Project Sentient Nexus: AI-Driven Knowledge Integration for Combat System Interoperability

Groundbreaker Solutions LLC
Jason L. Lind <> jason@groundbreaker.solutions <> +1 414.704.0718
5 February 2025

Introduction

Modern naval combat operations rely on the seamless integration of diverse combat systems, each with unique data formats, interfaces, and operational protocols. However, the lack of a unified framework for data normalization and interoperability creates significant challenges in real-time decision-making, sensor fusion, and command and control (C2) coordination. Current approaches to manually developing a Domain Specific Model (DSM) and ontology for combat systems are prohibitively complex, labor-intensive, and incapable of keeping pace with rapidly evolving battlefield technologies. This fragmentation leads to inefficiencies in data exchange, increased system integration costs, and vulnerabilities in mission-critical applications.

Our proposed architecture directly addresses these challenges by employing an AI-driven approach to automatically generate and maintain a standardized DSM. By leveraging advanced machine learning (ML) techniques, natural language processing (NLP), and ontological reasoning, our solution extracts, structures, and harmonizes data from multiple disparate sources. This ensures that all combat system elements—sensors, weapons, and communication systems—adhere to a unified data model, enabling real-time interoperability across Navy, Department of Defense (DoD), and allied forces. By transforming unstructured technical documentation into machine-readable JSON/XML schemas, we eliminate the need for extensive manual data normalization and provide an adaptable, future-proof foundation for system integration.

A key innovation in our approach is the incorporation of multiple schema frameworks beyond the traditional NIEM standard, including UCore for rapid data exchange, S1000D for structured technical documentation, NAF/DoDAF for mission architecture alignment, RDF/JSON-LD for semantic reasoning, and STIX/TAXII for cybersecurity readiness. By integrating these schemas, our solution not only enhances interoperability but also improves decision support through AI-driven knowledge graph generation and real-time threat intelligence processing. This multi-schema strategy ensures our AI-generated DSM remains flexible, scalable, and resilient against emerging cyber and operational threats.

Furthermore, our architecture introduces an automated pipeline that transforms arbitrary unstructured content—including PDFs, technical reports, and operational logs—into structured, queryable RDF/OWL-based ontologies. This enables seamless information retrieval, reasoning, and predictive analytics within military command networks. By integrating with the Stardog graph database, our system provides a dynamic, continuously evolving representation of combat system relationships, supporting real-time situational awareness and mission adaptability. The ability to convert SPARQL query results into NIEM XML/UCore JSON outputs ensures smooth integration with legacy and next-generation combat systems alike.

By automating the generation of combat system DSMs and ontologies, our solution significantly reduces the time, cost, and complexity of system integration. It establishes a standardized framework for future combat system developments while enhancing operational readiness and decision-making capabilities. As the Navy continues to prioritize trusted AI and autonomy in its modernization efforts, our AI-driven architecture provides a critical enabler for achieving seamless combat system interoperability and enhancing mission effectiveness in an increasingly complex and contested battlespace.

Phase I Technical Objectives

1. Develop a Concept for AI-Generated DSM and Ontology

Objective: Establish a methodology for automatically extracting, structuring, and standardizing a Domain Specific Model (DSM) and ontology from combat system technical documentation.

Our Solution:

Utilize Natural Language Processing (NLP) and Machine Learning (ML) models to extract structured knowledge from diverse unstructured sources (PDFs, XML, JSON, and text-based reports).

Implement ontology reasoning frameworks (RDF, OWL, JSON-LD) to define relationships between combat system components.

Develop a prototype AI-driven data extraction pipeline that converts raw data into a standardized machine-readable format.

2. Demonstrate Feasibility of AI-Driven DSM Extraction

Objective: Validate that AI-based techniques can successfully extract and normalize combat system data from heterogeneous sources.

Our Solution:

Implement multi-schema integration using NIEM (National Information Exchange Model), UCore, S1000D, NAF/DoDAF, RDF, and STIX, ensuring compatibility with DoD and allied systems.

Apply AI-based entity recognition models (spaCy, BERT, Stanford NLP) to extract critical data elements such as sensor IDs, weapon types, and communication protocols.

Test feasibility through modeling and simulation, mapping extracted data into NIEM-compliant JSON/XML schemas to ensure traceability and auditability.

3. Enable Automated Data Normalization and Interoperability

Objective: Ensure that extracted data is structured into a common format to facilitate seamless interoperability across combat systems.

