CUI

ATTENTION

Controlled by:	OUSD(R&E)
Controlled by:	Sponsoring DoD Agency/Component SBIR/STTR Program Office
CUI Categories:	Source Selection
Distribution Statement:	F
POC:	osd.ncr.ousd-r-e.mbx.sbir-sttr@mail.mil

"Distribution Statement F. Further distribution only as directed by [Sponsoring DoD Agency/Component SBIR/STTR Program Office] [Date of Determination: BAA End Date for SBIR/STTR Phase I Proposals or SBIR/STTR PII Proposal Receipt Date (Proposal Released to the Agency/Component)] or higher DoD authority."

ATTENTION

All individuals handling this information are required to protect it from unauthorized disclosure.

Handling, storage, reproduction, and disposition of the attached document(s) must be in accordance with 32 CFR Part 2002 and applicable agency policy.

Access to and dissemination of Controlled Unclassified Information shall be allowed as necessary and permissible to any individual(s), organization(s), or grouping(s) of users, provided such access or dissemination is consistent with or in furtherance of a Lawful Government Purpose and in a manner consistent with applicable law, regulations, and Government-wide policies.

Standard Form 901 (11-18)
Prescribed by GSA/ISOO | 32 CFR 2002

CUI

Small Business Innovation Research(SBIR) Program - Proposal Cover Sheet

Disclaimer

Knowingly and willfully making any false, fictitious, or fraudulent statements or representations may be a felony under the Federal Criminal False Statement Act (18 USC Sec 1001), punishable by a fine of up to $10,000, up to five years in prison, or both.

SBIR Phase I Proposal

Proposal Number:	**F244-0001-0034**
Proposal Title:	**KnowFlow: An Interactive, Mutable, Federated Knowledge Graph Environment**

Agency Information

Agency Name:	**USAF**
Command:	**AFMC**
Topic Number:	**AF244-0001**

Firm Information

Firm Name:	**Technergetics, LLC**
Address:	**114 Genesee Street , Utica, NY 13502-3508**
Website:	**http://www.techngs.com**
UEI:	**XWYVA88GZ1F5**
DUNS:	**078840730**
CAGE:	**6X4M8**
SBA SBC Identification Number:	**000700589**

Firm Certificate

OFFEROR CERTIFIES THAT:

1. It has no more than 500 employees, including the employees of its affiliates.	**YES**
2. Number of employees including all affiliates (average for preceding 12 months)	**75**
3. The business concern meets the ownership and control requirements set forth in 13 C.F.R. Section 121.702.	**YES**
4. Verify that your firm has registered in the SBAS Company Registry at www.sbir.gov by providing the SBC Control ID# and uploading the registration confirmation PDF:	**SBC_000700589**

Supporting Documentation:

- SBC_000700589_sbir.pdf

5. It has more than 50% owned by a single Venture Capital Owned Company (VCOC), hedge fund, or private equity firm	**NO**
6. It has more than 50% owned by multiple business concerns that are VOCs, hedge funds, or private equity firms?	**NO**
7. The birth certificates, naturalization papers, or passports show that any individuals it relies upon to meet the eligibility requirements are U.S. citizens or permanent resident aliens in the United States.	**YES**
8. Is 50% or more of your firm owned or managed by a corporate entity?	**NO**
9. Is your firm affiliated as set forth in 13 CFR Section 121.103?	**NO**
10. It has met the performance benchmarks as listed by the SBA on their website as eligible to participate	**YES**
11. Firms PI, CO, or owner, a faculty member or student of an institution of higher education	**NO**
12. The offeror qualifies as a: [ ] Socially and economically disadvantaged SBC [ ] Women-owned SBC [ ] HUBZone-owned SBC [ ] Veteran-owned SBC [ ] Service Disabled Veteran-owned SBC [**X**] None Listed	
13. Race of the offeror: [ ] American Indian or Alaska Native [ ] Native Hawaiian or Other Pacific Islander [ ] Asian [**X**] White [ ] Black or African American [ ] Do not wish to Provide	
14. Ethnicity of the offeror:	**NON-HISPANIC**
15. It is a corporation that has some unpaid Federal tax liability that has been assessed, for which all judicial and administrative remedies have not been exhausted or have not lapsed, and that is not being paid in a timely manner pursuant to an agreement with the authority responsible for collecting the tax liability:	**FALSE**
16. Firm been convicted of a fraud-related crime involving SBIR and/or STTR funds or found civilly liable for a fraud-related violation involving federal funds:	**NO**
17. Firms Principal Investigator (PI) or Corporate Official (CO), or owner been convicted of a fraud-related crime involving SBIR and/or STTR funds or found civilly liable for a fraud-related violation involving federal funds:	**NO**

Signature:

Printed Name	**Signature**	**Title**	**Business Name**	**Date**
Robert Harrod	Robert Harrod	President	Technergetics, LLC	02/12/2020

Audit Information

Summary:

Has your Firm ever had a DCAA review?	**YES**
	Last Audit Date: **04/16/2024**
Was your accounting system approved by the auditing agency?	**YES**
	Last Update Date: **09/19/2019**
Was a rate agreement negotiated with the auditing agency?	**YES**
	Last Update Date: **11/08/2023**
Was an overhead and/or cost audit performed?	**YES**
	Date of Overhead Audit: **06/18/2019**
	Date of Cost Audit: **06/18/2019**
Are the rates from the audit agreement used for this firms proposal?	**YES**

Firm Information:

Agency Firm:	**Defense Contract Audit Agency**
Address:	**5795 Widewaters Parkway 2nd Floor Dewitt , New York 13214**
Point of Contact (POC) Name:	**Tong Sun**
POC Phone:	**(571) 448-9307**
POC Email:	**xiaotong.sun.civ@mail.mil**

Upload a copy of the audit information:

- 01 DCAA MFR 03441-2018U10100004-final-Signed-1.pdf

VOL I - Proposal Summary

Summary:

Proposed Base Duration (in months):	**6**

Technical Abstract:

Data has emerged as a pivotal component of modern warfare, rivaling conventional weapons in its strategic significance. The seamless flow of information is integral to the operational effectiveness of our economy, military capabilities, and command structures. As we navigate increasingly complex global military dynamics, information superiority—the capacity to collect, process, and exploit operational data—becomes crucial for outpacing adversaries. The demand for a real-time strategic knowledge architecture capable of aggregating operational data is growing, evident in the use cases of the Department of the Air Force (DAF) Intelligence, Surveillance, and Reconnaissance (ISR). This architecture would facilitate rapid decision-making across various scenarios, including combat, humanitarian operations, and peacetime engagements. By converting data into actionable intelligence, we can enhance situational awareness, empower decision-makers, and optimize power projection through Multi-Domain Operations (MDO) across cyberspace and physical domains. This transformation informs faster tactical responses and sustains operational readiness in an evolving battlefield landscape.

Knowledge Graphs (KGs) have become essential tools across industries; in commercial settings, KGs often function as a unified, authoritative source of truth, even if deployed across distributed infrastructures. However, military applications demand a different approach. Operational units, such as individual squadrons or organizations, must be able to construct, adapt, and customize their KGs to address unique engagement scenarios. This is particularly exacerbated if a unit needs to operate in a contested environment under degraded network conditions, such that even a unit's KG may diverge from itself. This decentralized need inevitably leads to the creation of multiple disconnected or partially aligned KGs, increasing the risk of information silos and fragmented knowledge.

Creating a robust solution for interactive and editable KGs that can effectively manage the complexities of dynamic, real-time decision-making scenarios while ensuring interoperability across organizational units involves various engineering challenges. Our team possesses the expertise necessary to tackle these obstacles head-on. Furthermore, Technergetics has established a solid reputation in this domain, leveraging extensive experience in large-scale data collection and KG development to enhance the effectiveness of our solution.

In this Phase 1 effort, our combined team will assess the feasibility of a software ecosystem that brings these solutions together, along with a deep alternative analysis of current SOTA methodologies in dynamic KG alignment and reinforcement learning (RL), to advance the development of **KnowFlow: An Interactive, Mutable, and Federated KG Environment (IMF-KGE),** a novel solution for intelligence analysis, enhanced situational awareness, threat detection, and targeting operations.

Anticipated Benefits/Potential Commercial Applications of the Research or Development:

We propose to research and develop KnowFlow IMF-KGE to address the DAF need for adaptive, interactive, and dynamic KGs. Our effort focuses on applying version control principles to manage dynamic mutations within the IMF-KGE. With these principles in place, we can automate data ingestion, mapping, and recommendation features while providing users with a streamlined user interface (UI) to interact with the IMF-KGE. The result should be increased ML labeling efficiency, streamlined operational analytics, and knowledge collaboration to support “fight-tonight” readiness.

1. Research and design automatic data ingestion methodologies to translate tabular data stores into knowledge graph entries, specifically targeting situational awareness use-cases.
2. Research and design a KG version control system that tracks a federated KG state with data version control systems while maintaining rich context to support downstream recommendations tasks leveraging nanopublications.
3. Research and evaluate SOTA approaches for automatic KG recommendations based on user modeling with RL as well as graph completion algorithms leveraging KG embeddings.
4. Assess the viability of Cytoscape.js (U.S. National Human Genome Research Institute (NHGRI), n.d.) as a KG visualization and interaction mechanism once integrated with a viable KG version control system.

Technergetics' strategy for the technology offers opportunities for DoD and non-DoD customers to optimize product revenue. A key differentiator between the two customer markets is that DoD customers will receive unlimited rights developed within this effort. This benefit will reduce the barrier of entry for the DoD. The market's growth is primarily propelled by organizations' need to effectively manage and derive insights from large, complex datasets. The rising demand for personalized user experiences and the imperative for interoperability and seamless data sharing among diverse systems further fuel this expansion. Knowledge graphing technologies are finding applications across multiple sectors—including healthcare, media and entertainment, financial services, e-commerce, and manufacturing—where they enhance operational efficiency, reduce costs, and deliver tailored experiences to end users. Potential customers of interactive KGs span various industries and sectors, primarily driven by their need to analyze complex data relationships, derive insights, and improve decision-making processes. We envision initial primary customers as AFRL and DAF intelligence analysts. The global KG market is experiencing significant growth, driven by the increasing need for efficient data integration and management solutions. In 2022, the market was valued at approximately USD $1.3B and is projected to reach USD $3.6B by 2030, reflecting a compound annual growth rate of 14.2%. These projections underscore the growing adoption of KGs as organizations seek to harness complex data relationships for enhanced decision-making and operational efficiency.

0

List a maximum of 8 Key Words or phrases, separated by commas, that describe the Project:

Interactive Knowledge Graphs, Knowledge Acquisition, Knowledge Graph Versioning, Data Collection, Nanopublication Architecture, Data Visualizations, Federated Data Environment, AI/ML-Enhanced Nodes and Edges

VOL I - Proposal Certification

Summary:

1. At a minimum, two thirds of the work in Phase I will be carried out by your small business as defined by 13 C.F.R Section 701-705. The numbers for this certification are derived from the budget template. To update these numbers, review and revise your budget data. If the minimum percentage of work numbers are not met, then a letter of explanation or written approval from the funding officer is required. Please note that some components will not accept any deviation from the Percentage of Work (POW) minimum requirements. Please check your component instructions regarding the POW requirements.	**YES**
Firm POW	**67.88%**
Subcontractor POW	**32.12%**
2. Is primary employment of the principal investigator with your firm as defined by 13 C.F.R Section 701-705?	**YES**
3. During the performance of the contract, the research/research and development will be performed in the United States.	**YES**
4. During the performance of the contract, the research/research and development will be performed at the offerors facilities by the offerors employees except as otherwise indicated in the technical proposal.	**YES**
5. Do you plan to use Federal facilities, laboratories, or equipment?	**NO**
6. The offeror understands and shall comply with export control regulations.	**YES**
7. There will be ITAR/EAR data in this work and/or deliverables.	**NO**
8. Has a proposal for essentially equivalent work been submitted to other US government agencies or DoD components?	**NO**
9. Has a contract been awarded for any of the proposals listed above?	**NO**
10. Firm will notify the Federal agency immediately if all or a portion of the work authorized and funded under this proposal is subsequently funded by another Federal agency.	**YES**
11. Are you submitting assertions in accordance with DFARS 252.227-7017 Identification and assertions use, release, or disclosure restriction?	**NO**
12. Are you proposing research that utilizes human/animal subjects or a recombinant DNA as described in DoDI 3216.01, 32 C.F.R. Section 219, and National Institutes of Health Guidelines for Research Involving Recombinant DNA of the solicitation:	**NO**
13. In accordance with Federal Acquisition Regulation 4.2105, at the time of proposal submission, the required certification template, "Contractor Certification Regarding Provision of Prohibited Video Surveillance and Telecommunications Services and Equipment" will be completed, signed by an authorized company official, and	**YES**

included in Volume V: Supporting Documents of this proposal.

NOTE: Failure to complete and submit the required certifications as a part of the proposal submission process may be cause for rejection of the proposal submission without evaluation.

14. Are teaming partners or subcontractors proposed?	**YES**
15. Are you proposing to use foreign nationals as defined in 22 CFR 120.16 for work under the proposed effort?	**NO**
16. What percentage of the principal investigators total time will be on the project?	**20%**
17. Is the principal investigator socially/economically disadvantaged?	**NO**
18. Does your firm allow for the release of its contact information to Economic Development Organizations?	**YES**

Partners:

Partner Name	Partner Type	Point of Contact
Simage	Small Business	Dr Chase Yakaboski

VOL I - Contact Information

Principal Investigator

Name:	**Robert Harrod**
Phone:	**(315) 271-2096**
Email:	**robert.harrod@techngs.com**
Address:	**114 Genesee Street , Utica, NY 13502 - 1111**

Corporate Official

Name:	**Robert Harrod**
Phone:	**(315) 271-2096**
Email:	**robert.harrod@techngs.com**
Address:	**114 Genesee Street , Utica, NY 13502 - 1111**

Authorized Contract Negotiator

Name:	**Ms. Evi Jesaitis**
Phone:	**(315) 271-2096**
Email:	**evi.jesaitis@techngs.com**
Address:	**114 Genesee Street , Utica, NY 13502 - 1111**

Table of Contents

1.0. Identification and Significance of the Problem or Opportunity

Data has emerged as a pivotal component of modern warfare, rivaling conventional weapons in its strategic significance. The seamless flow of information is integral to the operational effectiveness of our economy, military capabilities, and command structures. As we navigate increasingly complex global military dynamics, information superiority—defined as the capacity to collect, process, and exploit operational data—becomes crucial for outpacing adversaries. The demand for a real-time strategic knowledge architecture capable of aggregating operational data is growing, evident for Department of the Air Force (DAF) Intelligence, Surveillance, and Reconnaissance (ISR). This architecture would facilitate rapid decision-making across various scenarios, including intelligence exploitation, combat, humanitarian operations, and peacetime engagements. By converting data into actionable intelligence, we can enhance situational awareness, empower decision-makers, and optimize power projection through Multi-Domain Operations (MDO) across cyberspace and physical domains. This transformation informs faster tactical responses and sustains operational readiness in an evolving battlefield landscape.

Knowledge Graphs (KGs) have become essential tools across industries, with prominent examples including Google Knowledge Graph [1], DBpedia [2], and Wikidata [3]. In the public sector, initiatives such as the NIH Biomedical Translator program [4] aim to build a federated KG of biomedical knowledge to accelerate drug discovery, repurposing, and translational research from the lab to the clinic. These efforts demonstrate the power of KGs to consolidate and structure vast amounts of information, enabling advanced search, reasoning, and data integration. However, existing KGs in commercial and research environments typically operate on acquisition timelines of weeks to months. This cadence is inadequate for highly dynamic, time-sensitive decision-making environments, such as those encountered by military personnel, where insights are needed within seconds or minutes.

In commercial settings, KGs often function as a unified, authoritative source of truth, even if deployed across distributed infrastructures. However, military applications demand a different approach. Operational units, such as individual squadrons or organizations, must be able to construct, adapt, and customize their own KGs to address unique engagement scenarios. This is particularly exacerbated if a unit needs to operate in a contested environment under degraded network conditions, such that even a unit's KG may diverge from itself. This decentralized need inevitably leads to the creation of multiple disconnected or partially aligned KGs, increasing the risk of information silos and fragmented knowledge. Therefore, if an interoperable approach is not addressed, even the best KG solution will become immediately useless.

Developing a solution to support interactive and editable KGs that can handle the real-time complexity of a dynamic and evolving decision-making scenario while maintaining interoperability between organizational units presents several engineering challenges. The core engineering and R&D challenges include:

1. **Large-Scale KG Management and Cognitive Burden:** Current KG systems primarily focus on visualization and query mechanisms, which are essential but insufficient in highly dynamic environments. Analysts and operators often need to directly manage, track, and modify the graph's content to maintain accuracy and relevance. This manual process can become time-consuming and cognitively overwhelming, creating barriers to adoption in high-pressure environments where fast decisions are critical.

