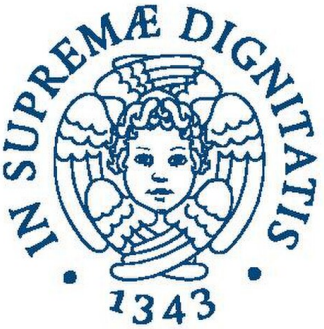

UNIVERSITÀ DI PISA

Visual Analysis of Suspicious Activities in the Oceanus Fleet

VAST Challenge 2024 Data Analysis - Mini-Challenge 2

Author: Giorgio Gobbin

Master's Degree in Data Science and Business Informatics
Student ID: 683071
Email: giorgiogobbin@gmail.com
Date: November 7, 2025

Contents

1 Introduction: Data and Suspicious Signatures **2**
1.1 VAST Challenge Context . . . 2
1.2 Data Description . . . 2
1.3 The Suspicious Signature . . . 2

2 Design Choices **2**
2.1 Layout and Interface . . . 3
2.2 Colors . . . 3
2.3 Style . . . 4
2.4 Shapes and Transformations . . . 4

3 Related Works (State of the Art) **5**
3.1 Automated Detection Systems . . . 5
3.2 Visual Analytics and Network-Based Approaches . . . 5

4 Detailed Visualization Description and Interaction **5**
4.1 Visual Components . . . 5
4.1.1 Geographic Map (Leaflet) . . . 5
4.1.2 Force-Directed Graph (D3) . . . 6
4.1.3 Timeline and Brushing (D3) . . . 6
4.2 Coordinated Analytical Workflow . . . 6

5 Use Case Example for an Analytical Task **7**
5.1 Task 1 (Q1/Q3): Identify the Signature of SouthSeafood Corp . . . 7
5.2 Task 2 (Q4): Find Other Actors with Similar Behaviors . . . 7

6 Conclusions and Reflections **8**

1 Introduction: Data and Suspicious Signatures

1.1 VAST Challenge Context

The VAST Challenge 2024 (Mini-Challenge 2) focuses on analyzing commercial fishing activities in the fictional Oceanus archipelago. The non-profit organization FishEye International suspects illegal fishing activities, particularly by the company **SouthSeafood Express Corp**. The goal of this project is to develop a visual analytics tool to analyze the provided data in order to identify, understand, and visualize patterns of illicit behavior.

1.2 Data Description

The dashboard analyzes three main datasets, loaded via D3.js:

- **mc2.json:** The main Knowledge Graph file, containing nodes (e.g., vessels, cargo documents) and links (e.g., transactions, transponder pings).
- **Oceanus Geography.geojson:** A GeoJSON file describing the archipelago's geometries, including islands, ports, and, most importantly, special interest zones (e.g., ecological preserves).
- **Oceanus Geography Nodes.json:** A mapping file that connects the graph's location nodes (e.g., `loc-123`) to geographical names (e.g., "Ghoti Preserve").

1.3 The Suspicious Signature

The primary objective of the dashboard is to visually communicate a specific pattern of suspicious behavior. Through analysis of the provided data structure and problem description, **I defined the "signature" of illegality** based on two key indicators:

1. **Incursions into Prohibited Zones:** The strongest indicator. This occurs when transponder pings from a vessel are physically located within a polygon geometry defined as an "Ecological Preserve" (the red zones on the map).
2. **Transponder Gaps:** Prolonged interruptions in signal transmission. In the implemented system, a gap is defined as a time difference greater than **12 hours** between two consecutive pings from the same vessel.

The dashboard is designed to allow an analyst to immediately see *which* companies, *where* (in which zones), and *when* (during which periods) these suspicious behaviors occur.

2 Design Choices

The interface was designed following the "Overview first, zoom and filter, then details-on-demand" (Shneiderman [1]) paradigm through a Coordinated Multiple Views system.