Our Solution:

Design a rule-based transformation engine to convert extracted data into standardized NIEM, RDF, and OWL formats.

Implement SHACL validation to enforce schema consistency and data integrity across combat system components.

Leverage SPARQL query capabilities to validate extracted relationships, ensuring accurate representation of real-world combat system structures.

4. Develop Initial AI Models for Ontology Creation and Data Integration

Objective: Build and test AI models that can infer relationships between combat system elements and align them with existing military data frameworks.

Our Solution:

Use ontology reasoning techniques (Description Logic, Semantic Web Technologies) to define relationships between sensors, weapons, and C2 systems.

Integrate Stardog knowledge graph database to provide a dynamic, queryable combat system representation with real-time inferencing.

Implement TAXII/STIX cybersecurity intelligence integration to identify potential cyber threats and enable real-time anomaly detection.

5. Provide Initial Prototype and Testing Plan for Phase II Development

Objective: Deliver a proof-of-concept model that demonstrates the ability to extract, process, and integrate data from diverse combat system sources.

Our Solution:

Develop an AI-powered ingestion and transformation pipeline that accepts raw technical documentation and produces NIEM-compliant outputs.

Simulate real-world data exchange scenarios with combat system sensor, weapon, and communication data sources.

Establish a roadmap for Phase II prototype development, including advanced AI training, real-world data trials, and user feedback-driven improvements.

Conclusion

Our AI-driven approach directly addresses Phase I technical objectives by automating DSM extraction, enabling seamless data normalization, and integrating multi-schema interoperability frameworks. Through a combination of NLP, ML, semantic reasoning, and cybersecurity-aware AI models, we establish a robust foundation for combat system integration, reducing the complexity, cost, and time required for interoperability efforts across the Navy and DoD.

Phase I Statement of Work

Multi-Schema Integration Strategy for AI-Generated DSM

1. Overview

To enhance the AI-generated **Domain Specific Model (DSM)** and ensure interoperability across **Navy, DoD, and allied combat systems**, this strategy incorporates **multiple schemas** beyond NIEM. The integration of **UCore, S1000D, NAF/DoDAF, RDF/JSON-LD, and STIX** will enable better **data exchange, ontology reasoning, and cybersecurity readiness**.

2. Core Schema: NIEM

Role of NIEM in the AI-Generated DSM

NIEM serves as the **foundational schema** for structured data exchange, ensuring:

- **Standardization of combat system interfaces (sensors, weapons, C2 systems)**

- **Interoperability with existing DoD information-sharing frameworks**
- **NIEM IEPD compliance for machine-readable AI outputs (JSON, XML)**

NIEM Integration Approach

- AI-extracted combat system elements will be mapped to **NIEM Military Operations (MilOps) and Maritime Domains**.
- AI-generated DSM will be structured into **NIEM-compliant JSON/XML schemas**.
- The system will ensure **traceability & auditability** through NIEM's nc:SourceDocument metadata.

3. Supplementary Schemas & Their Roles

3.1. UCore (Universal Core)

Purpose: Lightweight, rapid data exchange standard. **Benefit:** Enhances **real-time interoperability** between **combat systems & C2 nodes**. **Implementation:**

- AI-generated SITREPs will be formatted using **UCore XML structures** for faster transmission.
- AI-extracted DSM components will **map to UCore entity structures for rapid information sharing**.

3.2. S1000D (Technical Documentation Standard)

Purpose: Standard for **defense technical documentation**. **Benefit:** Improves **AI-driven data extraction** from **maintenance & system manuals**. **Implementation:**

- AI **parses S1000D XML content** to extract structured knowledge for DSM.
- AI-generated DSM components will align with **S1000D-based system attributes**.

3.3. NAF/DoDAF (Military Architecture Frameworks)

Purpose: Standard for **military system architecture representation**. **Benefit:** Ensures **DSM alignment with NATO & DoD mission planning frameworks**. **Implementation:**

- AI-generated **combat system relationships will be structured into DoDAF Operational Views (OVs) & System Views (SVs)**.
- DSM will integrate into **NAF 4.0-compliant architecture documentation**.

3.4. RDF/JSON-LD/OWL (Ontology & AI Reasoning)

Purpose: Enables **semantic reasoning & AI-driven knowledge graphs**. **Benefit:** Improves AI-driven **combat system ontology creation & reasoning**. **Implementation:**

- AI-generated ontology will be encoded in **RDF/OWL format for inferencing**.
- AI models will support **SPARQL queries for combat system relationship discovery**.
- JSON-LD will be used for **linked data integration** with external military datasets.

