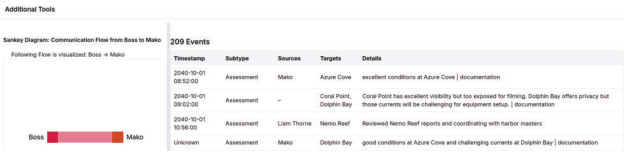

Figure 3: Left: Sankey diagram showing the communication count between entities retaining the flow direction. Right: Event Communication View displaying events and optional content.

each communication event: timestamp, source, target, and message content. Since this data represents the foundational ground truth of the knowledge graph, it is presented in its unaltered form, enabling users to validate events and draw unbiased conclusions.

Visualizations are linked, ensuring a cohesive and responsive analytical environment. For instance, selecting an entity in the node-link diagram will automatically update the "sender" filter, influencing all other views accordingly. Clicking on an event reveals its supporting evidence in the communication view, highlighting which communications were used to infer it. The information panel further displays attributes and metadata of the selected node or edge. The color coding of nodes and edges is explained in the legend and can be clicked to enable or disable them.

When a message in the Communication View is selected, similar messages are displayed. Similarity is determined by encoding the message content using the `BAAI bge-small-en-v1.5 model` and computing the distances between the resulting vector representations. Clicking on the sender or receiver sets the corresponding filters to that entity. Selecting a timestamp sets the start time for filtering, which is particularly useful in identifying pseudonyms when users want to view all subsequent messages irrespective of sender or receiver. The bar plot displays the binned hourly communication count for the entities currently visible in the filtered node-link diagram. Users can click or brush over bars to set the timestamp filter. These two visualizations can be seen in Figure 2.

The Sankey diagram visualizes the communication flow between senders and receivers. If both are unspecified, the diagram remains empty. If only one is defined, it displays all related communication flows accordingly. Clicking on a node or edge in the Sankey diagram adjusts sender and receiver filters, allowing users to explore message flows and relationships within the knowledge graph efficiently. Lastly, the event view shows all visible event nodes from the current node-link diagram, including their timestamps, sub-types, sources, targets, and details. When an event is selected, its corresponding evidence is displayed in the communication view, enabling direct validation of inferred events based on the original communication data (see Figure 3). Through this integrated and transparent approach, our tool facilitates an in-depth, unbiased exploration of the knowledge graph, empowering users to draw insights grounded firmly in the underlying data.

4 Results and Findings

In response to the tasks and questions posed by MC3, our visual analytics approach enabled us to extract several key insights from the knowledge graph. These findings are structured according to the challenge's outlined objectives.

4.1 Temporal Communication Patterns

Over the two-week observation period, we identified daily patterns in communication frequency. While the general volume fluctuates, significant changes in behavior are best understood when viewed in the context of specific groups or entities. In the first week, our analysis indicates that the entity referred to as "Boss" and a network of associated individuals coordinate preparations for an operation centered around Nemo Reef. Simultaneously, journalist Clepper Jensen, along with other authority figures, begins to detect irregularities and attempts to investigate the developing situation.

In the second week, the covert operation appears to be executed. Activities involving underwater operations and resource extraction are referenced. Although authorities become aware of suspicious activities during this period, they are unable to obtain concrete evidence. Using Clepper's knowledge graph, we found that Nadia Conti and Liam Thorne ensured that authorities were misguided.

4.2 Group Detection and Topical Associations

For the second task of finding groups and topics, we were able to extract themes and their corresponding members with the help of the similarity search and the content of the messages. For instance, we were able to find out that there are four city officials taking part in the communications, including one person named "Torres", who is not initialized as an entity in the system and therefore does not actively take part in the communication. There are many more semantic groups that we discovered.

4.3 Identification of Pseudonyms

To identify the use of pseudonyms among entities, we employed a combination of our similarity search functionality and the temporal animation feature in conjunction with the communication view. This approach allowed us to detect instances in which multiple entities shared similar communication content, context, or timing, indicating that they may in fact represent the same actor operating under different names.

Our analysis uncovered several such pseudonymous relationships. For example, Elise operates under the aliases Mrs. Money and The Accountant, Liam Thorne is also identified as The Middleman, but is not the only one using it, Sam and Kelly are associated with the pseudonyms The Intern and The Lookout, respectively, and Remora is using the alias Serenity. Many more can be found.

4.4 Activity and Role of Nadia Conti

With respect to the suspicion surrounding Nadia Conti, we found compelling evidence suggesting her renewed involvement in illicit activity on Oceanus. She appears to be instrumental in securing expedited permits and in misleading local authorities. Our analysis additionally indicates that Nadia is occasionally operating under the pseudonym Boss and plays a central role in the orchestration of the Nemo Reef operation. Communication patterns and inferred events strongly support the hypothesis that she is a key figure behind the unauthorized operations conducted in the protected reef area.

5 Conclusion

Our tool enables users to interactively explore the knowledge graph and derive insights relevant to the investigation. By presenting only objective, unprocessed data, primarily through the communication view, we empower users to perform unbiased analysis. While inferred events are included to assist in interpretation, their accuracy can be verified through supporting communication evidence.

Pseudonym identification remains a challenging task due to the potential reuse of aliases across different entities. Similarly, identifying entity groupings and thematic associations requires manual effort, though our similarity search greatly supports this process. Overall, our system facilitates a comprehensive and transparent exploration of the complex scenario presented in MC3.