2. **Provenance Tracking of User Changes:** As KGs evolve, ensuring accountability and traceability of changes is essential for trust and usability. However, most commercial off-the-shelf (COTS) solutions lack robust provenance tracking, particularly at the user level. Developing a method to log manual edits, additions, and deletions is key to ensuring the graph's integrity and usability over time.
3. **Federated Knowledge Fusion:** In military settings, individual units are likely to create, manage, and curate their own KGs tailored to their specific operational needs. However, critical insights and better situational awareness emerge when these distributed graphs are aligned and fused into a coherent operational picture. Achieving seamless integration and knowledge fusion among these federated KGs without sacrificing their local autonomy is a complex challenge requiring innovative alignment techniques and ontology management.
4. **Real-Time Knowledge Ingestion and Event Data Integration:** In dynamic environments, the ability to ingest, process, and incorporate real-time data, such as event streams, is crucial. This involves creating robust data pipelines that convert raw, tabular, or unstructured data into actionable knowledge that can be linked and queried within the KG. Proper data labeling, ontology alignment, and event mapping are non-trivial problems that need to be solved to make real-time data useful for situational awareness.
5. **AI/ML-Enhanced Node and Link Prediction:** To reduce users' manual burden, the system must incorporate AI/ML techniques that predict missing or relevant nodes and edges within the KG. This capability will allow users to focus on critical tasks while the system offers meaningful suggestions for graph updates.
6. **Scalability and Performance in Time-Constrained Operations:** Military operations often involve time-sensitive scenarios where latency in knowledge retrieval or graph updates can impact mission success. Ensuring the system can perform efficiently at scale without compromising accuracy or functionality will be critical. This requires optimizing the underlying graph architecture and the algorithms for query and update performance.

Our team is well-positioned to address each of these challenges. Technergetics has developed and matured the FuelAI Machine Learning (ML) data labeling platform, has integrated it with Team Awareness Kit (TAK) to process tactical data streams, and has deployed FuelAI to operational DoD networks for AF ISR use cases, addressing Challenge 4. Additionally, Technergetics is developing the Army Tech Marketplace (ATM), which aims to become the "Research Intelligence Network" for Army researchers to participate in the innovation economy through dynamic connections to Government agencies, industry, and investors. ATM R&D is implementing a recommender system to proactively predict potential KG relationships, addressing Challenge 5. Furthermore, Technergetics has the expertise and experience to optimize our solution for time-constrained operations and a proven track record of delivering solutions in such contexts, directly addressing the listed Challenge 6.

Simage has developed the D-SPARK platform for the National Institute of Health (NIH), specifically addressing Challenges 1, 2, and 3. Our D-SPARK platform allows organizations to set up multiple KGs, automatically aligning and implementing a LinkML [5] data model, a human-readable and extendable open-source ontology solution, while operating in an inherently federated architecture. Additionally, organizational users can edit and manipulate KGs in real-time with full provenance tracking through our RDF nanopublication model. This enables robust tracking and efficient semantic reasoning while abstracting away the complexities of managing RDF statements and governing ontologies.

In this Phase 1 effort, our combined team will assess the feasibility of a software ecosystem that brings these solutions together, along with a deep alternative analysis of current SOTA methodologies in dynamic KG alignment and reinforcement learning (RL), to advance the development of an Interactive, Mutable, and Federated (IMF) KG solution for intelligence analysis, enhanced situational awareness, threat detection, and targeting operations.

2.0. Technical Objectives

Objective 1: Knowledge Acquisition. Can we develop knowledge ingestion pipelines that allow users to ingest actionable data into respective KGs?

- **Key Result:** Establishes a foundational KG tailored to the specific use case of an analyst or organization, reducing the manual burden of data entry and minimizing errors from hand-curated statements.
- **Measure of Success:** Demonstrate the pipeline's capability to process raw situational awareness data (e.g., sensor data, event information) and translate this data into an initial KG using our existing Situational AWareness Linked Data (SAWlink) ontology and FuelAI research.

Objective 2: Knowledge Graph Versioning. Can we demonstrate a novel approach to track user KG modifications within an independent graph database to manage changes over time?

- **Key Result:** Develop a federated KG version control platform that tracks KG state with a data version control system, with updates contextualized via a nanopublication framework for effective downstream predictive recommendations.
- **Measure of Success:** Show enhanced capabilities for version control of KG states and establish a GovCloud nanopublication server for tracking contextualized analyst updates.

Objective 3: Knowledge Refinement. Can we develop a Reinforcement Learning (RL) data collection approach to gather user changes to KGs and train a future KG recommender?

- **Key Result:** Develop a prototype data collection methodology leveraging our nanopublication architecture to capture and structure analyst-driven KG modifications, creating a foundation for training KG refinement recommendations.
- **Measure of Success:** Validate a prototype RL agent's ability to improve KG accuracy within a simulated environment using a KG generated from our acquisition pipeline (Technical Objective 1).

Objective 4: User Experience. Can we design a user interface which provides users with methods to provide input to the KG, and knowledge managers with methods to manage user input and provenance for federated knowledge graph?

- **Key Result:** Design a prototype user interface (UI) which provides effective interaction with the KG, while preserving the depth of contextual information provided by these features.

- **Measure of Success:** Show ease of use in navigating and editing a federated KG through this intuitive, user interface and its associated visualizations to potential user and knowledge manager stakeholders.

2.1. Technical Approach and Feasibility

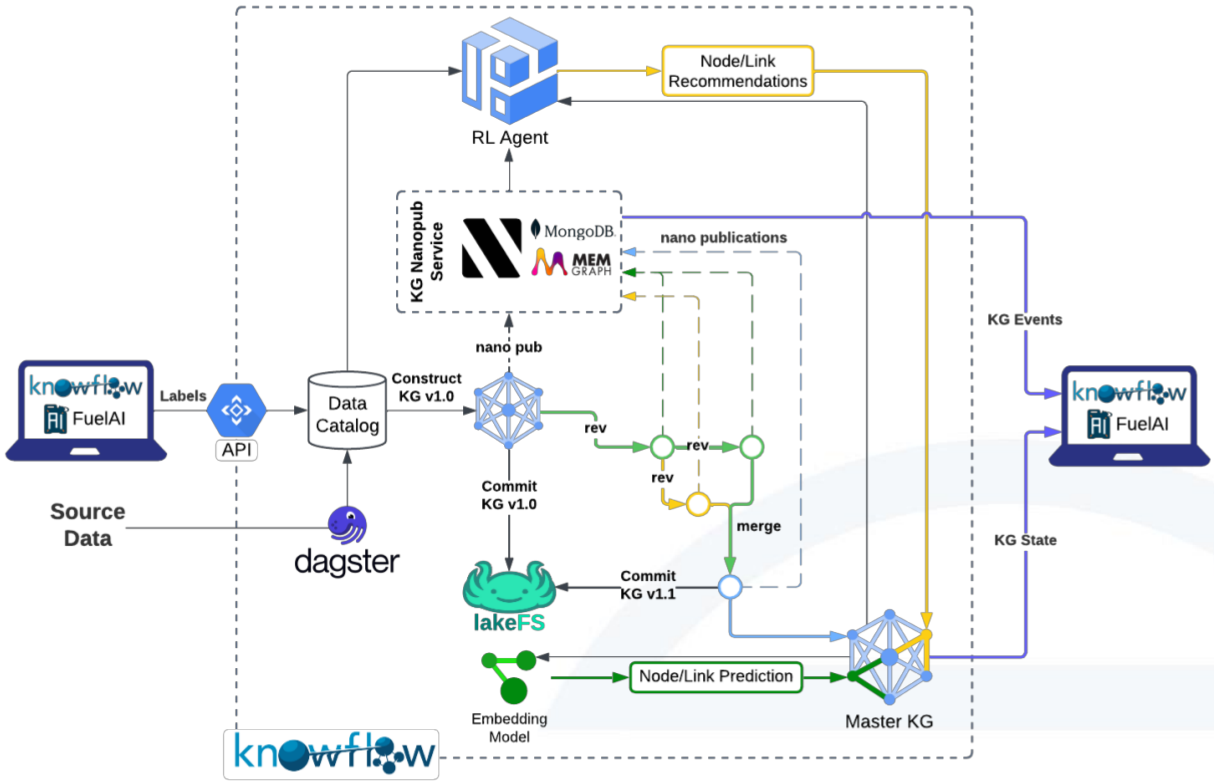

Figure 1: KG Tracking and Versioning with Nanopublications, all user edits are published as nanopubs

We propose to research and develop KnowFlow, an Interactive, Mutable, and Federated Knowledge Graph Environment (IMF-KGE), to address the DAF need for adaptive, interactive, and dynamic KGs. Our effort focuses on applying version control principles to manage dynamic mutations within the IMF-KGE. With these principles in place, we can automate data ingestion, mapping, and recommendation features, while providing users with a streamlined user interface (UI) to interact with the IMF-KGE. The result should be increased ML labeling efficiency, streamlined operational analytics, and knowledge collaboration to support "fight-tonight" readiness.

To accomplish this vision, our approach centers around research within four primary areas aligned with our technical objectives and depicted in Figure 1.

1. Research and design automatic data ingestion methodologies to translate tabular data stores into knowledge graph entries, specifically targeting situational awareness use-cases.
2. Research and design a KG version control system that tracks a federated KG state with data version control systems while maintaining rich context to support downstream recommendations tasks leveraging nanopublications.

3. Research and evaluate SOTA approaches for automatic KG recommendations based on user modeling with RL as well as graph completion algorithms leveraging KG embeddings.
4. Assess the viability of Cytoscape.js [1] for KG visualization and interaction mechanism once integrated with a viable KG version control system.

2.1.1. Scientific Feasibility

In this section, we present an initial scientific feasibility assessment of three key research areas to demonstrate that our approach is well-supported by current trends in academic literature. Since Research Area 2, which focuses on KG version control, is primarily implementation-oriented rather than research-driven, our review will focus on the scientific literature relevant to Research Areas 1, 3, and 4.

Aligning and Converting Flat Tabular Data into Knowledge Statements

One of the most common challenges organizations face with implementing KG-based solutions is that their data is not inherently organized in a graph structure [2]. Instead, data typically resides in relational databases, CSV files, or Excel sheets, all of which must be transformed into knowledge triples (subject-relationship-object) to become useful within a KG. A study by Hofer et al. (2023) [3] highlights this challenge, noting the significant, often manual, effort required for schema alignment and custom data integration. The authors point out that many organizations lack a unified schema, either across or within departments, which adds complexity to the harmonization process. Developing a schema from scratch is a cognitively demanding task, further complicating the adoption of KGs.

Moreover, in the context of the semantic web, the limitations of tabular data formats become even more apparent. Tabular data lacks intrinsic semantic content, making it difficult to automatically convert into meaningful graph structures. As a result, attempts to transform tabular data or real-time data streams into graphs without sufficient semantic context often lead to degraded graph quality [4]. These challenges, along with frequent feedback from organizations using KG solutions, indicate a clear need for robust methodologies to transform existing data assets into foundational KGs. Developing a schema or ontology mapping solution appears to be the most promising path forward, providing a structured approach to harmonize disparate data sources and streamline KG construction.

For example, Wu et al. [5] argue that without an underlying ontology that organizes nodes and edges based on background knowledge of concepts, a KG system is essentially just a data graph. While establishing an ontological foundation is a barrier to KG adoption for many organizations, once an ontology or data model is in place, the problem becomes more tractable. A rich body of research supports the conversion of raw sensor data [6] and text data [7] into semantic web triples, as well as the development of full data acquisition pipelines [2] [8] [9].

To formalize this process, we define KG construction as follows:

Key Definition 1: Knowledge Graph Construction. KG construction is a procedure, f, that maps a data source or data stream into a KG: $f: D \times f_k(D) \rightarrow G$ where D is the set of data sources and $f_k(D)$ is an ontology or data model representing the target structure. In our case, the SAWlink Model will serve as f_k and significant progress has already been made in building this ontological model for situational awareness. See Section 4.5 for a detailed introduction of our SAWlink Model.

With the KG construction problem well-defined, the function f generally takes three primary forms: (1) named entity recognition and extraction, (2) event recognition and extraction, and (3) entity linking. Each of these methods has been extensively studied in the semantic web community and has found successful applications in systems like TransOMCS [10] and ASER [11].

KG Recommendations via Embeddings, Graph Neural Networks, and Path-based Methods

Recent advances in Knowledge Graph (KG) embeddings and deep learning demonstrate the feasibility of using Graph Neural Networks (GNNs) and embedding-based methods for KG recommendations [12] [13] [14] [15] [16]. Embedding-based approaches convert KG entities and relationships into continuous vector spaces, enabling more efficient manipulation and interpretation of graph elements for prediction tasks such as entity classification, link prediction, and recommendation systems. In particular, recurrent knowledge graph embedding (RKGE) frameworks have been explored to integrate KGs with recommender systems, using pre-trained embeddings that represent KG elements in a latent vector space [17]. By incorporating these embeddings into a recommendation framework, it becomes possible to leverage relational patterns within the KG to inform recommendation processes.

Path-based methods in KG recommendation systems utilize specific sequences of relationships, or "meta-paths," to identify and recommend items based on their connections within the graph [16]. In a recommendation context, these meta-paths (e.g., user-item-user or user-item-category) enable the system to uncover patterns and relationships that link users to items in a meaningful way. By following these paths, the system can suggest items that are likely to be relevant, based on how similar entities or users are connected to one another within the KG structure.

Traditional path-based approaches often rely on manually designed meta-paths, which effectively capture domain-specific relational patterns [18]. However, these manually constructed paths are labor-intensive to develop and can lack flexibility when applied to varied or evolving datasets. To address these limitations, recent advances focus on using deep learning techniques to automatically identify useful paths. For instance, [19] demonstrate that machine learning algorithms can dynamically discover optimal paths, enhancing recommendation accuracy without the need for extensive manual intervention.

These path-based methods not only improve recommendation relevance but also provide a transparent framework that reveals why certain recommendations are made, as they are based on explicit relational pathways within the KG. By focusing on relationship patterns rather than standalone entities, path-based methods can dynamically adapt to new data in the KG, making them a valuable approach for building robust and contextually aware recommendation systems [20].

Further, recent research highlights the potential of combining KGs with Large Language Models (LLMs) to bridge structured and unstructured data [21, 22]. By aligning KG subgraph feature vectors with embeddings from unstructured data streams in latent space, it's possible to integrate new knowledge into the KG. For example, a Graph Convolutional Network (GCN) can aggregate features across connected nodes within a KG, while an LLM can extract semantic features from individual nodes. A hybrid approach could involve using the LLM for feature extraction at the

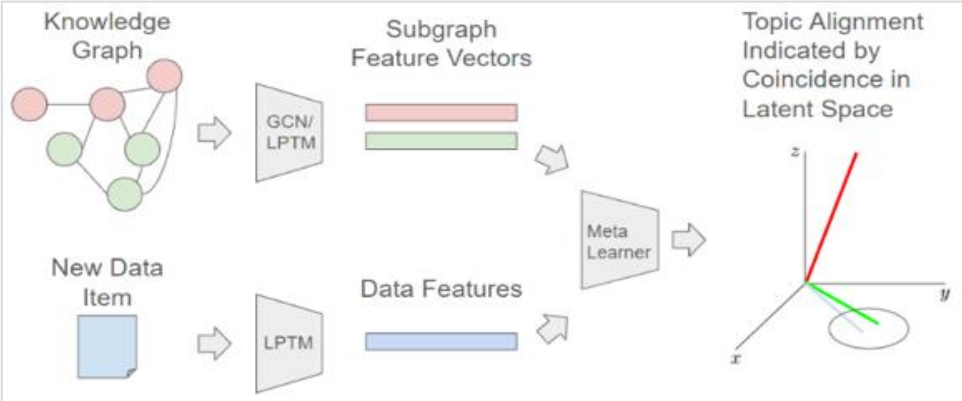

Figure 2: Algorithmic approach to aligning novel unstructured data (e.g. written reports, mission briefs, news stories, user inputs, etc...) with nodes in the pre-existing KG.

node level and the GCN to capture the broader structural context across the KG. On the data ingestion side, Language-Pretrained Models (LPTMs) like BERT, or multimodal models like CLIP, can convert unstructured data into vector embeddings that align with the KG's vector space. Although, there are significant downsides to an approach based solely on LLMs for KG construction (see Alternative Technical Approaches section).

These techniques underscore the feasibility of our approach: leveraging GNNs, KG embeddings, and deep learning to develop a recommendation system that automatically suggests KG refinements. Through these methods, we aim to achieve a dynamic KG that can evolve in response to new data, continuously enhancing situational awareness and analytic capabilities.

Knowledge Graph Recommendations via Reinforcement Learning

RL has become a cornerstone of AI research, achieving recent success in large language models (LLMs) like ChatGPT through a technique known as RL from Human Feedback. In RL, an agent interacts with an environment, exploring potential actions and receiving rewards or penalties with the goal of maximizing cumulative rewards [23].

An RL problem is typically formulated as a Markov Decision Process (MDP), represented by a four-tuple (S, A, T_a, R_a), where:

- S is the state space, representing all possible states of the environment in which the RL agent operates.
- A is the action space, where A_s denotes the actions available from a given state s.
- $T_a(s, s')$ is the transition function, either deterministic or probabilistic, defining the likelihood that action a at state s at time t will lead to state s' at time $t + 1$.
- $R_a(s, s')$ is the reward function, assigning a reward for transitioning from state s to s' as a result of action a.