2.1 Layout and Interface

The dashboard (visible in Figure 1) is divided into three main areas:

- **Control Panel (Left):** Contains global filters (company search, date filter, top N companies per ping, and a reset visualization button) and the details panel (`#info-box`), which updates contextually based on the user's selection.
- **Geographic Map (Right, Top):** The main view (Leaflet.js) that provides spatial context.
- **Bottom Widgets (Right, Bottom):** Two coordinated views: a force-directed graph for relationship analysis and a timeline for temporal context.

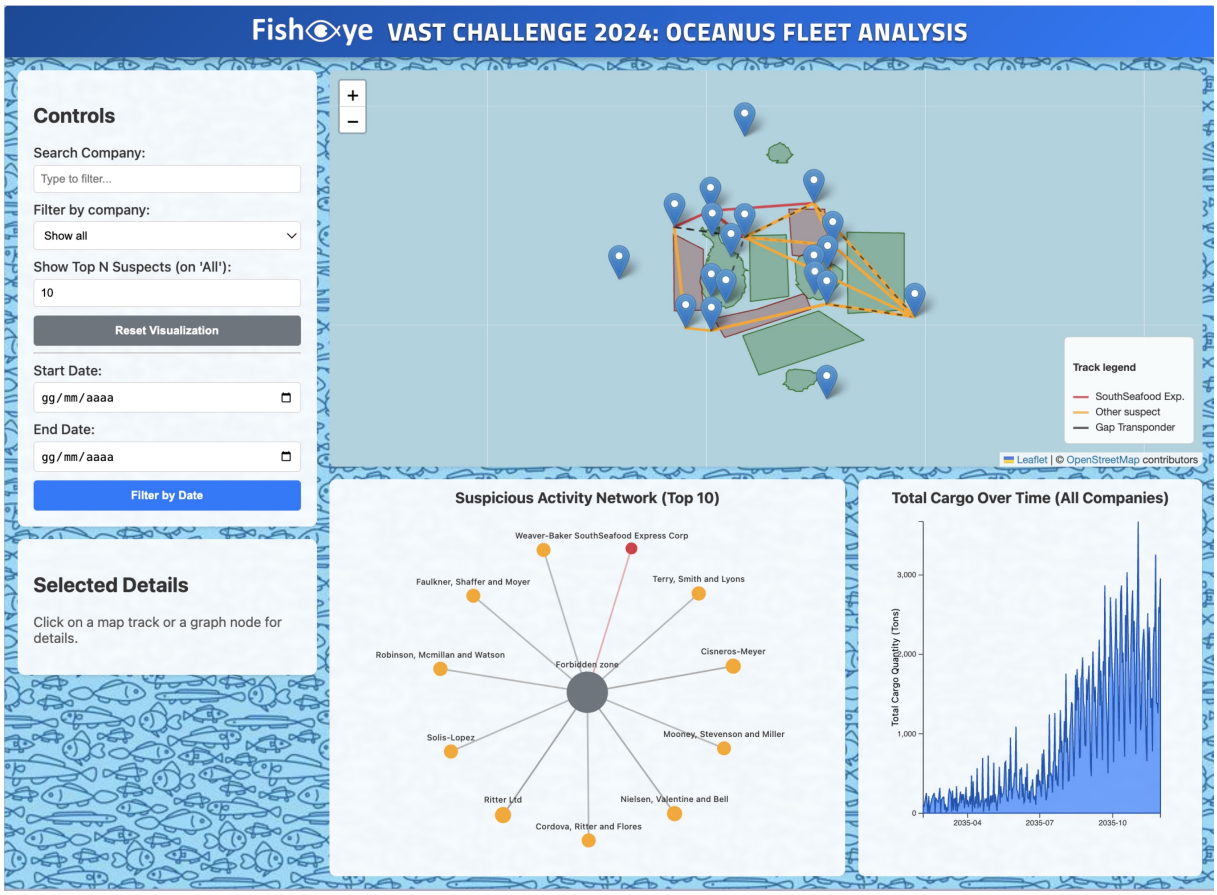

Figure 1: Overview of the dashboard with its coordinated multiple views.