3.5. STIX/TAXII (Cyber Threat Intelligence Standard)

Purpose: Facilitates **cyber threat intelligence & anomaly detection**. **Benefit:** Enhances **DSM's cybersecurity integration for real-time threat awareness**. **Implementation:**

- AI-generated DSM will produce **STIX reports for combat system cybersecurity events**.
- TAXII integration will allow **secure threat data sharing** with DoD security platforms.

4. Pipeline for Transforming Arbitrary Content to NIEM and Stardog

4.1. Data Ingestion

- Accept input from **unstructured content (PDF, text, reports, JSON, XML)**.
- AI-based **NLP models (spaCy, BERT, Stanford NLP)** extract entities from content.
- Structure extracted entities using **predefined NIEM entity mappings**.

4.2. Transformation to NIEM Model

- Use **rule-based entity mapping** to align extracted data with **NIEM XML schemas**.
- Apply **SHACL validation** to ensure NIEM-compliance.
- Example transformation:
  - **Input Text:** "The USS Zumwalt is equipped with an AN/SPY-3 radar."
  - **NIEM XML Output:**

```
<mil:CombatSystemComponent>
   <mil:SystemID>USS-Zumwalt</mil:SystemID>
   <mil:Sensor>
      <mil:SensorID>AN/SPY-3</mil:SensorID>
      <mil:Type>Radar</mil:Type>
   </mil:Sensor>
</mil:CombatSystemComponent>
```

4.3. Conversion to RDF/OWL for Stardog

- Convert NIEM XML to **RDF triples** for persistence in Stardog.
- Use **XSLT transformations** or **AI-based ontology alignment (OntoAlign, LogMap)**.
- Example RDF output:

```
:USS-Zumwalt a :CombatSystemComponent ;
   :hasSensor :AN-SPY-3 .

:AN-SPY-3 a :Radar ;
   :sensorType "AN/SPY-3" .
```

4.4. Querying Stardog via SPARQL

- Convert NIEM queries into SPARQL.

- Example SPARQL query to find all radars on combat systems:

```
SELECT ?system ?sensor WHERE {
  ?system a :CombatSystemComponent ;
        :hasSensor ?sensor .
  ?sensor a :Radar .
}
```

4.5. Response Transformation to NIEM/UCore

- Convert SPARQL query results into **NIEM XML/UCore JSON** for external system consumption.
- Example NIEM-compliant XML response:

```
<mil:CombatSystemComponent>
  <mil:SystemID>USS-Zumwalt</mil:SystemID>
  <mil:Sensor>
    <mil:SensorID>AN/SPY-3</mil:SensorID>
    <mil:Type>Radar</mil:Type>
  </mil:Sensor>
</mil:CombatSystemComponent>
```

5. Conclusion

Supporting additional schemas **beyond NIEM** ensures the AI-generated DSM is **flexible, secure, and widely applicable**. By integrating **UCore, S1000D, NAF/DoDAF, RDF, and STIX**, the solution will:

- Provide **multi-domain interoperability** for combat systems.
- Enhance **machine-readable AI-driven ontology generation**.
- Improve **cybersecurity resilience & threat detection capabilities**.

This multi-schema strategy will **increase adoption potential** for both **defense and civilian applications**, ensuring **long-term viability** of the AI-driven DSM in Phase II and Phase III deployments.

Comprehensive Assessment & Validation

Simulation-Based Testing and Metrics

To demonstrate the feasibility and robustness of the AI-generated Domain Specific Model (DSM), we will initially conduct simulation-based testing using representative data sets derived from unclassified combat system scenarios. In this stage, we will create controlled test cases that mirror real naval operations, covering various sensor outputs, communication protocols, and weapon configurations. Our simulation framework will stress-test the DSM on data volume, data variety (e.g., structured vs. unstructured inputs), and timing constraints relevant to combat operations. We will track key performance metrics—such as accuracy of entity extraction, latency in AI pipeline processing, throughput under high data loads, and resource utilization—to measure system performance. These metrics will be monitored through automated dashboards, ensuring a transparent view of the AI’s decision-making process and its resource demands. This approach allows us to quickly identify bottlenecks, validate schema coverage, and measure improvements as we iterate on the AI models.