The objective of RL algorithms is to find the optimal policy π, which specifies the action $\pi(s)$ that maximizes the expected cumulative reward from any state s. Solutions to this optimization problem generally fall into three categories: value-based methods, policy-based methods, and actor-critic methods.

A recent review by Huo et al. [24] examines the application of RL to four primary KG tasks: knowledge representation, extraction, fusion, and reasoning. Knowledge representation involves learning ontological structures or concepts either from raw data or from KGs with poor alignment. Knowledge extraction constructs KGs from structured (e.g., relational databases), semi-structured (e.g., tables, lists), or unstructured data (e.g., plain text). Knowledge fusion merges heterogeneous KGs or aligns different ontologies, while knowledge reasoning aims to infer new conclusions, correct errors, and identify gaps within the KG. While RL has proven successful across these areas, our research will focus on knowledge reasoning, as it will allow us to infer and suggest recommendations to analysts.

Several methods have been published that leverage RL for tasks like inferring missing edges or answering specific queries by traversing paths within the KG, thus forming predictive connections between entities [25, 26, 27, 28]. These methods generally use graph traversal and rule-based logic to infer missing relationships. While we plan to incorporate these established techniques, we hypothesize that another approach could provide additional value. Existing methods focus on static, curated KGs, but if we consider a KG's construction or editing history, a new sequential decision-making process emerges. We believe that RL, particularly in the context of recommender systems, could be applied to the KG's dynamic history through nanopublications that capture the KG's evolution over time.

Limited research exists on applying RL to the dynamic histories of KGs. However, there is substantial literature supporting RL's utility in traditional recommender system settings [29]. Mapping the recommender system problem onto an MDP framework for KG recommendations is straightforward, where the analyst and the federated KG serve as the environment, and the MDP components are as follows:

- S: A set of states defined by the analyst's preferences and the KG change log.
- A: A set of actions representing recommendations, such as nodes or edges to update (add, edit, delete).
- $T_a(s, s')$: The transition function encoding the probability of moving from s to s' if action a is taken.
- $R_a(s, s')$: The reward assigned based on analyst feedback (e.g., accept or reject) in response to a recommendation.

Our research aims to explore this RL-driven recommender system within the context of dynamic KGs, testing the feasibility of using analyst interactions and nanopublications to inform KG updates. Given RL's success in recommendation settings, we believe this approach could offer a powerful mechanism for continuous, user-guided KG refinement that adapts to new insights and evolving data.

Human Computer Interaction with Large Knowledge Graphs

Recent research in human-computer interaction (HCI) highlights several challenges associated with visualizing and working with large graphs, including KGs and other types of graph data [30]. First, many organizations, even small companies, deal with graphs containing billions of edges, creating significant scalability challenges. The sheer size of these graphs often exceeds the limitations of current software, and although advances in graph databases have eased some of these issues, setting up and managing large-scale graph solutions still requires expert knowledge.

Visualization of large graphs presents additional hurdles, as scalability issues make navigation slow and cumbersome. Graph layouts frequently need recalculation, and users are often confronted with the "hairball" effect, where the visualization becomes an opaque, tangled mess. Creating a manageable visualization requires careful filtering to reduce complexity, often necessitating proficiency in a graph query language to apply meaningful filters that align with the user's intended analysis.

Despite these challenges, several open-source KG visualization tools are available that can handle large and dynamic graphs. For instance, Cytoscape.js [1] is a widely used graph interaction and visualization tool, especially popular in large-scale life science and bioinformatics applications. It has been successfully employed to visualize vast biomedical knowledge bases, such as Reactome [31], Pathway Commons [32], and Ensembl [33]. Cytoscape.js is also used by many Fortune 500 companies, including Amazon and IBM, as well as by universities and non-profits. Other open-source tools like Gephi [34] and Graphviz [35] offer alternatives but have more limited features and are, in our view, less extensible than Cytoscape.js, as they are not directly built as JavaScript libraries.

Given the success of these tools in large-scale domains, applying them to situational awareness KGs appears feasible. The primary HCI challenge to address is abstracting away the complexity of graph query languages to enable users to create and interact with meaningful visualizations without encountering the "hairball" effect. Developing user-friendly interfaces and filtering options will be crucial to ensure the visualization aligns with users' analytical goals while maintaining usability and clarity.

2.1.2. Data-to-KG Ingestion

KGs are only as useful as the underlying data model supporting them. For KGs to provide meaningful and actionable insights, the data feeding into them must be harmonized into a coherent, and structured model. Further, if these KGs are to be dynamic, this data model must also be extendable over time. However, most organizations lack a unified data model across their curated datasets and especially their real-time data streams, creating a significant barrier to deriving actionable knowledge from a KG. Without an initial alignment or fusion over these diverse data sources, any resulting KG may lack the consistency and relevance needed for real-time situational awareness use-cases.

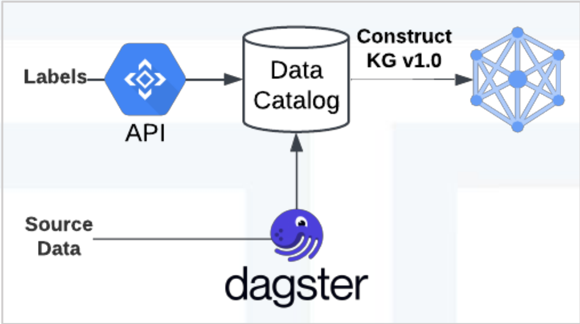

Figure 3: Data to KG Ingestion

To circumvent this challenge, our approach beings with research and design efforts focused on first exploring viable solutions to automate this translation between flat, tabular data into structured KG entries. We propose to extend Dagster as a basis for our Data-to-KG pipelines. Dagster is an open-source data orchestrator for building, scheduling, and monitoring data pipelines. It's designed to help data engineers and data scientists manage complex workflows involving data ingestion, transformation, and analytics. With Dagster, we can build (or extend) Python assets, and compose pipeline "graphs" to collect, transform, and ingest streaming data. Our approach is based upon successful implementations for the DoD; we have extended Dagster to import 700,000 entities and 1.4M relationships from open-source data sources to our ATM KG. Figure 4 is the Dagster visualization for our ATM ingestion pipeline for Federal Procurement Data System (FPDS) data, a publically available Government contract database.

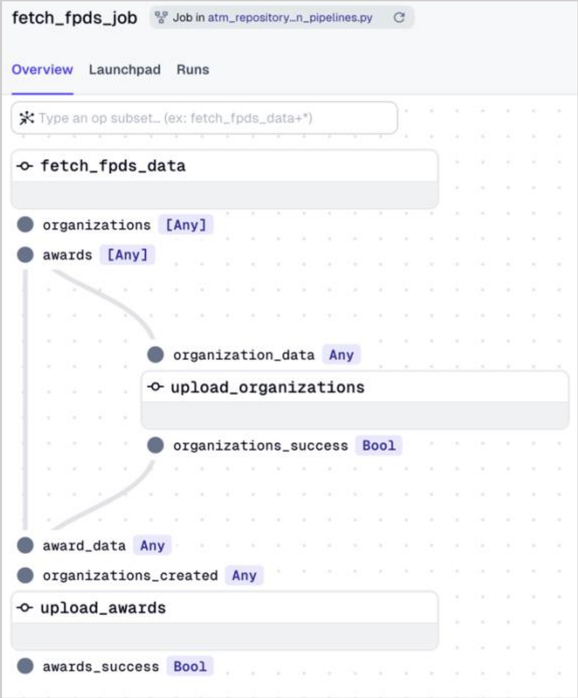

Figure 4: Dagster FPDS Ingestion Pipeline

This R&D could equip the DAF units with foundational KGs built from their existing data, minimizing the manual label of creating a KG from scratch. Such foundational KGs could then serve as an immediate operational starting point, enabling organizations to leverage data they already maintain before adding more complex real-time data sources. We have included a code snippet of the ATM FPDS Dagster ingestion function in Volume 5 of this proposal (KG-atm-dagster-code-snippet.pdf).

2.1.3. KG Provenance Versioning with Nanopublications, Context, and Snapshots

Our team has been developing an RDF nanopublication architecture to track changes in mutable, federated KGs based on recent work in FAIR nanopublications for science [36, 37, 38]. While our SAWlink model offers built-in provenance tracking making it possible to monitor changes directly within the KG itself, we opted to decouple these two frameworks for performance, scalability, and clarity of purpose. Tracking every change and provenance detail within a KG would introduce significant complexity, making the graph cumbersome and ultimately impeding query performance. This separation ensures that the SAWlink model can remain lightweight and efficient for situational awareness tasks, while the nanopublication server provides a dedicated mechanism for managing and tracking edits.

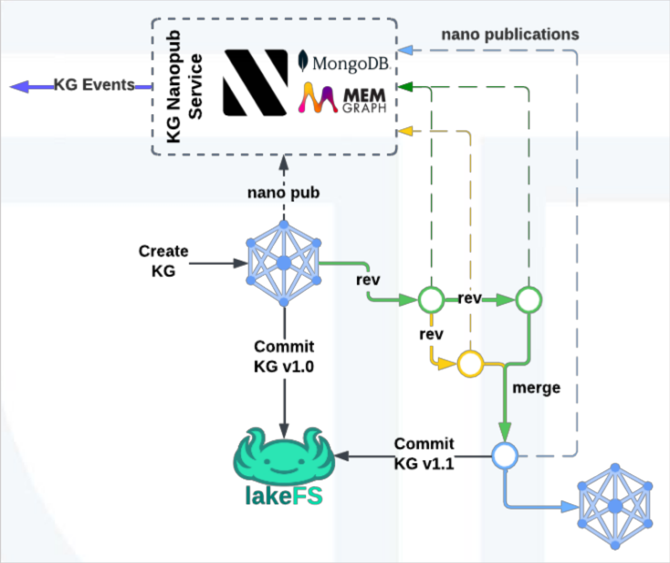

Figure 5: KG Versioning Subsystem

Our approach treats the nanopublication server as analogous to a git server, while the KGs represent the source code or files tracked by the server. However, unlike source code, knowledge graphs are highly dynamic and context-dependent, meaning that the sequence and purpose of changes carry just as much weight as the changes themselves. This is where traditional event logs or change tracking systems fall short. Solutions such as Neo4j and Memgraph alone support change logs and event tracking, but these logs often lack the semantic context necessary to explain the “why” behind a change—much like commit messages in Git provide essential insight into a codebase’s evolution.

Tracking just the raw changes without capturing the rationale, assumptions, and goals behind each modification leaves the system blind to the intent of the users, creating cognitive load for analysts when they need to interpret the edits later. In fast-moving operational environments, such as those involving threat detection or situational awareness for the DoD, this can become a critical bottleneck. Analysts would have to manually piece together why a KG was updated, leading to inefficiencies and a lack of trust in AI-augmented workflows.

The nanopublication approach directly addresses these challenges by providing self-contained, contextualized “knowledge updates.” A nanopublication records not only what change was made (e.g., adding or removing an edge or node) but also why it was made and who made it, along with any supporting evidence or metadata. Each nanopublication contains the core subject-predicate-object triple along with provenance details, offering a concise but complete representation of the edit. This modular structure supports semantic reasoning and traceability, making the KG evolution intelligible and actionable for both users and AI systems.

Furthermore, our nanopublication architecture lays the groundwork for intelligent KG refinement workflows. By aggregating nanopublications over time, we can identify patterns in user behavior and common edit paths, allowing the system to recommend future updates automatically. This feedback loop between users and the system ensures that the KG continues to improve, becoming more accurate and relevant with each interaction.

KG Snapshots to Ensure Data Integrity

Our approach emphasizes a robust, DoD-compliant solution for version controlling the KG state, capturing snapshots of nodes, edges, and properties at designated intervals. By integrating an open-source data version control system like lakeFS, we aim to create efficient, manageable “snapshots” that balance storage demands with data accuracy. These snapshots act similarly to database backups, capturing the full state of the KG during key points in its evolution while maintaining manageable size constraints.

To accomplish this, we propose a snapshotting architecture analogous to a Git workflow, with KG merge requests (KG MRs) serving as a mechanism to periodically update the KG state. KG MRs allow for multiple nanopublications to be incorporated at once, ensuring that each update is contextualized with its own nanopublication metadata. This structure offers a pragmatic trade-off, allowing us to manage the potentially large binary files that represent KG snapshots while still leveraging the modular functionality of nanopublications to maintain granularity.

Versioning Process Artifacts to Ensure Data Integrity

Within this architecture, version control is critical for maintaining data integrity across the KG’s lifecycle. By treating each KG snapshot as a distinct artifact within the version control system, we can ensure that every modification is captured with historical context. This level of versioning provides traceability, allowing analysts and operators to audit changes, revert to previous states if necessary, and confirm that the KG’s evolution aligns with mission requirements. This approach also addresses data integrity concerns, offering transparency and accountability for each change introduced into the KG.

General Data Persistence

The snapshotting strategy also addresses the broader challenge of data persistence. By establishing intervals for capturing KG state, we avoid the overhead of continuous versioning while ensuring essential information is preserved for future analysis. This approach supports long-term data retention policies by allowing specific KG versions to be archived in alignment with regulatory or mission-critical requirements. Furthermore, maintaining persistence through a structured snapshotting schedule ensures that data remains accessible and interpretable across shifts in technology and operational demands, supporting continuity in KG utilization.

2.1.4. KG Refinement Recommender System using RL and GNNs

To enable continuous and adaptive updates to the KG or federated KGs, we propose a KG refinement recommender system that will provide real-time suggestions for modifications based on two key modalities: (1) reinforcement learning (RL) applied to analyst nanopublication histories, and (2) pretrained graph embedding models that generate node and link predictions based on the current KG state. By combining these approaches, we aim to establish a feedback loop where recommendations are initially derived from graph embeddings and presented to the analyst for review. Each recommendation, whether accepted or rejected, is recorded as a nanopublication. As these nanopublications accumulate, the RL agent can learn from this feedback history to continually refine its recommendations, ensuring the KG remains relevant, accurate, and aligned with organizational needs.

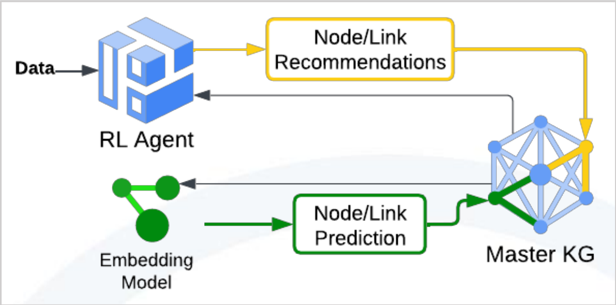

Figure 6: KG Recommendations and Prediction

In this RS, we will model the decision-making process as an MDP with the following components:

- S: A set of states $(p, h, G, e) \in S$ defined by a tuple of the analyst's preferences p, the KG change log (analyst's histories) h, the current KG state G, and the current recommendations from the GNN embedding model e.
- A: A set of actions representing recommendations. Actions include accepting or ignoring embedding model suggestions, along with additional recommendations the embedding model might miss.
- $T_a(s, s')$: The transition function encoding the probability of moving from s to s' if action a is taken.
- $R_a(s, s')$: The reward assigned based on analyst feedback (e.g., accept or reject) in response to a recommendation.

Due to the large state space defined by the tuple (p, h, G, e) we plan to use deep Q-learning as our initial RL approach, supported by evidence of its effectiveness in handling high-dimensional state spaces. Selecting appropriate embedding models will also be critical. We will explore publicly available graph embedding models on platforms like Hugging Face and fine-tune as necessary to optimize their relevance for our application.

For validation, we will set up a simulated training environment using our data-to-KG ingestion pipeline (Section 2.1.2). This will generate an initial KG from real-world, SAWlink-aligned data, representing situational awareness use cases common to intelligence analysts. We will then design reward functions to mimic different analyst behaviors, such as:

- **Maximizing Graph Completeness**: Rewarding the RL agent for completing as many relevant connections as possible.
- **Minimalist Focus**: Rewarding the agent for reducing the graph to a subgraph of critical relations most relevant to the analyst.

These workflows can be encoded within the MDP reward function, allowing us to test and refine the RL agent's adaptability across different decision-making scenarios.

Finally, we will run deep Q-learning to optimize the RL agent's policy, adjusting the embedding model as needed to achieve an optimal policy under these simulated scenarios. To assess the effectiveness of the KG refinement recommender system, we will focus on two primary metrics:

1. **Average Reward Convergence**: By observing the RL agent's reward patterns over time, we can determine how effectively it learns to optimize its actions in alignment with the ground truth KG. Convergence on a high average reward would indicate that the agent has successfully internalized a policy for generating accurate and valuable recommendations.
2. **Practical Utility of Recommendations**: Beyond reward convergence, the true value of the recommender system lies in its ability to provide actionable insights that enhance situational awareness. We will evaluate the utility of the agent's recommendations in simulated scenarios, measuring how well they align with real-world analyst needs and contribute to improved decision-making.