2.2 Colors

The choice of colors is fundamental for the pre-attentive encoding of data. I used a semantic palette:

- **Red (#dc3545):** Used exclusively for **SouthSeafood Express Corp**. I selected red to immediately mark the primary investigative target across all views (tracks on the map, nodes in the graph), leveraging the unconscious association of red with caution or error.

- **Orange (#FFA500):** Used for **Other Suspects**. A company or vessel is colored orange if it has recorded at least one ping in a prohibited zone. This choice draws the analyst's attention to other actors exhibiting similar high-risk behavior, akin to an "attention" or "warning" sign.
- **Blue (#007bff):** The default color for non-suspicious vessels or for aggregated data (e.g., "Total Cargo" in the timeline). This is the dashboard's main theme color, chosen to maintain a color scheme that evokes the ocean, a philosophy also reflected in the stylized fish background.
- **Dashed Black (#222222):** Used on the map to indicate transponder "Gaps", visually highlighting the second suspicious behavior.

2.3 Style

The project's general typographic style relies on a modern system font (`system-ui`) to ensure maximum legibility and native performance, applied to the entire document body (`html, body`) with a standard dark gray text color (`#333`). The most distinctive text element is the main application title (`header h1`). For this title, a specific font, "Titillium Web," is imported from Google Fonts, giving it a more technical and appealing look. This title is made visually prominent through several CSS rules: it has a significantly larger size (`1.8em`), a bold weight (`700`), and is transformed entirely to uppercase (`text-transform: uppercase`). Furthermore, it features increased letter spacing (`letter-spacing: 1px`) and a complex text shadow (`text-shadow`) to ensure its legibility and give it depth against the header's gradient background. The header also contains the FishEye logo to maintain a direct relationship with the investigating organization. Other headings in the dashboard, such as those for widgets (`h3`), are more subdued, with a size of `1.1em` and the standard `#333` color. Control labels (`label`) have a medium weight (`font-weight: 500`), while texts specific to the visualizations, like the node labels in the graph, are very small (`10px`) to avoid overcrowding the chart.

2.4 Shapes and Transformations

- **Map (Leaflet):** Vessel tracks are rendered as `L.polyline`. Prohibited zones (`Ecological Preserve`) are semi-transparent red polygons to allow visibility of the tracks underneath.
- **Force-Directed Graph (D3):** The raw data (many pings) is transformed into an **aggregated bipartite network** (Companies ↔ Zones). The nodes are circular and represent the companies or the zones. Their size is mapped to the square root of the number of pings, normalizing visual perception. Nodes are linked to the Forbidden zone with gray links.
- **Timeline (D3):** The data (pings or cargo) is aggregated by day and displayed as an area chart, which is effective for showing trends and peaks over time. If the analyst filters by a company, the graph switches to display the number of suspicious pings for that company over time.

3 Related Works (State of the Art)

Research on visual analytics for maritime surveillance and illegal, unreported, and unregulated (IUU) fishing focuses mainly on two complementary directions: automated anomaly detection and interactive human-centered visualization.

3.1 Automated Detection Systems

Systems such as Global Fishing Watch integrate AIS and satellite data (e.g., SAR imagery) to identify vessels operating with transponders turned off or deviating from normal routes [4]. Recent studies show that over 6 % of global fishing activity may occur while AIS is intentionally disabled [5]. Machine learning models, including CNN and LSTM architectures, have also been proposed to classify vessel behavior and infer possible illegal activities [6]. While effective at large scale, these approaches often lack interpretability and interactive support for analysts.

3.2 Visual Analytics and Network-Based Approaches

Methods that emphasize human-in-the-loop exploration of spatio-temporal and relational data. Graph and map-based systems allow analysts to detect collaboration networks or transshipment links between vessels [7], while anomaly-visualization tools highlight temporal gaps or spatial incursions that may indicate illicit behavior [8]. These systems support sense-making through coordinated multiple views and direct interaction rather than full automation.