Real-World Data Integration Trials

Building on the findings from simulation-based testing, the next phase involves real-world data integration trials. During this phase, we will ingest and normalize data feeds from actual or near-real combat system logs, technical documentation, and operational reports (as permitted under security constraints). We will collaborate with Subject Matter Experts (SMEs) to ensure the data sets accurately reflect current naval platforms, such as emerging weapon systems and sensor arrays. By validating our ontology against real operational data, we can assess how effectively the AI pipeline handles edge cases, newly introduced formats, and incomplete or noisy data. This iterative process will allow us to refine both the AI extraction logic and the underlying ontology, ultimately ensuring that the resulting DSM is resilient, up to date, and capable of scaling to the Navy's evolving requirements.

Synchronization with NIEM-Compliant Systems

A critical aspect of our validation effort is confirming that the AI-generated DSM—and all associated data transformations—correctly align with NIEM (National Information Exchange Model) standards. To achieve this, we will implement a two-stage NIEM compliance check. First, our transformation engine will convert extracted data into NIEM XML or JSON (leveraging the MilOps and Maritime domains, among others) and validate it against NIEM schemas using an automated schema validation tool (e.g., SHACL rules or NIEM-specific validators). Second, we will design a test harness that simulates a typical NIEM-compliant data consumer—such as an external Navy or DoD system—that requires strict NIEM-conforming inputs. By passing AI-extracted data from our pipeline into this simulated consumer, we will measure interoperability, error rates, and completeness of the exchanged information. This process also includes round-trip testing: converting data to NIEM-compliant formats, ingesting it into the knowledge graph, and then re-exporting it to confirm fidelity of the original data. Incorporating NIEM validation throughout the pipeline ensures that when the DSM is deployed at scale, it can seamlessly communicate with existing and future NIEM-based Navy systems.

User Feedback and Trustworthiness Validation

Throughout these stages, we will integrate operator and SME feedback via targeted review sessions and user acceptance testing. Our platform will track user interactions with the transformed data, highlighting points of confusion or mismatch between user expectations and AI outputs. We will conduct structured interviews and distribute usability surveys to gather qualitative feedback on DSM clarity, ease of data queries, and interface readability. Additionally, to support trusted AI initiatives, we will implement transparency measures such as confidence scoring, explainable AI features, and audit logs that capture how the system arrived at specific schema mappings. These tools will help build user confidence, enable continuous improvement of the AI models, and bolster compliance with DoD "Responsible AI" directives. The combined results of simulation-based testing, real-world trials, NIEM synchronization, and user feedback will form a comprehensive validation loop, confirming the system's accuracy, efficiency, and readiness for operational deployment.

Principal Investigator – Jason L. Lind : President / Chief Architect @ Groundbreaker Solutions LLC

Jason Lind's extensive experience in **canonical data model (CDM) development, AI-driven architectures, and large-scale system integration** makes him uniquely suited to lead the AI-generated Domain Specific Model (DSM) initiative for combat system interoperability. His prior work in **defense, financial systems, and cybersecurity** directly aligns with the challenges outlined in the SBIR topic. Below are key areas where his expertise applies:

1. Canonical Data Model (CDM) Development and Data Standardization

Jason has led multiple **data modeling and system integration** projects, including:

- **State of New Mexico (KPMG):** Spearheaded the development of a **canonical data model (CDM) for Medicaid data modeling**, influencing federal data standards.
- **United Airlines (Merchandising IT):** Developed an enterprise-wide **AI-driven data exchange framework**, enabling interoperability between multiple pricing and merchandising systems, resulting in a **30% increase in per-seat revenue**.
- **Northern Trust (Distributed Frameworks):** Designed a **distributed framework for migrating 800+ enterprise applications** into a unified architecture.

How it applies to this SBIR:
Jason's **CDM development experience** directly translates to the AI-generated DSM challenge by providing:

- A **structured methodology** for harmonizing **combat system data models**.
- A **proven approach** for integrating **multiple schema standards** (NIEM, UCore, DoDAF, S1000D, RDF, STIX).
- A deep understanding of **ontology-driven decision support systems**, critical for AI reasoning in military applications.