Through this iterative process of data collection, training, and evaluation, the KG refinement recommender system can develop a robust understanding of KG maintenance needs, enabling it to feasibly adapt to new data sources and evolving situational contexts. By capturing and structuring analyst interactions via nanopublications, we can create a feedback loop that ensures the system's recommendations remain relevant and continually improve over time. This approach positions the recommender system as an integral component for maintaining the KG's integrity and operational value across diverse scenarios.

A key consideration when building an operational recommender system will be mitigating the bias of the recommendations. As one of the primary data sources in an RS is human interactions, both explicit and implicit, biases are bound to arise. Much work has been done to understand sources of bias in recommender models [39]. Potential sources of bias include popularity bias, position bias, selection bias, and temporal bias. For example, popularity bias occurs when a handful of items receive the most user interaction, resulting in these items being recommended more frequently than other, less popular items. Training methods such as introducing regularization terms have been shown to alleviate popularity bias in modern RS [39]. Cumulatively, these biases can lead to low-quality recommendations, unfair promotion of certain items, and an unsatisfactory user experience. By considering these potential factors, we can achieve both more performant models and more fair and equitable models that can prevent unwanted consequences in our system, such as intelligence product/source preference.

2.1.5. Knowledge Graph Visualization

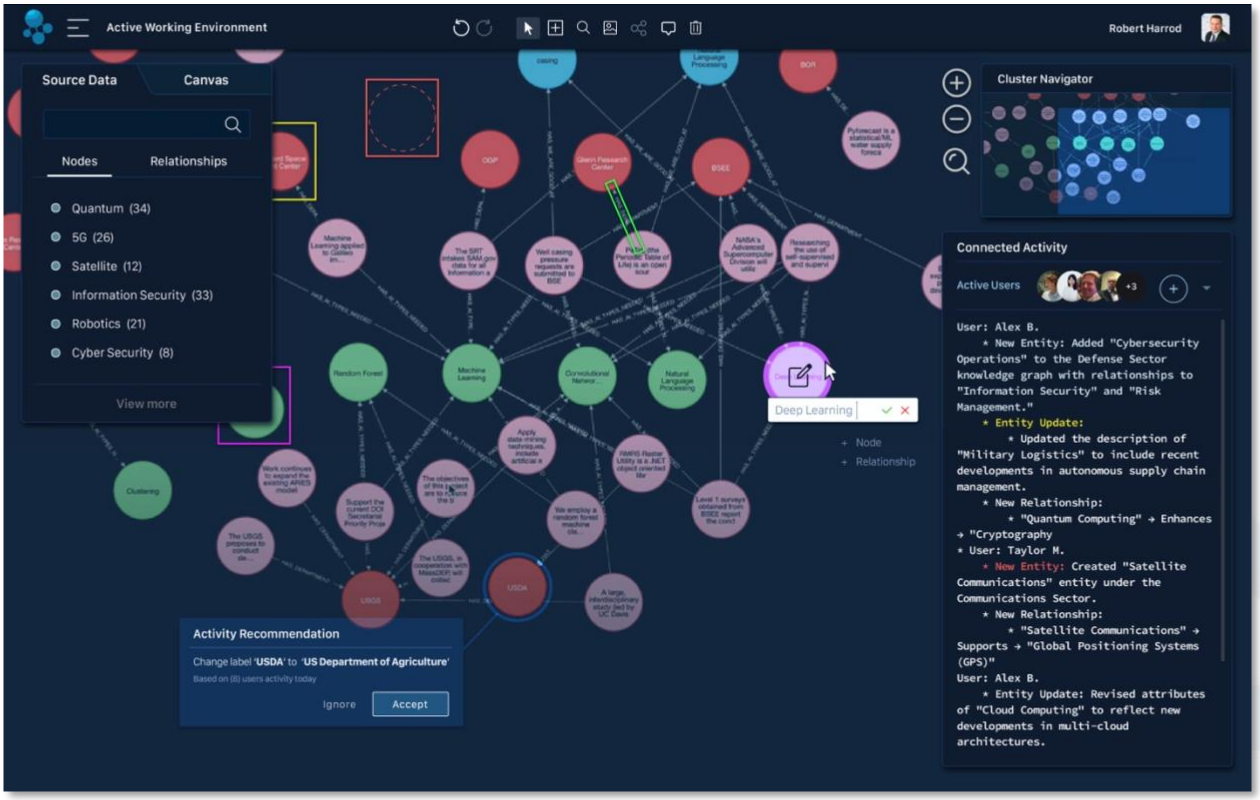

Figure 7: Interactive Knowledge Graph Visualization Concept

Designing knowledge graph visualization requires a balance of clarity, interactivity, and scalability. Nodes and edges should be distinct, with colors, shapes, or sizes representing different entities and relationships, while keeping labels minimal to avoid clutter. Scalability can be achieved with techniques like progressive loading, clustering, and filters, allowing users to explore complex graphs in layers. Interactivity—such as zoom, pan, tooltips, and highlighting paths—lets users engage dynamically, and node selection helps focus on specific details without overwhelming the display.

For larger graphs, layout is essential to maintain readability, positioning nodes logically to reflect importance or connection strength. Aesthetic choices, like minimalism, dark mode, and a consistent color scheme, enhance user experience by reducing visual strain, while responsive design and accessibility ensure usability across devices and for all users. Clear navigation aids like breadcrumbs or mini maps help users orient within complex graphs, while context tools like lineage indicators can add transparency. In essence, a well-designed knowledge graph visualization organizes complex relationships intuitively, making large datasets both informative and user-friendly. In addition, it will be necessary to provide effective visualization of user updates that have been drafted to the federated KG.

During our past performance on similar tasks, we assessed multiple design systems for specific user interaction requirements. As a result of these assessments which considered the costs/benefits in producing effective UIs for similar requirements, we recommend the Tailwind UI design system. Tailwind UI is a collection of professionally designed, responsive UI components built with Tailwind CSS. It offers pre-built components like buttons, forms, and navigation elements that are fully customizable using Tailwind's utility classes, enabling rapid development with a consistent, polished design. Tailwind UI's flexibility allows developers to adapt each component to match their unique style, making it a great choice for web applications that need quick, visually appealing, and scalable interfaces. Additionally, Tailwind CSS's utility-first approach keeps code maintainable, responsive, and lightweight, ideal for building modern, mobile-friendly applications.

In addition to Tailwind UI, we recommend analyzing Cytoscape [1] to visualize our proposed IMF-KGE. Originally developed for biological research, Cytoscape is open-source, and is widely used for visualizing molecular interaction networks, biological pathways, and genetic data. It has since evolved to support a wide range of network analysis applications, including social networks, transportation networks, and more. Cytoscape.js, Tailwind UI, and the React Javascript library are foundational open-source frameworks which can accelerate the development of a prototype capability for DAF. Figure 7 presents a concept that can be achieved with our approach presented.

2.2. Alternative Technical Approaches

KG Construction from Raw Text via Commercial Large Language Models (LLMs)

While commercial LLMs like GPT-4 and ChatGPT show remarkable capability in natural language processing tasks, recent research highlights substantial limitations when directly applying them to KG construction, particularly without a well-defined ontological framework.

Zhu et al. [22] GPT-4 and ChatGPT underperformed compared to benchmark non-LLM methods on core KG tasks such as entity, relation, and event extraction. Similarly, Gao et al. [40] LLMs struggled with event extraction, trailing behind task-specific models that use sequence-to-sequence and BERT-based methods. These findings illustrate a key issue: while LLMs are adept at generating coherent, contextually appropriate text, they often lack the precision and structure required for effective KG construction from raw text.

A primary drawback to using commercial LLMs for KG construction lies in their broad, generalized training. Commercial LLMs are designed for versatility across many language tasks rather than optimized performance in highly specialized applications. As a result, they may introduce ambiguities or inaccuracies when tasked with identifying and structuring complex relationships in domain-specific data. This is particularly concerning in fields like situational awareness, where KGs must be precise and reliable.

Furthermore, the absence of a robust ontological framework exacerbates these issues. LLMs can produce inconsistent or contextually irrelevant entity and relationship mappings without a guiding ontology. An ontology like SAWlink provides the necessary structure, ensuring that data is organized and interpreted within a consistent framework that aligns with the operational goals of situational awareness. Without this framework, reliance on LLMs for KG construction could result in fragmented or contradictory information that lacks coherence and utility.

In contrast, traditional task-specific models that leverage established ontologies offer better control over entity and relationship definition, ensuring the resulting KG is accurate and contextually relevant. Although commercial LLMs may contribute to specific tasks within the KG pipeline, such as enhancing searchability or generating human-readable summaries, their direct use for KG construction remains problematic without foundational ontological support. Our approach, therefore, focuses on integrating these LLM capabilities selectively and coupling them with SAWlink's robust ontology to ensure the KG retains both structural integrity and operational relevance.

2.3. Technical Risk Mitigation

Phase I Risk	Phase I Mitigation
Will a suitable Knowledge Graph (KG) be available for Phase I R&D?	Technergetics and Simage are actively developing the SAWlink model for AFRL, which is desirable for the Phase I effort. This allows the research team to begin R&D on day 1, rather than relying on the Government to provide a relevant KG.
Will there be enough relevant data?	More relevant data is always desirable for operational success. However, our research will be successful based upon available datasets and simulations as described in this proposal.
How can we mitigate Recommender Bias?	To address recommender bias, we will implement recent research techniques, such as feedback-loop correction and diversity-promoting algorithms. These methods will ensure that user feedback does not disproportionately influence future recommendations. Additionally, we'll introduce a periodic auditing process to identify and correct any emerging biases, ensuring that recommendations remain fair, representative, and aligned with diverse user needs. [41].
Will KnowFlow be compatible with DoD networks?	Technergetics will develop DoD-ready software based on experience deploying multiple software applications to DoD Platform One, DAF CLOUDWorks, BESPIN, and AF Cloud One.

3.0. Phase I Statement of Work Outline

3.1. Scope

The scope of this effort is to establish the KnowFlow software proof of concept and will assess the feasibility of different approaches to allow for user modifications to a dynamic Knowledge Graph (KG). Research will be conducted into novel methods to predict additional changes to the graph based on the user's input. This effort will also research version control principles to manage dynamic KG updates by users. Additionally, we will design a streamlined user interface (UI) to interact with the dynamic KG environment applying best-practice HCI principles.

3.2. Task Outlines

3.2.1. Milestone 1: Kickoff Presentation to Government Stakeholders

Within 30 days of the contract award, we will provide the Government with a project kickoff presentation that outlines the project plan and introduces the project team. Feedback from stakeholders will be collected and R&D action items will be documented.

3.2.2. Milestone 2: Demonstrate Data Ingestion Pipeline

KGs are only as useful as the underlying data model supporting them. For KGs to provide meaningful and actionable insights, the data feeding into them must be harmonized into a coherent and structured model. This task will extend our data pipeline architecture and demonstrate a prototype for automated translation between flat, tabular data into structured KG entries.

3.2.3. Milestone 3: Demonstrate KG Nanopublications and Version Control

This task will research a proof-of-concept architecture which can track and manage changes to mutable and federated KGs, akin to source code version control systems (e.g. Git). Our recommended nanopublication approach provides a loosely coupled provenance trail of contextualized knowledge updates to support downstream recommendation tasks. KG versioning provides traceability, allowing analysts and operators to audit changes, revert to previous states if necessary, and confirm that the KG's evolution aligns with mission requirements.

3.2.4. Milestone 4: Present Knowledge Graph Recommender R&D

This milestone will focus on researching and designing a KG refinement recommender system that provides adaptive suggestions for KG modifications. We will explore simulated scenarios within a RL framework, incorporating pretrained open-source graph embedding models to generate initial node and link predictions. Our goal is to validate the architecture in a controlled environment, laying the groundwork for potential future applications in real-world contexts. Evaluation will focus on two primary metrics, as outlined in our technical approach, to measure the system's effectiveness across two simulated user behavior models: maximizing graph completeness and focusing on critical relations.

3.2.5. Milestone 5: Design and Present User Interface Prototype

Designing interactive knowledge graph visualization(s) requires a balance of clarity, interactivity, and scalability. This task will focus on designing dynamic knowledge graph visualization(s) that intuitively organize complex relationships, provide insights into federated KG changes happening in the environment, and make large datasets informative and user-friendly. This design will feature user interface components that make interactions with the KG version control system straightforward.

3.2.6. Milestone 6: Final Phase I Technical Report

We will provide a final Phase I demonstration summarizing the R&D performed in these task areas. We will also deliver a final technical report on the progress made during this effort. This task focuses on preparing and presenting the final demonstration, as well as the compilation of the final technical report. Final artifacts to be delivered may include software mock-ups, software roadmaps, presentations, and a prioritized task list of potential features for Phase II applications.

3.3. Milestone Schedule and Deliverables

Table 1: Milestones for KnowFlow: IMF-KGE Phase I

Milestone Schedule	Timeframe	Objective	Deliverables
Milestone 1 – Develop and Deliver Kickoff Presentation with Government Stakeholders	Award + 30 Days	Objective All	Progress Report
Milestone 2 – Demonstrate Data Ingestion Pipeline	Award + 60 Days	Objective #1	Progress Report Demonstration
Milestone 3 – Demonstrate KG Nanopublications and Version Control	Award + 90 Days	Objective #2	Progress Report Demonstration
Milestone 4 – Present Knowledge Graph Recommender R&D	Award + 150 Days	Objective #3	Progress Report Demonstration
Milestone 5 – Design and Present User Interface Prototype	Award + 120 Days	Objective #4	Progress Report Demonstration
Milestone 6 – Final Phase I Technical Report	Award + 180 Days	Objective All	Final Report

3.3.1. Progress Reports

Progress reports are due monthly at a minimum. They shall be concise documents describing progress in meeting end-user needs. Each status report shall be no longer than 3 pages and/or 10 slides. The status reports shall include progress toward Phase I objectives and key results.

3.3.2. Final Report with SF 298

The Phase I final report is due upon completion of the technical effort. It shall be submitted via email to the cognizant AF Contracting Officer. The first page of the final report will be a single-page project summary, identifying the work's purpose, providing a brief description of the effort accomplished, and listing potential result applications. Phase I report will also describe the technology and anticipated applications/benefits for Government and/or private sector use. Final Report will include SF 298, "Report Documentation Page," with the first sentence reading, "Report developed under SBIR contract for topic "Department of Defense (DoD) SBIR 2024.4, Topic Number AF244-0001". The abstract will identify the purpose of the work and briefly describe the work conducted, the findings or results, and the potential applications of the effort as outlined by applicable guidance.

4.0. Related Work

Technergetics possesses extensive expertise in developing Machine Learning (ML) software solutions, and data analysis solutions which leverage Knowledge Graphs to interrelate data in a way that makes it easily accessible, interpretable, and usable. We regularly navigate DoD software accreditation requirements and have deployed multiple software applications to DoD networks.

4.1. Improved Data Collection and Knowledge Graphing in the TAK Ecosystem

Collecting and organizing sensor data from tactical mobile devices is crucial for creating accurate and context-aware machine-learning models that can run efficiently and securely on the devices.

This data provides the foundation for training and fine-tuning models tailored to specific missions and their unique environments. However, this data only becomes actionable when harmonized across devices, users, and organizations. To achieve this, we developed the SAWlink Model, which offers TAK stakeholders a unified source of truth for defining labels and context within a flexible, easily extendable ontology. This model enables advanced knowledge graph reasoning over previously flat and disorganized feature labels, enhancing TAK decision-making capabilities.