In this context, the developed dashboard aligns with the second approach: it empowers the analyst to explore spatial, temporal, and relational patterns interactively, combining automated data processing with visual reasoning to validate suspicious behaviors.

4 Detailed Visualization Description and Interaction

The dashboard's power lies in the coordinated interaction between the views.

4.1 Visual Components

4.1.1 Geographic Map (Leaflet)

- **Visualization:** The map shows the geography of Oceanus on a World map. I chose to show the relevant area immediately, but the analyst can zoom in or out at their discretion. Prohibited zones are in red, other zones are in green because they are not restricted. Vessel tracks are drawn and semantically colored (red, orange). Transponder gaps are dashed lines. Ports are displayed as blue pins.
- **Interaction:**
  - **Click on Prohibited Zone:** Selects the zone and opens a popup in the details panel (`#info-box`) showing: 1) The fish species present in that zone (manually mapped) and 2) A list of all vessels that pinged in that zone (filtered by date), grouped by company.

- **Click on Track:** This is a key interaction. Clicking on a track **filters the entire dashboard** (force graph, timeline, and info panel) by that vessel's *company*, not just the individual vessel. This facilitates high-level, company-based analysis.
- **Hover:** If the analyst hovers the mouse over an area or port, the name of the place appears as a tooltip.

4.1.2 Force-Directed Graph (D3)

- **Visualization:** Shows a "suspicion flow" network. Nodes are Companies (red/orange) or Zones (gray). Links represent the number of pings from a company in a zone. The size of the node is proportional to the number of pings.
- **Interaction (Hover & Click):**
  - **Hover on Node:** Triggers a "fade-out" of other nodes, clearly highlighting the selected node and its direct connections. A tooltip shows the total ping count and the entity's name.
  - **Click on Node (Company):** Like the map, this filters the entire dashboard for that company.

4.1.3 Timeline and Brushing (D3)

- **Visualization (Modal):**
  - **Default Mode:** Shows the "Total Cargo Over Time" (in tons), addressing Q2 of the challenge (associating cargo).
  - **Filtered Mode:** Shows the "Number of Suspicious Pings" for the selected company, isolating the temporal pattern of illicit behavior.
- **Interaction:** The user can hover the mouse over the graph, and a tooltip shows the date and the associated cargo quantity or ping count. The user can also use the **brush** to select a specific date range, which then acts as a temporal filter for the Map and the Force Graph. The corresponding start and end dates are updated in the control panel.

4.2 Coordinated Analytical Workflow

The analytical workflow is driven by a "global filter" (the selected company) and a "temporal filter" (the brush).

1. The analyst gains an **overview** (Map and Force Graph in "All" mode). They immediately identify the red node (SouthSeafood) and its strong links to the forbidden zones.
2. The analyst **filters** by clicking the "SouthSeafood" node.
3. The **map** updates, showing only the red tracks.
4. The **timeline** updates, showing peaks of suspicious pings for SouthSeafood.

5. The analyst **zooms** in time by using the brush or date inputs. The graph visually indicates the chosen period with two red dotted lines, allowing for a focused analysis of the map data for that exact period.

5 Use Case Example for an Analytical Task

I describe a use case that answers questions Q1, Q3, and Q4 of the VAST Challenge.

5.1 Task 1 (Q1/Q3): Identify the Signature of SouthSeafood Corp

1. **Initial Analysis:** On load (Figure 1), the analyst observes the Force Graph. The SouthSeafood Express Corp node is red and has a visible link to the "Forbidden zone" node. This immediately identifies the primary suspect.

2. **Filter by Company:** The analyst clicks the "SouthSeafood" node in the graph or the red track on the map.