2. AI and Semantic Knowledge Graphs for Data Integration

Jason has significant expertise in **semantic modeling, ontology development, and AI-driven data integration**, including:

- **StratML 3.5 Development:** Led the expansion of the **Strategy Markup Language (StratML)** standard, **catching the attention of the DoD** due to its potential for improving **machine-readable command structures**.
- **MultiPlex.studio (AFWERX & DoD Engagements):** Developed **AI-driven ontology frameworks** for military and cyberwarfare applications, competing in **AFWERX Multi-Domain Operations** and **USAF Cyberspace Dominance** initiatives.
- **Stardog RDF/OWL Implementations:** Implemented **knowledge graphs** in multiple projects using **SPARQL and RDF/OWL**, allowing for **AI reasoning and dynamic query capabilities**.

How it applies to this SBIR:

- AI-generated DSM will **leverage RDF/OWL-based ontologies**, aligning with Jason's experience.
- **Semantic interoperability** between combat systems will benefit from **his expertise in linked data and federated knowledge systems**.
- AI models for **combat system reasoning and real-time decision support** can be directly built upon **his prior work with knowledge graphs**.

3. Large-Scale Military and Defense Systems Integration

Jason has **direct experience working with the DoD and USSF**, including:

- **United States Space Force (USSF):** Developed a **real-time telemetry and tracking portal** for **weather balloon instrumentation**, integrating multiple **PostgreSQL** databases with **SignalR-based live updates**.
- **Cyber Safety Harbor (Quantum Cybersecurity Initiative):** Led efforts to **rearchitect military-grade authentication systems** using **Hardware Unique Factor (HUF) and Secure Execution Environments (SEE)** for **quantum-resistant identity management**.
- **US Army SBIR (mil.API.fit):** Developed a proposal for **real-time wearables and IoT** in **Army Basic Training**, focusing on **data aggregation, AI analysis, and adaptive training models**.

How it applies to this SBIR:

- **AI-driven combat system DSM** aligns with Jason's work **standardizing DoD systems**.
- **Cyber-resilience and security-first approaches** are informed by his work in **cybersecurity and quantum-resistant architectures**.
- His **ability to rapidly prototype AI-driven decision support tools** ensures an effective **Phase I prototype** leading into Phase II.

4. Expertise in Enterprise Software Architecture for Data-Intensive Systems

Jason has **architected scalable data-driven platforms** in highly regulated industries, including:

- **Neuberger Berman (High-Yield Fixed Income):** Designed a **real-time mortgage default projection system** capable of analyzing **complex financial data models at scale**.
- **Guggenheim Partners (Enterprise Modernization):** Led the **transformation of critical financial modeling applications** into a **unified data architecture**.
- **Mesirow Financial (Currency Trading Compliance):** Designed a **compliance-driven currency trading framework** with **auditable AI-based decisioning**.

How it applies to this SBIR:

- **High-performance AI models for combat system data ingestion** align with his experience building **real-time, data-intensive enterprise systems**.
- **Ontology-based decision frameworks for interoperability** are **directly applicable** from his **financial compliance AI projects**.
- His **background in data transformation at scale** ensures efficient conversion of **combat system technical documentation** into **machine-readable AI-driven DSMs**.

5. Strategic Impact and Proposal Development for DoD Contracts

Jason has a **long history of successful technical proposal development**, including:

- **AFWERX Fusion 2019 Finalist** for **Multi-Domain Operations (MDO) Fog Computing Solutions**.
- **NSA and USD-R&E RFPs and RFIs:** Developed technical responses and architecture roadmaps for **national security applications**.
- **US Army, USAF Cyberspace Dominance, and Undersecretary of Defense for R&E engagements:** Successfully delivered **technical whitepapers, strategy documents, and AI architecture proposals**.

How it applies to this SBIR:

- His ability to **align technical solutions with DoD strategic priorities** ensures **a well-positioned AI-driven DSM proposal**.
- **Experience securing DoD funding and navigating government contracts** directly benefits **this SBIR's transition into Phase II and Phase III**.
- His expertise in **trusted AI and autonomy frameworks** aligns with **SBIR modernization priorities**.

Conclusion

Jason Lind's background in **CDM development, AI-driven knowledge systems, military system interoperability, and cybersecurity** makes him the ideal leader for this SBIR project. His **experience in developing scalable, AI-powered data frameworks for mission-critical systems** aligns directly with the **AI-generated DSM and ontology challenges**. His **track record of delivering DoD-focused AI architectures, large-scale system integrations, and ontology-based decision frameworks** ensures that this project is technically feasible, strategically aligned, and mission-ready for **Phase II deployment and beyond**.