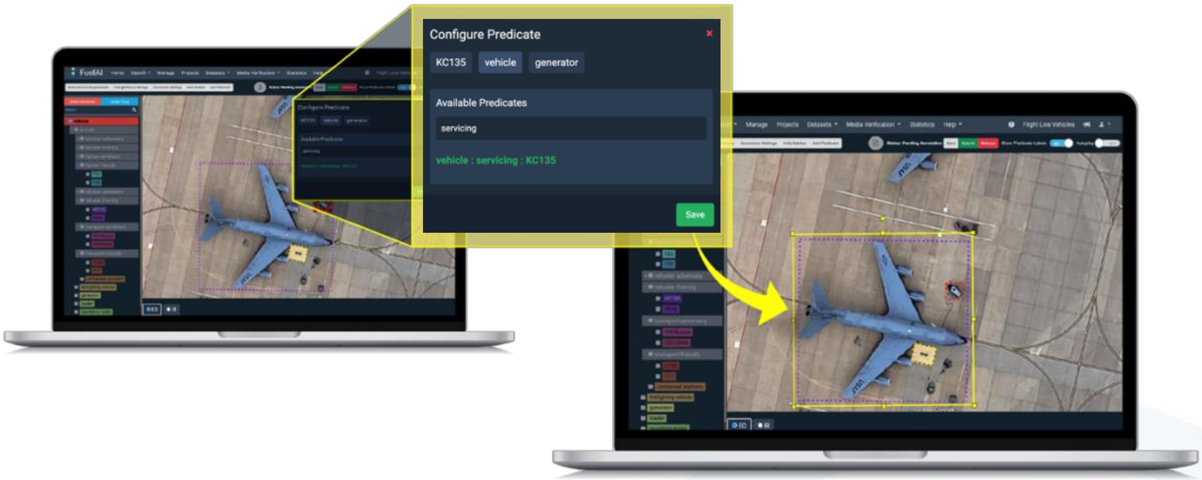

Figure 8: FuelAI Supports Predicate Labeling with SAWlink Ontologies

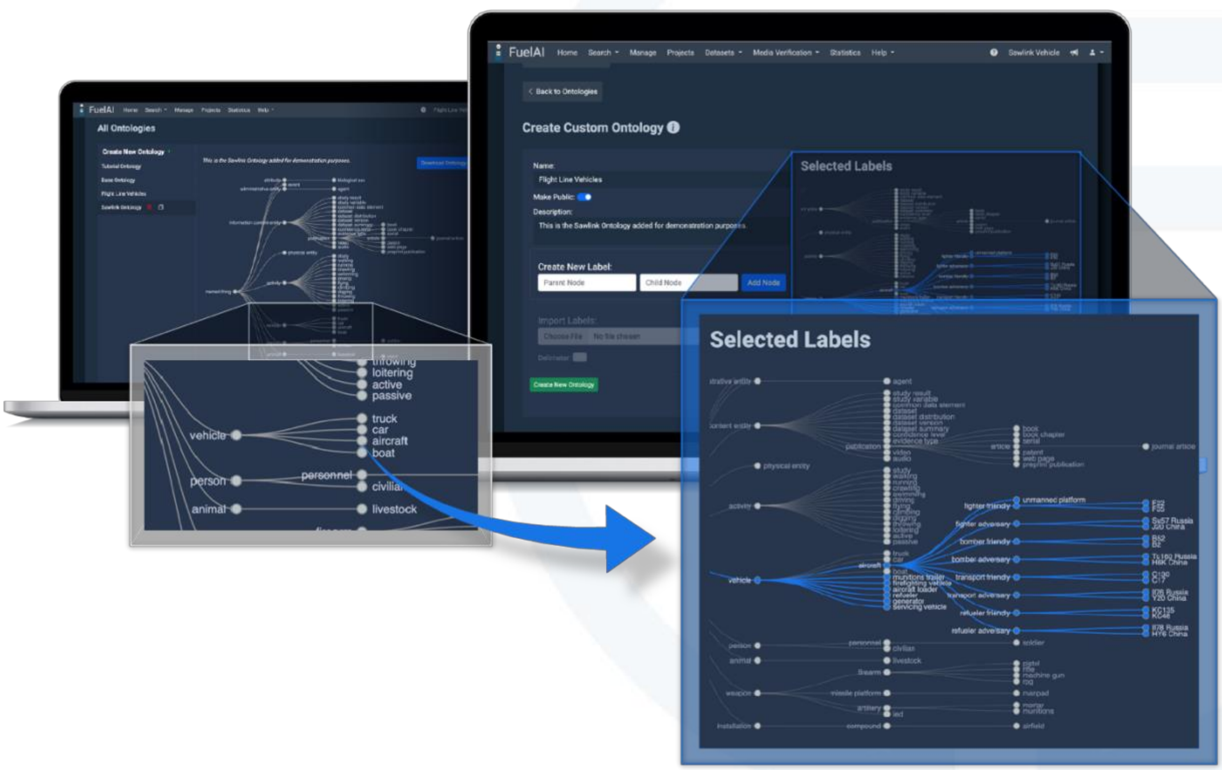

Figure 9:FuelAI Custom Ontologies Extend SAWlink for TAK

We have realized impressive capabilities in this effort with FuelAI and SAWlink, which relate to the requirements in this solicitation. SAWlink is the new upper ontology in FuelAI geared toward situational awareness use cases. A new ontology collaboration system provides an interactive way to tailor the ontology to a specific labeling effort while maintaining links to the primary SAWlink ontology as a single source of truth across the platform, aiding in data discovery (see Figure 9). SAWlink is digested in FuelAI by interpreting the YAML file containing FuelAI configuration elements, including label definitions/descriptions, prefix mappings, predicates, etc. FuelAI now uses SAWlink predicates for annotating relationships between labels, adding contextual depth to our training data (see Figure 8). This allows for a dynamic rendering of the ontology and flexibility in handling future/other versions of the ontology.

4.2. Army Tech Marketplace (ATM)

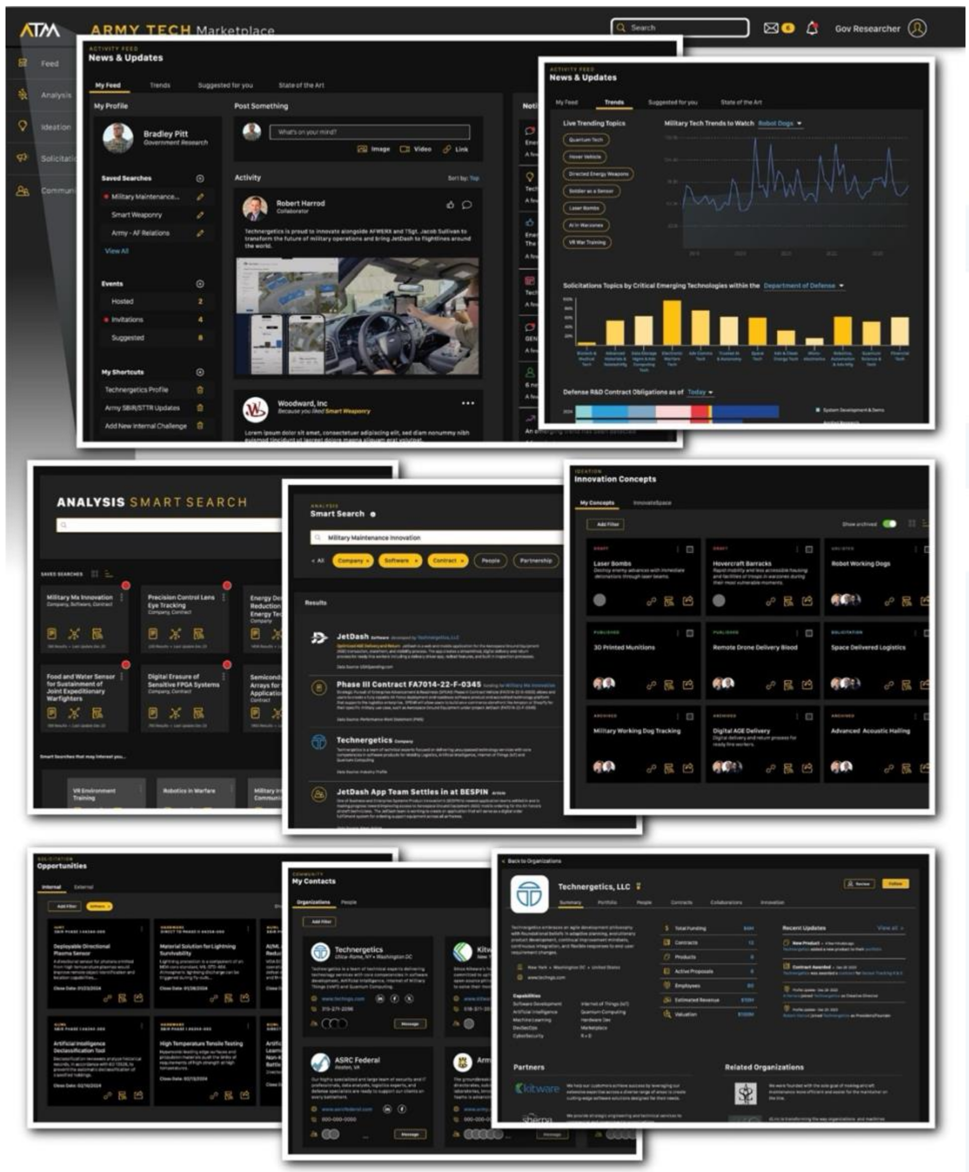

Figure 10: Army Tech Marketplace Web Application

Technergetics is the prime contractor for the ATM Phase II SBIR. ATM is an Innovation Research Intelligence (RI) Platform that offers a unique and engaging experience for Army researchers and innovators. ATM's goals are to 1. connect the Army with the innovation economy and 2: facilitate a greater awareness of innovation (e.g., structured knowledge). This R&D focuses on developing the ATM KG to allow the system to describe relationships between innovation entities from various input sources. ATM has transformed 700,000 entities and 1.4M relationships from open-source data sources to our ATM KG.

4.3. FuelAI: Data Labeling Platform for Training AI/ML Algorithms

Video: [FuelAI Introduction on YouTube]

FuelAI is our operational web-based data-annotation tool, built to DoD requirements, and does not entail software licensing costs. Since 2020, FuelAI has been used to create training datasets for DAF AI model development, for mission-critical use for Automated Target Recognition (ATR) for military drones/Unmanned Ariel Vehicles (UAV) and Battle Damage Assessment (BDA). FuelAI features DoD-procured enhancements to efficiently train computer vision (CV), natural language processing (NLP), and audio models; over 6M data labels are available and are increasing daily. FuelAI is now the preferred data-annotation tool for the DAF Chief Data and AI Office (CDAO) to build its inventory of AI data. FuelAI has a Certificate to Field (CtF) and Authority to Operate (ATO) and is actively used for DoD missions on the Joint Worldwide Intelligence Communication System (JWICS) and SIPRNET. FuelAI has contributed to countless ML model training projects over the last five years.

4.4. WingmanAI: AI Algorithm Marketplace, Evaluation, and Deployment Platform

Video: [WingmanAI: Introduction on YouTube]

State-of-the-art (SOTA) moves quickly for AI models, as evident in the recent progress achieved for Large Pre-Trained Models (LPTMs). Traditional DoD acquisition cannot respond quickly enough to test, validate, and onboard SOTA models for operational use cases. WingmanAI, a license-free web-based AI model marketplace, like HuggingFace. Another application based upon our Modellum framework, WingmanAI, is a web application that addresses this deficiency where model developers can introduce vulnerability-mitigated models to the DoD market in the security domain of their choice. In parallel, operational users define use cases where AI may benefit the platform. From there, rapid collaboration occurs between model developers and operational users to validate models on representative operational data for the defined operational use cases. Ultimately, WingmanAI streamlines the rapid acquisition of AI models while remaining accountable to the operational requirements that justify the acquisition, providing transparency for DoD stakeholders. WingmanAI is available on JWICS for AI mission support in operational ISR missions.

4.5. The SAWlink Model

The SAWlink Model is a data model designed to organize events, labels, and assets (e.g., videos, images) to support KG-based Situational Awareness (SAW) applications. This model acts as a bridge, integrating disparate ontological domains and heterogeneous data sources under a unified structure. SAWlink builds upon many of the foundational concepts first developed for Biolink, adapting them to meet the needs of the situational awareness domain. The Biolink Model [42] is a linked data framework also leverages the Linked Data Modeling Language (LinkML) [43] to harmonize biomedical knowledge.

The SAWlink Model is organized into an upper ontological hierarchy composed of defined classes, properties, predicates, and associations (see Table 2 for examples and descriptions of these elements). Knowledge aligned with the SAWlink model is represented as a set of associations, analogous to Web Ontology Language (OWL) [44] axioms and Resource Description Framework (RDF) statements. Each association captures a predicate relationship (a SAWlink model predicate) between a subject and object (SAWlink model classes), forming a core triple. An example of such an association is shown in Figure 11.

The subject and object of each association represent instances of concepts defined explicitly in the SAWlink model, such as events, vehicles, or people. Their Internationalized Resource Identifiers (IRIs) are sourced from widely adopted ontologies (e.g., WikiData, DBpedia, Sapient), ensuring that the model remains both extensible and interoperable with community standards. Additionally, associations include slots for capturing supplementary metadata, such as provenance or evidence supporting the assertion, further enhancing the model's utility for tracking and validating information.

Table 2: SAWlink Model elements, definitions, and examples

SAWlink Model Element	Definition	Example
Class	High-level categories representing core concepts of interest to situational awareness such as events, people, places, things, etc., arranged in a class hierarchy.	sawlink:Event, sawlink:Vehicle, sawlink:Weapon
Predicate	Objects that define the action being carried out by the subject on the object of the core triple and define how these two classes are related to each other.	sawlink:related_to, sawlink:affects, sawlink:contains
Node Property	A set of attributes that can be regarded as a characteristic or inherent part of an instance of a Named Thing (e.g., sawlink:NamedThing)	sawlink:name, sawlink:latitude, sawlink:longitude
Edge Property	A set of attributes that can be regarded as a characteristic or inherent part of a statement, association, or edge.	sawlink:qualifier, sawlink:timepoint, sawlink:has_evidence
Core Triple	The domain knowledge of an association expressed by the subject and object nodes and the predicate connecting them together.	sawlink:Video sawlink:contains sawlink:Event
Association	Classes that define a relationship between two domain concepts, constrained and qualified by edge attributes.	sawlink:VideoToEvent, sawlink:AudioToEvent

Type	A kind of value that tells what operations can be performed such as strings or integers, but can be extended to custom types like quotient or unit.	URI, CURIE, string, integer, sawlink:Quotient, sawlink:Unit

The SAWlink model addresses several key challenges that currently hinder, or may eventually obstruct, the interoperability of KG solutions designed for situational awareness:

1. The need for expertise to transform data between tabular, RDF, and graphical models.
2. Inconsistent application of ontologies or taxonomies at the organizational unit level, along with variations in the identifiers used to store node instances within a KG.
3. The lack of a standard approach to model the intersection of ontological domains and data labeling tasks, such as relationships between information assets (e.g., videos, images) and the events these assets contain or share.

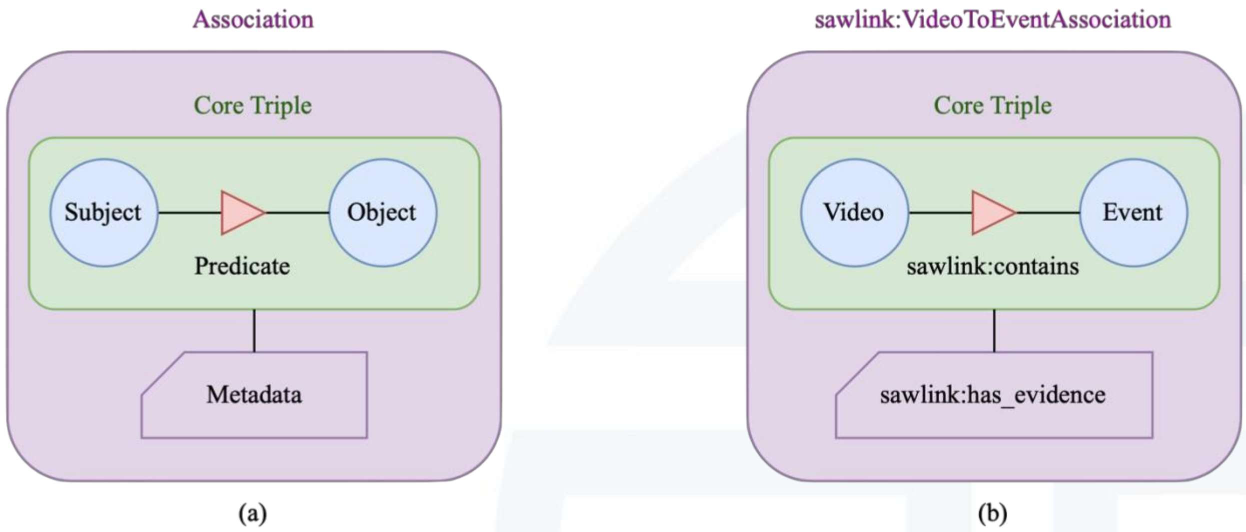

Figure 11: (a) Shows the Association architecture of a SAWlink association. First, a subject class and object class are related to one another via a predicate class forming the "core triple" or statement in the KG. This is further wrapped into an association class with the inclusion of metadata about the core triple such as publication denoting evidence of the assertion or other types of provenance. (b) An instantiated example of a situational awareness use-case in which a Video asset contains a certain type of event.

A key advantage of the SAWlink model is that it builds upon the LinkML framework, allowing it to be distributed in a variety of formats, including YAML, JSON-Schema, Python/Java Data Classes, and RDL. This flexibility makes the model accessible to a broad range of developers and database engineers working in government, academia, and private industry. Additionally, the model can be visually represented through Unified Modeling Language (UML) diagrams as seen integrated with FuelAI in Figure 9. By leveraging LinkML, model elements in SAWlink can be validated using existing toolchains such as JSONSchema validation and SQL constraints, accelerating alignment between tabular data, ontologies, and KGs.

One common issue in ontology development, particularly within large organizations or even individual units—is the creation of multiple, overlapping ontologies for the same domain. For example, the SAPIENT specification [45], developed by the UK to standardize AI and autonomy in networked multisensory systems, defines a controlled taxonomy with significant overlap with other situational awareness ontologies [46, 47, 48]. This creates a dilemma for KG developers: Which ontological concepts should I use? The problem becomes even more challenging when dealing with lower-level ontologies, such as those defining vehicle models or personnel unit callsigns. Developers must determine which vocabulary is most appropriate while also ensuring that concepts from the chosen ontology are applied correctly to class instances.