3. **Filter Result (Figure 2):**

   - The **Info Panel** updates, showing the total number of suspicious pings for the entire company (e.g., "Found 150 pings...").
   - The **Map** shows only SouthSeafood's tracks (red), many of which visibly enter the prohibited zones (red polygons). Dashed black lines are also visible, indicating transponder gaps.
   - The **Timeline** switches to "Suspicious Pings: SouthSeafood", highlighting the peaks of illegal activity over time.

4. **Conclusion (Q1/Q3):** SouthSeafood's "signature" is a combination of (a) frequent incursions into the "Ghoti Preserve" zone and (b) the pervasive use of transponder gaps.

5.2 Task 2 (Q4): Find Other Actors with Similar Behaviors

1. **Reset and Observe:** The analyst presses "Reset Visualization" to return to the global view.

2. **Identification (Graph):** Observing the Force Graph, the analyst notices another node, this time **Orange** (e.g., "Ritter LTD"), which is also connected to "Forbidden zone".

3. **Filter by New Suspect:** The analyst clicks the orange "Ritter LTD." node.

4. **Filter Result (Figure 3):**

   - The **Map** updates, now showing only the orange tracks of "Ritter LTD". The analyst can visually confirm that this company also entered the prohibited zones.
   - The **Timeline** shows the suspicious pings for "Ritter LTD", allowing for a temporal comparison with SouthSeafood.

