CHAPTER 4: DIVING BENEATH THE SURFACE

Mission 4: Pattern Detection

OPERATION NAME: NEURAL NEXUS
CLASSIFICATION: TLP RED
MISSION FILE: CHAPTER 4 // DIVING BENEATH THE SURFACE

Analysts:
Recent collection has linked the Phantom Group to open source material including posts, shipping records, and metadata. Alone, none of it confirms malicious activity. Together, it suggests a coordinated effort to blend illicit behavior into legitimate online noise.
Use machine learning pattern detection to uncover anomalies that may indicate the Phantom Group is preparing for future operations. Use a dataset you already monitor, or the practice files attached to this mission.

Director of AI Intelligence Agency
Office of Intelligence

FILES Practice datasets (optional)

File	Download name	What's in it
Port calls, Jun to Sep 2026	ch4_port_calls.csv	493 fictional vessel port calls: identifiers, operators, ports, AIS gaps, declared cargo.
Domain registrations, Aug to Oct 2026	ch4_domain_registrations.csv	54 fictional domain registrations: registrar, hosting, IP, TLS certificate hash, timing.

Download these files from the worksheet page at deepinsightbook.com.

Everything here is fictional and built so it can't resolve to anything real. Domains use .example, IPs and ASNs are documentation ranges, and the IMO numbers deliberately fail the checksum.

Analyst (that's you)

Date

Tools used

Dataset(s) used

ML technique (clustering, anomaly detection...)

ANOMALIES Document two or three

Anomaly 1

What makes it unusual?

Why could it be relevant to Phantom Group activity?

What additional collection would validate it?

Confidence: ○ Low ○ Moderate ○ High

Anomaly 2

What makes it unusual?

Why could it be relevant to Phantom Group activity?

What additional collection would validate it?

Confidence: ○ Low ○ Moderate ○ High

Anomaly 3

What makes it unusual?

Why could it be relevant to Phantom Group activity?

What additional collection would validate it?

Confidence: ○ Low ○ Moderate ○ High

BIAS CHECK Compare across subsets

Subsets compared (region, language, timeframe...)

- ☐ Missing geographic coverage
- ☐ Language or translation bias
- ☐ Assumptions baked into the tool
- ☐ Collection artifact mistaken for behavior
- ☐ Other

The bias concern I found, and what it changes about my findings

WRITEUP Recommendations and limits

Recommended next steps for collection or monitoring

Limitations or uncertainties with ML use here