Defining a SAWlink Model class involves directly specifying the class's id_prefixes construct. This design encourages the reuse of existing ontologies while remaining flexible enough to incorporate new ones on the fly by simply adding an additional prefix. In this way, SAWlink functions as an upper-level ontology, governing the overarching, shared, and abstract concepts within the situational awareness domain that are unlikely to change over time (e.g., sawlink:Event, sawlink:Truck). At the same time, it leaves more specialized, lower-level cocepts to be managed by domain-specific ontologies (e.g., dbr:Ford_F_series from the DBpedia ontology). For example, for a city class, the SAWlink model currently recommends using dbr (the DBpedia Resource ontology) as the preferred vocabulary for city instances.

Additionally, each element in the SAWlink model is mapped, where applicable, to equivalent elements from other ontologies or models. These mappings leverage terms from the Simple Knowledge Organization System (SKOS) namespace to describe relationships between SAWlink elements and external concepts, indicating whether the relationship is exact, broad, narrow, close, or related. These mappings enhance computability and interoperability, enabling the automatic harmonization and integration of disparate data sources, which makes the model and its derivative KGs more connected and functional.

4.6. D-SPARK: A Versatile Platform for Knowledge Discovery and Situational Awareness

Originally developed as GenNIFER at Dartmouth College by Dr. Chase Yakaboski, the Discovery-Support Platform for Accelerating Research and Knowledge (D-SPARK) began as a Gene Regulatory Network (GRN) analysis tool in collaboration with the NIH NCATS Translator Program (NIH award OT2TR003436). Over time, the platform was expanded to become a flexible tool for broader data science applications.

D-SPARK provides customizable discovery-to-knowledge pipelines and built upon a federated knowledge graph hosting architecture, allowing multiple KGs to be integrated and managed sim to support knowledge visualization, and analysis. It also supports the design and implementation of LinkML-based data models, like SAWlink, to provide structured, domain-specific knowledge representations that align with a user's ontological needs. The platform also features a user-friendly interface with Cytoscape-based visualizations and an iterative search API, enabling users to explore data-driven insights and overlay analytical results with KG-informed knowledge.

In partnership with Technergetics, D-SPARK has been adapted for the TAK data collection effort, leveraging its federated architecture and flexible data modeling capabilities to integrate TAK data streams within customized situational awareness KGs. This evolution from NIH-inspired origins to situational awareness demonstrates D-SPARK's scalability and relevance for dynamic, large-scale data environments, supporting complex data alignment and exploration across interconnected KGs.

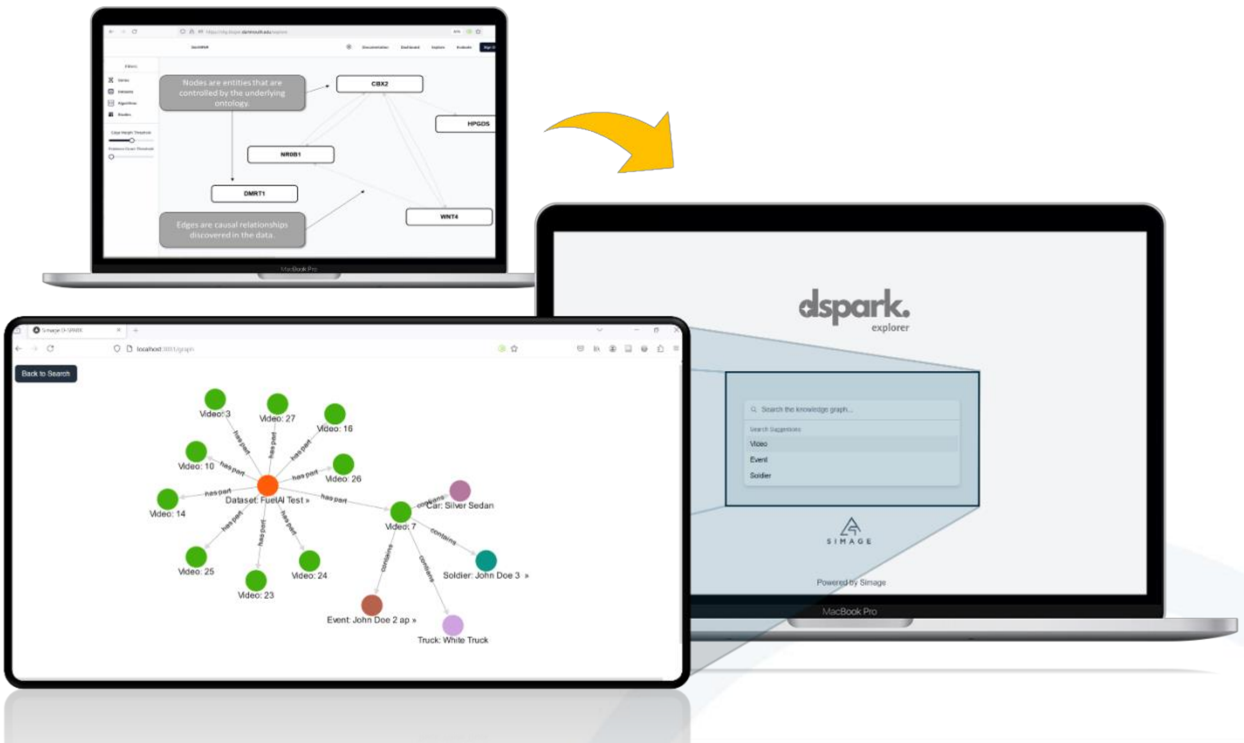

Figure 12: Evolution from NIH inspired GenNIFER to conduct GRN analysis to DSPARK a more general framework to translate data into actionable knowledge. Shown is integration with the SAWlink ontology over a sample situational awareness use case.

5.0. Relationship with Future Research or R&D

This R&D relates directly to our potential Army Tech Marketplace (ATM) Phase II-B FY26 R&D effort; however, the potential for future R&D extending this innovation is nearly endless. Organizations face significant challenges in managing, analyzing, and deriving actionable insights from vast, complex datasets as data generation expands across diverse sources- social media, IoT devices, and digital platforms. Semantic knowledge graphs offer a robust solution by organizing and structuring data to facilitate advanced search, discovery, and insight extraction. Their structured, interconnected data representation is highly compatible with AI and machine learning advancements, serving as valuable input for model training and improving AI-driven capabilities. The rapid increase in internet traffic further accelerates the demand for semantic knowledge graphing. As more people and devices connect online, data volume grows exponentially, often in unstructured formats that hinder traditional data management and analysis. Semantic knowledge graphs address these challenges by organizing data into graph structures representing relationships between disparate data points, enabling more efficient data analysis, insight generation, and informed decision-making.

This explosion of online data also heightens the need for personalization and targeted recommendations. Semantic knowledge graphs create a unified, multidimensional view of individual users by integrating data from various sources—browsing history, social media interactions, and purchase patterns—enabling precise recommendations, personalized marketing, and tailored user experiences. Moreover, the growing adoption of IoT further emphasizes the need for semantic knowledge graphs. As many devices generate real-time data streams, semantic knowledge graphs facilitate integrating and analyzing these IoT data sources, yielding more profound insights and enhancing decision-making capabilities across connected system

6.0. Commercialization Strategy

Technergetics' strategy for the technology offers opportunities for DoD and non-DoD customers to optimize product revenue. A key differentiator between the two customer markets is that DoD customers will receive data rights developed within this effort. This benefit will reduce the barrier of entry for the DoD. In the commercial sector, semantic knowledge graphing technologies are finding applications across multiple sectors—including healthcare, media and entertainment, financial services, e-commerce, and manufacturing—where they enhance operational efficiency, reduce costs, and deliver tailored experiences to end users. Our strategy for commercializing this applied R&D involves pursuing transition paths for military and commercial applications; our commercialization strategy addresses the following questions:

(1) **What is the first product planned to incorporate the proposed technology or R&D?** FuelAI, the Army Tech Marketplace (ATM), and Team Awareness Kit (TAK) are the first products planned to incorporate the proposed technology.

(2) **Who are the probable customers, and what is the estimated market size for this technology?**
We envision the Intelligence Community (IC) becoming our early adopters. The global commercial knowledge graph market is experiencing significant growth, driven by the increasing need for efficient data integration and management solutions. In 2022, the market was valued at approximately USD $1.3B and is projected to reach USD $3.6B by 2030, reflecting a compound annual growth rate (CAGR) of 14.2% [49]. These projections underscore the growing adoption of knowledge graphs across various industries, including healthcare, finance, and e-commerce, as organizations seek to harness complex data relationships for enhanced decision-making and operational efficiency.

(3) **How much money is needed to bring this technology to market and how will it be raised?** This product is anticipated to be ready for a customer introduction campaign by the end of the Phase I feasibility study and follow-on Phase II R&D effort. This campaign will engage DoD customers, and potentially customer customers, for applications of this platform. Funds will be raised through venture capital investment if required. This product is anticipated to experience customer market sustainability when funding totals $3-4M.

(4) **Does your firm have the necessary marketing expertise, and if not, how will your firm compensate?**
Technergetics was established in 2013 and boasts a workforce of over 70 employees. Over the past five years, Technergetics has garnered an impressive $39.9M in contract awards and has successfully commercialized SBIR technology. These efforts include two Phase III contract vehicles with a $198M contract ceiling open until 2027, accruing over $10.4M in allocated Task Orders. Moreover, Technergetics has two AI-based applications that have been daily fixtures in operational DoD environments, generating recurring annual income for the company for the last five years. With a solid financial track record spanning 11 years,

Technergetics is a trusted DoD software partner.

Technergetics' success as a small business requires senior-level staff to perform multiple job functions, including product development, advocacy, and business development. Our company is structured as an innovative software/hardware company committed to positively transforming our customers operations with technology.

Our team's diverse experience in the commercial product markets, leading large DoD contracts, and relationship building will prove helpful in jumping-starting the marketing strategy. Technergetics is now fortunate to have Mr. Kevin Brown leading our business development and customer engagement corporate capability. Mr. Brown is a 25-year retired DAF CMSgt Veteran who held strategic logistics leadership positions in all 9 Major Commands (MAJCOMs) and served as Air Transportation career field MAJCOM functional lead twice. He has experience in military policy at the DAF and DoD levels and is experienced in translating change management and innovation adoption across multiple career fields. He is adept at translating strategic goals into achievable team tasks and milestones and has verified experience developing over 18+ innovation software/hardware solutions for both DoD and commercial markets. Ms. Alisa Ferrara is our Creative Director and oversees product design and company-wide marketing initiatives. She is an experienced visual and concept designer with 15+ years of experience across various industries. Her background is heavily in graphic design, video production, and all things marketing. Technergetics' founder, Mr. Robert Harrod, has a diverse background in the defense and commercial sectors. Mr. Harrod has led the development of multiple successful hardware/software platforms, focusing on workflow standardization, tactical situational awareness and sensors, mobile device platform architectures, artificial intelligence/machine learning tools, and edge applications. Mr. Harrod is a subject matter expert in various programming languages, as well as has extensive knowledge in Platform as a Service (PaaS), Infrastructure as a Service (IaaS), and Infrastructure as code (IaC). Mr. Harrod also has qualifications in the transition processes required to meet requirements for the Risk Management Framework (RMF) and to achieve a CtF and ATO. Many of the solutions engineered by Mr. Harrod and his development teams have enabled analytics, visualization, and advancements in human-machine decision-making to aid in measuring mission effectiveness and future mission planning. Mr. Harrod's technical and management experience spans nearly two decades of successful client delivery of software/hardware platforms and cutting-edge research and development, supporting many highly important research programs across the DoD, DAF, and AFRL.

(5) **Who are probable competitors, and what price/quality advantage is anticipated by your firm?**

Stardog Union, and their Enterprise Knowledge Graph Platform (Stardog) is likely the most formidable small business competitor in this space. Stardog is built upon a proprietary technology stack on top of the proprietary DataBricks platform, as opposed to our approach, which leverages open-source, SBIR-funded, and Government-off-the-Shelf (GOTS) solutions. Stardog is an evolution of Extract Transform Load (ETL) and is sufficient for answering basic business analytics questions. Stardog does not provide functions to predict new relationships or nodes in a KG.

KnowFlow takes a different approach, focusing on organizational collaboration on knowledge with a completely new paradigm. Our paradigm creates federated awareness of knowledge evolution across a community between interested parties. Along with our innovative approach, we anticipate competitive advantages through cost savings, flexibility, and faster innovation. Since our codebase is customizable, we can adapt it to specific needs without being locked into a single vendor. The collaborative community behind open-source drives rapid updates and quick bug fixes, while our focus on DoD customers encourages a stronger cyber security posture. This approach attracts skilled developers and offers powerful tools without heavy investment, leveling the playing field in tech innovation. Additionally, Government customers do not incur any licensing costs to sustain the software organically.

In addition, Technergetics boasts 18 years of direct engagement with the DoD Intelligence Community and the AFRL. Our team also includes in-house intelligence subject matter experts with over 20 years of experience as data analysts within the DAF, ensuring that our solutions are well-informed and contextually relevant for defense operations.

We plan to pursue a follow-on Phase II effort with AFRL or AFWERX to develop further capabilities. The following transition paths may unlock the next opportunity for this project:

- <u>Follow-on Phase II or D2P2:</u> Continue to collaborate with the initial DAF agency, AFRL or AFWERX
- <u>DoD Commercial Solutions Offering (CSO)</u>: Pursue CSO Phase II with non-DAF customers, such as the following or other Program Offices with an interest in intelligence analysis:
    - Navy -- Naval Supply Systems Command (NAVSUP), Office of Naval Research (ONR)
    - Army--Program Executive Office Command, Control, Communications-Tactical (PEO C3T), PEO Enterprise Information Systems (PEO EIS), PEO–Intelligence, Electronic Warfare & Sensors (PEO IEW&S), Project Manager, Intelligence Systems & Analytics (PM IS&A), Project Manager, Electronic Warfare and Cyber (PM EW&C), and Project Director for Sensors Aerial Intelligence (PD SAI)
    - Marines -- Marine Corps Systems Command (MCSC)
- <u>TAK Product Center</u>: Partner with the TAK product center to define and fund transition opportunities with civilian use cases
- <u>SBIR Partnerships</u>: Develop SBIR partnerships for Federal Government transition efforts
- <u>Follow-on Task Order on one of Technergetics existing DoD Phase III Contract Vehicles</u>: Extend R&D with Technergetics Phase III contracts for Global Logistics Resources (AMC contract FA7014-22-D-0006) or DoD Enterprise E-Commerce Exchange/Transaction Platform (HAF/A4L contract FA7014-22-D-0005) for similar requirements
- <u>New Phase III</u>: Partner with Government customer to extend R&D to a new related Phase III

7.0. Key Personnel

	**Mr. Robert Harrod -** ***<u>Principal Investigator</u>*** (US Citizen, TS/SCI cleared) Mr. Harrod founded Technergetics, LLC in 2013, a small business focused on maximum impact for its clients. As the President and Chief Architect of Technergetics, Mr. Harrod has over 24 years of experience creating products for the defense and commercial technology sectors. Mr. Harrod received a Bachelor’s degree in Computer Science in 2001 from the Rochester Institute	

	of Technology. He has architected solutions for the DoD since 2006 and has led strategic efforts for ACC/A5/2D, AFRL/RI, Army ASA/ALT, and other DoD customers. Mr. Harrod is the Principal Investigator for several SBIR Phase I, Phase II, and Direct-to-Phase II efforts over the last decade. He is published for his research in AI and Computer Vision **[50]**.	
	**Gregory Hyde** - ***<u>Knowledge Graph ML Researcher</u>*** (US Citizen) Greg Hyde is a PhD candidate in Computer Engineering at the Thayer School of Engineering at Dartmouth College. His research focuses on uncovering hidden abstractions within decision-making processes, particularly within Reinforcement Learning (RL) and Inverse Reinforcement Learning (IRL) frameworks. He has significant experience designing and analyzing knowledge graphs as a member of the NIH Translator Program and designing RL agents to model user behavior and intent for the Office of Naval Research. Before pursuing his PhD at Dartmouth, he earned his Master's and Bachelor's degrees in Computer Science from the University of Wisconsin-Whitewater. His research has been published in ML and Knowledge Graphs [51, 52].	
	**Dr. Chase Yakaboski** - ***<u>Knowledge Graph Research Scientist</u>*** (US Citizen) Dr. Yakaboski is an AI/ML PhD research scientist with nearly a decade of experience in AI/ML solutions across diverse domains. He is a Fullbright Scholar and current Harvard Research Fellow. Dr Yakaboski is also the co-principal investigator on our Improved Data Collection and Knowledge Graphing in the TAK Ecosystem SBIR effort. He recently led the architecture, development, and implementation of GenNIFER for the NIH NCATS Translator program, connecting single-cell RNA-sequencing data to existing biological knowledge extracted from literature and organized into a federated knowledge base. Before pursuing his PhD, he was a lead developer for DAF SEEK EAGLE, where he developed the current version of the Common Advanced Safe Escape Software (CASES) tool integrated into the Joint Mission Planning Software (JMPS). CASES currently supports the DAF through aircraft-specific Mission Planning Environments (MPEs) and is utilized by over 20 FMS partner countries. He has published and briefed his AI research in Knowledge Graphs as well as probabilistic graphical models at major international AI/ML conferences. [53, 54, 55]	
	**Ms. Alisa Ferrara** - ***<u>Principal Product Designer</u>*** (US Citizen) Ms. Ferrara is an experienced visual and concept designer with 20+ years of experience across various industries. Her background is heavily in graphic design, video, and marketing. She is the Technergetics creative director for all customer and company-wide design initiatives. She is also the Product Manager for our Army Tech Marketplace knowledge platform.	