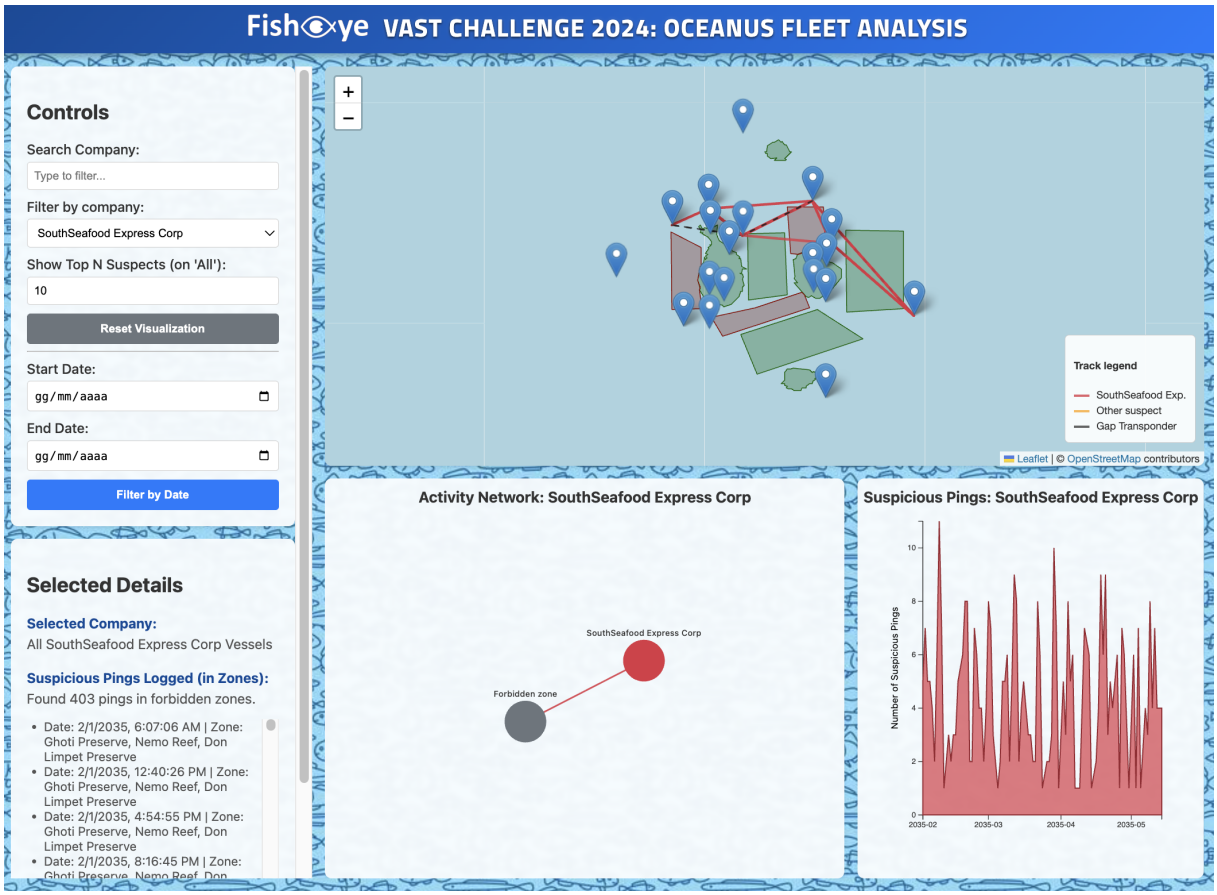

Figure 2: Dashboard filtered for "SouthSeafood Express Corp". The map shows only red tracks, and the timeline shows their suspicious pings.

5. **Conclusion (Q4):** The workflow (Reset → Observe Graph → Filter on Orange Node) allows for the discovery and visual validation of other actors exhibiting behaviors similar to SouthSeafood's, providing FishEye with new investigative leads.

6 Conclusions and Reflections

This project resulted in an interactive dashboard that effectively supports the analytical tasks required by the VAST Challenge MC2.

Reflection (Data Version): I chose the `mc2.json` version because, although structurally more complex, it was the only one containing the necessary relationships to connect companies to pings and cargo to transactions, allowing me to fully address all challenge questions.

Reflection (New Techniques): The challenge necessitated the development of hybrid visualization techniques. Instead of a generic graph (which would have been visually unreadable), I implemented an *aggregated* and *task-specific* graph view (Companies ↔ Zones). The most innovative interaction is that clicking a map track filters the entire analysis not by the single entity (vessel), but by its parent class (company),

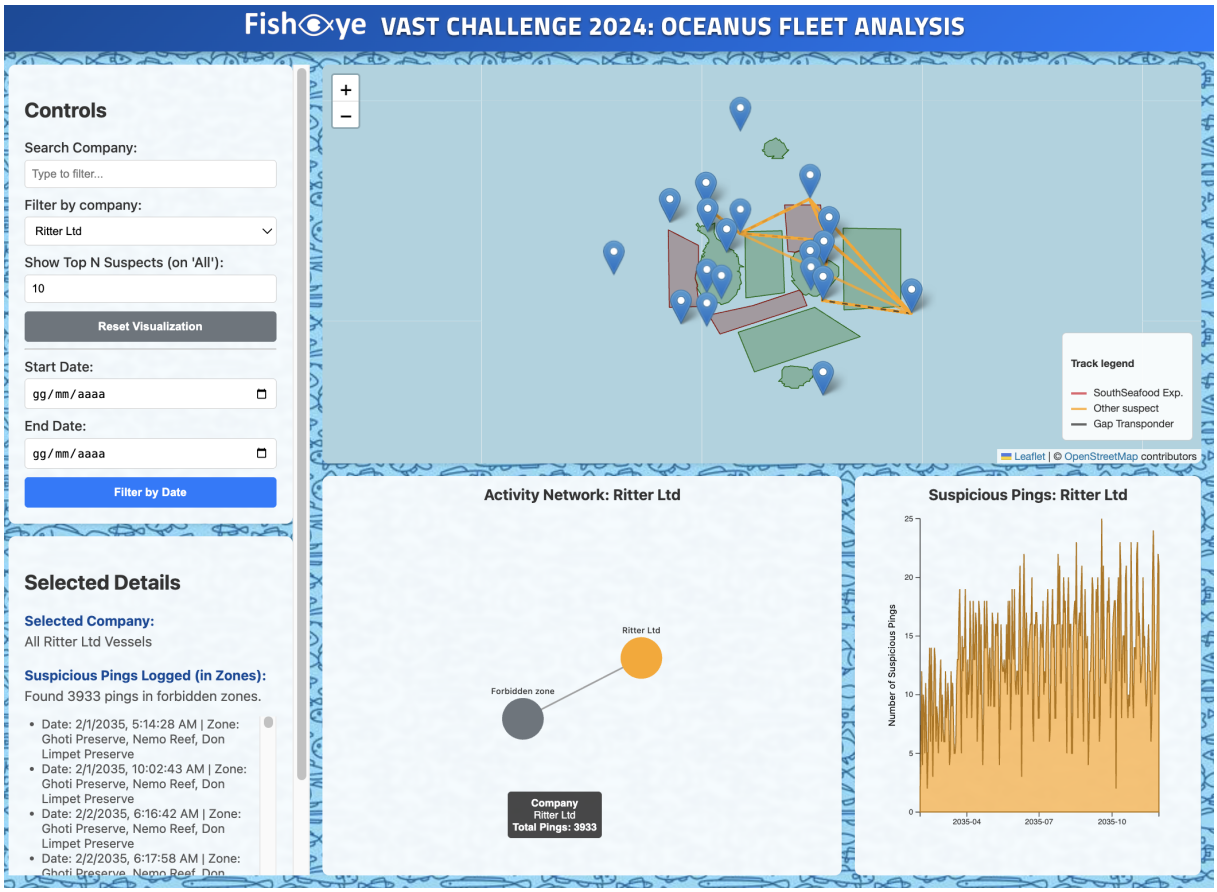

Figure 3: Dashboard filtered for another suspect (e.g., "Ritter LTD"). The orange tracks visually confirm similar behavior.

unifying the analysis at a higher level of abstraction.

Reflection (Most Difficult Part): The greatest difficulty was managing question Q2 (cargo), as the "port-exit" records were inherently disconnected from individual vessel tracks. My pragmatic solution was to have the timeline in its default mode show the total cargo trend, allowing the analyst to *infer* a correlation by visually comparing these peaks with a company's suspicious pings timeline. Another difficulty was the lack of explicit data on the fish species collected by specific companies. This information would have significantly improved the comparative analysis. However, I included the species present in the forbidden zones so that the analyst can estimate the potential risk based on the routes visible on the map. In conclusion, the developed dashboard provides a powerful analytical workflow for FishEye to efficiently identify, analyze, and discover new suspicious behaviors within the Oceanus ecosystem.

References

[1] Shneiderman, B. (1996). *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations.* In Proceedings of the IEEE Symposium on Visual Languages (pp. 336-343).

[2] Bostock, M., Ogievetsky, V., & Heer, J. (2011). *D3: Data-Driven Documents*. IEEE Transactions on Visualization and Computer Graphics, 17(12), 2301-2309.

[3] VAST Challenge 2024. *Mini-Challenge 2: Oceanus*. Available online.

[4] Global Fishing Watch. (2024). *Joint Analytical Cell: Data, Intelligence and Tools to Combat Illegal Fishing*. Available online: `https://globalfishingwatch.org/`.

[5] Welch, H., Kroodsma, D., & Miller, N. A. (2022). *Hot Spots of Unseen Fishing Vessels*. Science Advances, 8(32), eabm5307.

[6] Zhao, Q., Chen, Y., & Liu, H. (2024). *Fishing Vessel Behavior Classification Using CNN-LSTM Models*. Applied Sciences, 14(5), 2294.

[7] Park, A. J., & Stamato, S. Z. (2020). *Social Network Analysis of Global Transshipment: A Framework for Discovering Illegal Fishing Networks*. In *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2020.

[8] Rodríguez, J. P., Fernández, E., & Pérez, D. (2022). *Identification of Suspicious Behaviour Through Anomalies in the Tracking Data of Fishing Vessels*. arXiv preprint, arXiv:2204.06781.