	**Ms. Caitlin Genna - *Full-Stack Developer*** (US Citizen, Secret cleared) Ms. Genna is a skilled full-stack developer in an Agile team supporting the DoD operational ML software solutions. She has a Bachelor's Degree in Computer Science. She has been published for her research on developing and optimizing data mining models involved in identifying and analyzing audio attributes [56].	

8.0. Foreign Citizens

Technergetics does not plan to employ any non-US citizens and/or dual citizens for any Phase I efforts.

9.0. Facilities/Equipment

Technergetics primarily intends to perform the R&D for this effort on-site at the company facilities. Our facilities meet all existing Environment, Safety, and Occupational Health (ESOH) standards required by applicable environmental laws and regulations of Federal, state, and local Governments. Our physical offices provide software and hardware engineering resources for optimal team collaboration. Connected to the internet with an enhanced fiber network, our facilities enable on-site and secure remote teamwork maintained for Cybersecurity Maturity Model Certification (CMMC) compliance. Our engineers have the proper resources to perform R&D and agile software development. Technergetics employs seventy-eight people; 65% of our staff hold active DoD security clearances ranging from TS/SCI to Secret.

Our methodology enables a repeatable and logical approach to distributed teams to implement the Software Development Lifecycle (SDLC). Technergetics specifies all software development processes in a Software Development Plan (SDP); all staff are trained semi-annually on coding standards, Integrated Development Environment (IDE) tooling requirements, secure coding requirements/reviews, and adherence to Government Open-Source Software (OSS) requirements. In addition, we maintain guides and require semi-annual reviews of recommended DevSecOps processes and corporate cloud-based Continuous Integration/Continuous Delivery (CI/CD) tooling to support these processes, which applies automated compile-time Software Composition Analysis (SCA), among other best practices.

Our methodology enables the planning, creation, implementation, validation, and tracking of user stories in a streamlined manner. Our corporate, GovCloud-based GitLab environment stores source code/artifacts and provides managed CI/CD pipelines for source code quality. Trufflehog (secret detector), SonarQube (code smells/80%+ unit test coverage), Fortify (static code analysis), Docker File verification (docker-lint), Twistlock and Anchore (container scanning), Open Web Application Security Project (OWASP) DC (dependency checking), OWASP ZAP (penetration testing), Cypress (end-to-end testing) and other tooling are used to continuously assess code.

10.0. Subcontractors/Consultants

10.1. Simage Autonomy LLC (Simage)

Simage is a Dartmouth startup company founded by Chase Yakaboski, a PhD candidate at Dartmouth's Thayer School of Engineering. We specialize in developing cutting-edge tools that accelerate and enhance scientific discovery through explainable and trustworthy AI solutions. Our core expertise is crafting solutions for knowledge graph construction, deployment, and efficient reasoning, complemented by causal inference techniques to unveil data-driven motifs, primarily in biology and engineering. Our central mission is to enhance the understanding and discovery of intricate data-derived and knowledge-supported relationships in data to uncover novel and actionable findings. As a New Hampshire-based small business, we are also deeply committed to advancing innovation and promoting sustainable growth.

11.0. Prior, Current, or Pending Support of Similar Proposals/Awards

A proposal for similar or equivalent work has not been submitted to other DoD or government agencies

12.0. References

[1] U.S. National Human Genome Research Institute (NHGRI), "Cytoscape Graph library for visualisation and analysis," [Online]. Available: https://js.cytoscape.org/.

[2] J. W. Q. L. H. P. a. X. W. L. Zhong, "A Comprehensive Survey on Automatic Knowledge Graph Construction," *ACM Comput. Surv.*, vol. 56, no. 62, pp. 1-94, 2023.

[3] D. O. A. S. H. K. a. E. R. M. Hofer, "Construction of Knowledge Graphs: Current State and Challenges," *Information*, vol. 15, no. 8, 2024.

[4] N. Abdelmageed, "Towards Transforming Tabular Datasets into Knowledge Graphs," in *The Semantic Web: ESWC 2020 Satellite Events*, 2020.

[5] M. W. a. X. Wu, "On big wisdom," *Knowledge and Information Systems*, vol. 58, no. 1, pp. 1-8, 2019.

[6] N. Konstantinou, "Converting Raw Sensor Data to Semantic Web Triples: A Survey of Implementation Options," *International Journal of Sensors Wireless Communications and Control*, vol. 2, no. 1, pp. 44-52, 2012.

[7] M. R. W. P. I. J. a. J. B. K. Hassanzadeh, "“T2R: System for Converting Textual Documents into RDF Triples," in *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, 2013.

[8] S. P. E. C. P. M. a. P. S. Y. S. Ji, "A Survey on Knowledge Graphs: Representation, Acquisition, and Applications," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 2, pp. 494-514, 2022.

[9] P. C. a. H. Paulheim, "Knowledge graph refinement: A survey of approaches and evaluation methods," *Semant. web*, vol. 8, no. 3, pp. 489-508, 2017.

[10] D. K. Y. S. a. D. R. H. Zhang, "TransOMCS: from linguistic graphs to commonsense knowledge," in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, 2021.

[11] H. Z. e. al., "ASER: Towards large-scale commonsense knowledge acquisition via higher-order selectional preference over eventualities," *Artificial Intelligence*, vol. 309, p. 103740, 2022.

[12] B. G. I. H. a. E. K. S. Liu, "INDIGO: GNN-Based Inductive Knowledge Graph Completion Using Pair-Wise Encoding," in *Advances in Neural Information Processing Systems*, 2021.

[13] A. F. S. B. D. M. T. D. N. a. E. D. S. A. C. M. Mancino, "KGTORe: Tailored Recommendations through Knowledge-aware GNN Models," in *Proceedings of the 17th ACM Conference on Recommender Systems*, New York, NY,, 2023.

[14] C. G. e. al., "A Survey of Graph Neural Networks for Recommender Systems: Challenges, Methods, and Directions," *ACM Trans. Recomm. Syst*, vol. 1, no. 1, 2023.

[15] Z. G. J. L. a. G. L. Q. Chen, "Knowledge-enhanced Multi-View Graph Neural Networks for Session-based Recommendation," in *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2023.

[16] Q. G. e. al., "A Survey on Knowledge Graph-Based Recommender Systems," *IEEE Transactions on Knowledge and Data Engineering,* vol. 34, no. 8, p. 3549–3568, 2022.

[17] Z. Sun, J. Yang, J. Zhang, A. Bozzon, L.-K. Huang and C. Xu, "Recurrent Knowledge Graph Embedding for Effective Recommendation," in *Proceedings of the 12th ACM Conference on Recommender Systems*, New York, NY, 2018.

[18] B. H. W. X. Z. a. P. S. Y. C. Shi, "Heterogeneous Information Network Embedding for Recommendation," *IEEE Transactions on Knowledge and Data Engineering,* vol. 31, no. 2, pp. 357-370, 2019.

[19] M. Z. X. X. W. L. a. M. G. H. Wang, "Knowledge Graph Convolutional Networks for Recommender Systems," in *The World Wide Web Conference*, 2019.

[20] C. T. a. A. Krouska, "Path-Based Recommender System for Learning Activities Using Knowledge Graphs," *Information,* vol. 14, no. 1, 2023.

[21] L. L. Y. W. C. C. J. W. a. X. W. S. Pan, "Unifying Large Language Models and Knowledge Graphs: A Roadmap," *IEEE Transactions on Knowledge and Data Engineering,* vol. 36, no. 7, p. 3580–3599, 2024.

[22] Y. W. X. C. J. e. a. Zhu, "LLMs for knowledge graph construction and reasoning: recent capabilities and future opportunities," *World Wide Web,* vol. 27, no. 58, August 2024.

[23] R. S. S. a. A. G. Barto, Reinforcement learning: An introduction., MIT press, 2018.

[24] Q. H. e. al., "Knowledge Graph Based on Reinforcement Learning: A Survey and New Perspectives," *IEEE Access,* 2024.

[25] R. D. e. al., "Go for a Walk and Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning," in *International Conference on Learning Representations*, 2018.

[26] H. Z. a. H. M. P. P. Tiwari, "DAPath: Distance-aware knowledge graph reasoning based on deep reinforcement learning," *Neural Networks,* vol. 135, pp. 1-12, 2021.

[27] R. L. a. X. Cheng, "DIVINE: A Generative Adversarial Imitation Learning Framework for Knowledge Graph Reasoning," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*.

[28] T. H. a. W. Y. W. W. Xiong, "DeepPath: A Reinforcement Learning Method for Knowledge Graph Reasoning," in *arXiv*, 2018.

[29] T. C. a. B. F. M. M. Afsar, "Reinforcement Learning based Recommender Systems: A Survey," *ACM Comput. Surv.,* vol. 55, no. 7, pp. 1-145, 2022.

[30] A. M. S. S. J. L. a. M. T. Ö. S. Sahu, "The ubiquity of large graphs and surprising challenges of graph processing: extended survey," *The VLDB Journal,* vol. 29, no. 2, p. 595–618, 2020.

[31] M. M. e. al., "The Reactome Pathway Knowledgebase 2024," *Nucleic Acids Research,* vol. 52, no. D1, 2023.

[32] I. R. e. al., "Pathway Commons 2019 Update: integration, analysis and exploration of pathway data," *Nucleic Acids Research,* vol. 48, no. D1, 2019.

[33] P. W. H. e. al., "Ensembl 2024," *Nucleic Acids Research,* vol. 52, no. D1, 2023.

[34] S. H. a. M. J. “. [1] M. Bastian, "Gephi: an open source software for exploring and manipulating networks," in *Proceedings of the international AAAI conference on web and social media*, 2009.

[35] E. R. G. a. S. C. North, "An open graph visualization system and its applications to software engineering," *Software: Practice and Experience*, vol. 30, no. 11, p. 1203–1233, 2000.

[36] A. G. a. J. V. P. Groth, "The anatomy of a nanopublication," *Inf. Serv. Use*, vol. 30, no. 1-2, pp. 51-56, 2010.

[37] B. M. e. al., "The value of data," *Nat. Genet.*, vol. 43, no. 4, p. 281–283, 2011.

[38] P. E. B. M. L. N. a. M. K. T. Kuhn, "Broadening the Scope of Nanopublications," in *The Semantic Web: Semantics and Big Data*, 2013.

[39] J. e. a. Chen, "Bias and debias in recommender system: A survey and future directions.," ACM Transactions on Information Systems 41.3, 29 Dec 2021. [Online]. Available: https://arxiv.org/pdf/2010.03240.pdf.

[40] H. Z. C. Y. a. R. X. J. Gao, *Exploring the Feasibility of ChatGPT for Event Extraction*, arXiv, 2023.

[41] M. Färber, M. Coutinho and S. Yuan, "Biases in scholarly recommender systems: impact, prevalence, and mitigation," *Scientometrics 128*, p. 2703–2736, 2023.

[42] D. R. e. a. Unni, "Biolink Model: A universal schema for knowledge graphs in clinical, biomedical, and translational science," *Clinical and translational science*, vol. 15.8, pp. 1848-1855.

[43] "LinkML," [Online]. Available: https://github.com/linkml/linkml.

[44] C. e. a. Bock, "OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax (Second Edition)," W2C, 11 12 2012. [Online]. Available: https://www.w3.org/TR/owl2-syntax. [Accessed 05 11 2024].

[45] "SAPIENT Network of autonomous sensors and effectors –, Interface control document BSI Flex 335," British Institute of Standards (BSI), 31 Mar. 2024. [Online]. [Accessed 29 Oct. 2024].

[46] W. G. S. M. W. R. a. W. S. N. Baumgartner, "BeAware!—Situation awareness, the ontology-driven way," *Data Knowl. Eng.*, vol. 69, no. 11, pp. 1181-1193, 2010.

[47] P. D. C. a. G. G. J. P. A. Almeida, "Towards an Ontology of Scenes and Situations," in *IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, 2018.

[48] C. J. M. a. K. B. M. M. Kokar, "Ontology-based situation awareness," *Information Fusion*, vol. 10, no. 1, pp. 83-98, 2009.

[49] "Semantic Knowledge Graphing Market Size, Share & Trends Analysis Report By Data Source, By Knowledge Graph Type, By Task Type, By Application, By Organization Size, By Industry Vertical, By Region, And Segment Forecasts, 2023 - 2030," Grand View Research, 2023. [Online]. Available: https://www.grandviewresearch.com/industry-analysis/semantic-knowledge-graphing-market-report. [Accessed 4 11 2024].

[50] A. Fitwi, Y. Chen, S. Han and R. Harrod, "Estimating Interpersonal Distance and Crowd Density with a Single-Edge Camera," 29 09 2021. [Online]. Available: https://www.mdpi.com/2073-431X/10/11/143.

[51] G. Hyde and E. Santos Jr, "Detecting Hidden Triggers: Mapping Non-Markov Reward Functions to Markov," 20 01 2024. [Online]. Available: https://arxiv.org/abs/2401.11325.

[52] E. e. a. Santos, "Validation of Double Transition Model by Analyzing Reward Distributions," in *IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT).* , 2020.

[53] C. Yakaboski, G. Hyde, C. Nyanhongo and E. Santos, "AI for Open Science: A Multi-Agent Perspective for Ethically Translating Data to Knowledge.," 27 10 2023. [Online]. Available: https://openreview.net/forum?id=vhEUYsNrGa.

[54] C. &. S. J. E. Yakaboski, "Learning the Finer Things: Bayesian Structure Learning at the Instantiation Level," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023.

[55] C. A. &. S. E. Yakaboski, " Efficient Reasoning upon Fusion of Many Data Sources," in *The International FLAIRS Conference Proceedings* , 2021.

[56] C. Genna, "Mining Modern Music: The Classification of Popular Songs," *Advances in Data Science and Information Engineering: Transactions on Computational Science and Computational Intelligence,* pp. 609-616, 30 10 2021.

DSIP Defense SBIR/STTR Innovation Portal

Cost Volume

SBIR Phase I Proposal

Proposal Number	F244-0001-0034
Topic Number	AF244-0001
Proposal Title	KnowFlow: An Interactive, Mutable, Federated Knowledge Graph Environment
Date Submitted	11/06/2024 10:55:04 AM

Firm Information

Firm Name	Technergetics, LLC
Mail Address	114 Genesee Street , Utica, New York, 13502
Website Address	http://www.techngs.com
UEI	XWYVA88GZ1F5
Duns	078840730
Cage	6X4M8

Total Dollar Amount for this Proposal	$139,555.74
Base Year	$69,777.87
Year 2	$69,777.87
Technical and Business Assistance(TABA)- Base	$0.00
TABA- Year 2	$0.00

Base Year Summary

Total Direct Labor (TDL)	$35,082.76
Total Direct Material Costs (TDM)	$0.00
Total Direct Supplies Costs (TDS)	$0.00
Total Direct Equipment Costs (TDE)	$0.00
Total Direct Travel Costs (TDT)	$0.00
Total Other Direct Costs (TODC)	$2,400.00
G&A (rate 18.16%) x Base (TDL+TOH)	$6,371.03
Total Firm Costs	$43,853.79
Subcontractor Costs	
Total Subcontractor Costs (TSC) 1	$20,755.35
Total Subcontractor Costs (TSC)	$20,755.35
Cost Sharing	-$0.00
Profit Rate (8%)	$5,168.73
Total Estimated Cost	$69,777.87
TABA	$0.00

Year 2 Summary

Total Direct Labor (TDL)	$35,082.76

Total Direct Material Costs (TDM)	$0.00
Total Direct Supplies Costs (TDS)	$0.00
Total Direct Equipment Costs (TDE)	$0.00
Total Direct Travel Costs (TDT)	$0.00
Total Other Direct Costs (TODC)	$2,400.00
G&A (rate 18.16%) x Base (TDL+TOH)	$6,371.03
Total Firm Costs	$43,853.79
Subcontractor Costs	
Total Subcontractor Costs (TSC) 1	$20,755.35
Total Subcontractor Costs (TSC)	$20,755.35
Cost Sharing	-$0.00
Profit Rate (8%)	$5,168.73
Total Estimated Cost	$69,777.87
TABA	$0.00

Base Year

Direct Labor Costs

Category / Individual-TR	**Rate/Hour**	**Estimated Hours**	**Fringe Rate (%)**	**Fringe Cost**	**Cost**
Computer and Information Research Scientist/ Principal Investigator (Evi Jesaitis)	$99.03	64	38.81	$2459.75	$8,797.67
Computer Occupations, All Other/ Creative Director	$75.27	24	38.81	$701.09	$2,507.57
Software Developer/ Software Engineer II	$54.47	195	38.81	$4122.26	$14,743.91
Computer and Information Research Scientist/ ML Researcher II	$72.12	66	38.81	$1847.32	$6,607.24

Subtotal Direct Labor (DL)	$32,656.39
Labor Overhead (rate 7.43%) x (DL)	$2,426.37
Total Direct Labor (TDL)	**$35,082.76**

Subcontractor Costs

Subcontractor- Simage Autonomy LLC

Subcontractor/Consultant Budget Information

Category / Individual-TR	**Rate/Hour**	**Estimated Hours**	**Cost**
Computer and Information Research Scientist (SME)	$77.00	269.55	$20,755.35

Subtotal Subcontractor Labor (SL)	$20,755.35

Other Direct Cost

Type	**Vendor**	**Cost**

Total Subcontractor Other Direct Costs 1	$0.00
Total Subcontractor Costs (TSC) 1	**$20,755.35**
Total Subcontractor Costs (TSC1)	**$20,755.35**

Other Direct Costs

AWS Gov Cloud	$2,400.00
Total Other Direct Costs (TODC)	**$2,400.00**

G&A (rate 18.16%) x Base (TDL+TOH)	$6,371.03
Cost Sharing	-$0.00
Profit Rate (8%)	$5,168.73
Total Estimated Cost	$69,777.87
TABA	$0.00

Year 2

Direct Labor Costs

Category / Individual-TR	**Rate/Hour**	**Estimated Hours**	**Fringe Rate (%)**	**Fringe Cost**	**Cost**
Computer and Information Research Scientist/ Principal Investigator (Evi Jesaitis)	$99.03	64	38.81	$2459.75	$8,797.67
Computer Occupations, All Other/ Creative Director	$75.27	24	38.81	$701.09	$2,507.57
Software Developer/ Software Engineer II	$54.47	195	38.81	$4122.26	$14,743.91
Computer and Information Research Scientist/ ML Researcher II	$72.12	66	38.81	$1847.32	$6,607.24

Subtotal Direct Labor (DL)	$32,656.39
Labor Overhead (rate 7.43%) x (DL)	$2,426.37
Total Direct Labor (TDL)	**$35,082.76**

Subcontractor Costs

Subcontractor- Simage Autonomy, LLC

Subcontractor/Consultant Budget Information

Category / Individual-TR	**Rate/Hour**	**Estimated Hours**	**Cost**
Computer and Information Research Scientist (SME)	$269.55	77	$20,755.35

Subtotal Subcontractor Labor (SL)	$20,755.35

Other Direct Cost

Type	**Vendor**	**Cost**

Total Subcontractor Other Direct Costs 1	$0.00
Total Subcontractor Costs (TSC) 1	**$20,755.35**
Total Subcontractor Costs (TSC1)	**$20,755.35**

Other Direct Costs

AWS Gov Cloud	$2,400.00
Total Other Direct Costs (TODC)	**$2,400.00**

G&A (rate 18.16%) x Base (TDL+TOH)	$6,371.03
Cost Sharing	-$0.00
Profit Rate (8%)	$5,168.73
Total Estimated Cost	$69,777.87
TABA	$0.00

Explanatory Material Relating to the Cost Volume
The Official From the Firm that is responsible for the cost breakdown
Name: Kevin Brown
Phone: (419) 344-0070
Phone: kevin.brown@techngs.com
Title: Proposal Owner

If the Defence Contracting Audit Agency has performed a review of your projects within the past 12 months, please provide: Yes
Audit Agency Name: Defense Contract Audit Agency
Audit Agency POC: Tong Sun
Address: 5795 Widewaters Parkway 2nd Floor, Dewitt, New York,13214
Phone: (571) 448-9307
Email: xiaotong.sun.civ@mail.mil
Select the Type of Payment Desired: Partial payments

Cost Volume Details

Direct Labor

Base

Category	Description	Education	Yrs Experience	Hours	Rate	Fringe Rate	Total
Computer and Information Research Scientist	Principal Investigator	Bachelor's Degree	22	64	$99.03	38.81	$8,797.67
Computer Occupations, All Other	Creative Director	Associate's Degree	14	24	$75.27	38.81	$2,507.57
Software Developer	Software Engineer II	Bachelor's Degree	3	195	$54.47	38.81	$14,743.91
Computer and Information Research Scientist	ML Researcher II	Master's Degree	10	66	$72.12	38.81	$6,607.24

Are the labor rates detailed below fully loaded? **NO**

Provide any additional information and cost support data related to the nature of the direct labor detailed above.
The cost model requires two years of data, but the solicitation is only six months. I allocated three months in year one and three in year two to meet DSIP's percentage completion needs.

Labor rate Documentation:

- 2024_Provisional_Rates.pdf

Direct Labor Cost ($): $32,656.39

Year2

Category	Description	Education	Yrs Experience	Hours	Rate	Fringe Rate	Total
Computer and Information Research Scientist	Principal Investigator	Bachelor's Degree	22	64	$99.03	38.81	$8,797.67
Computer Occupations, All Other	Creative Director	Associate's Degree	14	24	$75.27	38.81	$2,507.57
Software Developer	Software Engineer II	Bachelor's Degree	3	195	$54.47	38.81	$14,743.91

Computer and Information Research Scientist	ML Researcher II	Master's Degree	10	66	$72.12	38.81	$6,607.24

Are the labor rates detailed below fully loaded? **NO**

Provide any additional information and cost support data related to the nature of the direct labor detailed above.
We benchmark labor costs to ensure competitive compensation, attract talent, and ensure financial sustainability.

Labor rate Documentation:

- Techngs_2024_Provisional_Rates.pdf

Direct Labor Cost ($): $32,656.39

Sum of all Direct Labor Costs is($): $65,312.78

Overhead

Base

Labor Cost Overhead Rate (%) **7.43**

Apply Overhead to Direct Other Cost? **NO**

Overhead Comments:
Overhead expenses encompass administrative costs linked to government contract management, training, certifications, and IT support.

Overhead Cost ($): **$2,426.37**

Year2

Labor Cost Overhead Rate (%) **7.43**

Apply Overhead to Direct Other Cost? **NO**

Overhead Comments:
Overhead expenses encompass administrative costs linked to government contract management, training, certifications, and IT support.

Overhead Cost ($): **$2,426.37**

Sum of all Overhead Costs is ($):	**$4,852.74**

General and Administration Cost
Base

G&A Rate (%):	**18.16**
Apply G&A Rate to Overhead Costs?	**YES**
Apply G&A Rate to Direct Labor Costs?	**YES**
Apply G&A Rate to Subcontractor Costs?	**NO**
Apply G&A Rate to Other Direct Costs?	**NO**
Please specify the different cost sources below from which your company's General and Administrative costs are calculated. **General and administrative costs consist of general business expenses for running and executing operations.**	
G&A Cost ($):	**$6,371.03**

Year2

G&A Rate (%):	**18.16**
Apply G&A Rate to Overhead Costs?	**YES**
Apply G&A Rate to Direct Labor Costs?	**YES**
Apply G&A Rate to Subcontractor Costs?	**NO**
Apply G&A Rate to Other Direct Costs?	**NO**
Please specify the different cost sources below from which your company's General and Administrative costs are calculated. **General and administrative costs consist of general business expenses for running and executing operations.**	
G&A Cost ($):	**$6,371.03**
Sum of all G&A Costs is ($):	**$12,742.06**

Subcontractor/Consultants
Base

Subcontractor/Consultant:
Simage Autonomy LLC

Budget Contact Name	Budget Contact Title	Budget Contact Phone	Budget Contact Email
Dr. Chase Yakaboski	President	(850) 218-7437	chase@simage.ai

Do you have a letter of commitment from the subcontractor/consultant? **YES**

Document uploaded for the letter of commitment:

- Attach_4_Commitment_Letter.pdf

Are you able to provide detailed budget information for this subcontractor/consultant? **YES**

Total Cost($) : **$20,755.35**

Detailed Budget Information				
Labor Category	**Description**	**Hours**	**Rate**	**Cost**
Computer and Information Research Scientist	SME	269.55	$77.00	$20,755.35

Additional Costs		
Type	**Amount**	**Explanation**
Overhead:		
G&A:		
Profit:		

Other Direct Costs			
Category	**Description**	**Vendor**	**Cost**

Year2

Subcontractor/Consultant:
Simage Autonomy, LLC

Budget Contact Name	Budget Contact Title	Budget Contact Phone	Budget Contact Email
Dr. Chase Yakaboski	President	(850) 218-7437	chase@simage.ai

Do you have a letter of commitment from the subcontractor/consultant? **YES**

Document uploaded for the letter of commitment:

- Attach_4_Commitment_Letter.pdf

Are you able to provide detailed budget information for this subcontractor/consultant? **YES**

Total Cost($) : **$20,755.35**

Detailed Budget Information				
Labor Category	**Description**	**Hours**	**Rate**	**Cost**
Computer and Information Research Scientist	SME	77	$269.55	$20,755.35

Additional Costs		
Type	**Amount**	**Explanation**
Overhead:		
G&A:		
Profit:		

Other Direct Costs			
Category	**Description**	**Vendor**	**Cost**

Total Subcontractors/Consultants Cost ($): **$41,510.70**

ODC-Other

Base

Description: AWS Gov Cloud

Vendor: AWS

Quantity: 3

Total Cost ($): $2,400.00

Competitively Sourced? yes

Exclusive for this Contract? yes

Supporting Comments:

Supporting Documents:

- AWS Cost Estimate.pdf

Year2

Description: AWS Gov Cloud

Vendor: AWS

Quantity: 3

Total Cost ($): $2,400.00

Competitively Sourced? yes

Exclusive for this Contract? yes

Supporting Comments:

Supporting Documents:

- AWS Cost Estimate.pdf

ODC-Summary

Base

Do you have any additional information to provide? **NO**

Year2

Do you have any additional information to provide? **NO**

Profit Rate/Cost Sharing

Base

Cost Sharing ($): **-**

Cost Sharing Explanation:

Profit Rate (%): **8**

Profit Explanation:

Total Profit Cost ($): **$10,337.46**

Year2

Cost Sharing ($):	**-**
Cost Sharing Explanation:	
Profit Rate (%):	**8**
Profit Explanation:	
Total Profit Cost ($):	**$10,337.46**
Total Proposed Amount ($):	**$139,555.74**

Technergetics, LLC

DISCLAIMER: Information provided herein is privileged and confidential, and not subject to disclosure, pursuant to 15 U.S.C. 638 (k)(4) and 5 U.S.C. 552. This information shall only be used or disclosed for evaluation purposes.

SBIR Company Commercialization Report

Total Investments:	Total Sales:	Total Patents:	Government Designated Phase III Funding:
$5,990,000.00	$5,990,000.00	0	$5,990,000.00

Company Information

Address:

114 Genesee Street
Utica, NY 13502-3526
United States

SBC Control ID: SBC_000700589

Company Url: http://www.techngs.com

Company POC

Title: President
Full Name: Robert Harrod
Phone: 3152712096 **Ext** 701
Email: robert.harrod@techngs.com

Commercialization POC

Title: President
Full Name: Robert Harrod
Phone: (315) 271-2096 **Ext** 701
Email: robert.harrod@techngs.com

Additional Company Information

% Revenue for last fiscal year from SBIR/STTR funding:	Total revenue for last fiscal year:
18.0%	$5,000,000 - $19,999,999
Year Founded:	**# Employees Currently:**
2013	62
Year first Phase I award received:	**# SBIR/STTR Phase I Awards:**
2021	2
Year first Phase II award received:	**# SBIR/STTR Phase II Awards:**
2020	6
# Employees at first Phase II award:	**Mergers and Acquisition within past 2 years:**
30	No
Spin-offs resulting from SBIR/STTR:	**IPO resulting from SBIR/STTR | Year of IPO:**
No	No | N/A
Patents resulting from SBIR/STTR | #Patents:	**List of Patents:**
No | N/A	
Woman-Owned:	**Socially and Economically Disadvantaged:**
No	No
HUBZone-Certified:	**SBC majority-owned by multiple VCOC, HF, PE firms | By what percent (%):**
No	No | N/A

Additional Investment From

	Last Submitted Version (11-21-2022 02:28 PM)	Current Version
DoD contracts/DoD subcontracts	$5,150,346.00	$5,990,000.00
Angel Investors	$0.00	$0.00
Venture Capital	$0.00	$0.00
Self Funded	$0.00	$0.00
Private Sector	$0.00	$0.00
Other Federal Contracts/Grants	$0.00	$0.00
Other Sources	$0.00	$0.00
Additional Investment	$0.00	$0.00
Total Investment	$5,150,346.00	$5,990,000.00

Phase III Sales To

	Last Submitted Version (11-21-2022 02:28 PM)	Current Version
DoD or DoD prime contractors	$5,150,346.00	$5,990,000.00
Private Sector	$0.00	$0.00
Export Markets	$0.00	$0.00
Other Federal Agencies	$0.00	$0.00
Additional commercialization by 3rd Party Revenue	$0.00	$0.00
Other Customers	$0.00	$0.00
Additional Sales	$0.00	$0.00
Total Sales	$5,150,346.00	$5,990,000.00

Government Phase III Contracts

	Last Submitted Version (11-21-2022 02:28 PM)	Current Version
Funding Obligated	$5,150,346.00	$5,990,000.00

Commercialization Narrative

Technergetics has received two Phase III contracts in FY22.

Our Phase III contract with Air Force District of Washington (AFDW) and Air Mobility Command (AMC), Contract #FA7014-22-D-0006, valued at $99M, is commercializing technology to modernize the Aerial Port into the Aerial Port of the Future.

Our second Phase III contract with Air Force District of Washington (AFDW) and HAF/A4L (Tesseract), Contract # FA7014-22-D-0005, valued at $99M, is commercializing Storefront capabilities for Air Mobility. This contract follows a Direct to Phase 2 contract, contract #FA8649-22-P-0844, awarded by AFWERX on 04 May 2022. Technergetics, LLC has received $2,100,346 in awards/sales on this Phase III contract. Once the SBIR.gov site updates with the Phase 2 contract #FA8649-22-P-0844, the Phase III award information will be entered.

Commercialized Awards

- Listed below are the sales revenue and investment details resulting from the technology developed under these SBIR/STTR awards.

URSUS: The Digital Platform for DoD Commerce — 1 of 2

Agency/Branch:	Department of Defense/Air Force	**Manufacturing related**	No | N/A
Program/Phase/Year:	SBIR/Phase II/2022	**Subsidiaries**	N/A
Topic #:	AF221-DCSO1	**Other contributing SBIR/STTR awards**	N/A
Contract/Grant #:	FA8649-22-P-0844	**Used in Federal or acquisitions program?**	No
Achieved a cost saving or cost avoidance?:	No		

Additional Investment From		**Phase III Sales To**	
DoD contract/subcontract:	$2,340,000.00	**Dod or DoD prime contractors:**	$2,340,000.00
Other Federal contract/grants:	$0.00	**Other Federal Agencies:**	$0.00
Angel Investors:	$0.00	**Private Sector:**	$0.00
Venture Capital:	$0.00	**Export Market:**	$0.00
Self-Funded:	$0.00	**3rd Party Revenue:**	$0.00
Private Sector:	$0.00	**Other Customers:**	$0.00
Other Sources:	$0.00		
Investment Total:	**$2,340,000.00**	**Sales Total:**	**$2,340,000.00**

Government Designated Phase III Contracts

Funding Agreement /	Agency	Project Title	Year Awarded	Funding Obligated

Contract #				
FA7014-22-D-0005	USAF	Strategic Pursuit of Enterprise Advancement & Readiness (SPEAR)	2022	$2,340,000.00

UPgrade: Digital Transformation for Space-Available (S/A) Passenger Experience

2 of 2

Agency/Branch:	Department of Defense/Air Force
Program/Phase/Year:	SBIR/Phase II/2021
Topic #:	AF203-CSO1
Contract/Grant #:	FA8649-21-P-1495
Achieved a cost saving or cost avoidance?:	No
Manufacturing related	No | N/A
Subsidiaries	N/A
Other contributing SBIR/STTR awards	N/A
Used in Federal or acquisitions program?	No

Additional Investment From

DoD contract/subcontract:	$3,650,000.00
Other Federal contract/grants:	$0.00
Angel Investors:	$0.00
Venture Capital:	$0.00
Self-Funded:	$0.00
Private Sector:	$0.00
Other Sources:	$0.00
Investment Total:	**$3,650,000.00**

Phase III Sales To

Dod or DoD prime contractors:	$3,650,000.00
Other Federal Agencies:	$0.00
Private Sector:	$0.00
Export Market:	$0.00
3rd Party Revenue:	$0.00
Other Customers:	$0.00
Sales Total:	**$3,650,000.00**

Government Designated Phase III Contracts

Funding Agreement / Contract #	Agency	Project Title	Year Awarded	Funding Obligated
FA8649-21-P-0287	USAF	Precision Logistics & Advancement in Transportation Optimization (PLATO)	2022	$3,650,000.00

CERTIFICATE OF COMPLETION

THIS CERTIFICATE IS PRESENTED TO

Kevin Brown, Technergetics, LLC

FOR SUCCESSFULLY COMPLETING FRAUD, WASTE AND ABUSE TRAINING AND MEETING ALL REQUIREMENTS SET FORTH BY THE OFFICE OF SMALL BUSINESS PROGRAMS

Oct 07, 2024

COMPLETION DATE

Oct 07, 2025

EXPIRATION DATE