

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: F244-0001-0048
Proposal Title: IKOS: Interactive Knowledge Graphs for Open-Source Intelligence Analysis

Agency Information

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

Firm Information

Firm Name: The Bulls Run Group, Inc
Address: 9207 Bulls Run Pkwy, Bethesda, MD 20817-2403
Website: https://www.bulls.run
UEI: KDH9SC281WL8
DUNS: 129524731
CAGE: 8LSH4
SBA SBC Identification Number: 001972933

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) | 10 |
| 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_001972933 |

Supporting Documentation:

- [SBA Registration.pdf](#)

5. It has more than 50% owned by a <u>single</u> Venture Capital Owned Company (VCOC), hedge fund, or private equity firm	NO
6. It has more than 50% owned by <u>multiple</u> 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	YES
12. The offeror qualifies as a:	
<div><input type="checkbox"/> Socially and economically disadvantaged SBC</div> <div><input type="checkbox"/> Women-owned SBC</div> <div><input checked="" type="checkbox"/> HUBZone-owned SBC</div> <div><input checked="" type="checkbox"/> Veteran-owned SBC</div> <div><input checked="" type="checkbox"/> Service Disabled Veteran-owned SBC</div> <div><input type="checkbox"/> None Listed</div>	
13. Race of the offeror:	
<div><input type="checkbox"/> American Indian or Alaska Native</div> <div><input type="checkbox"/> Native Hawaiian or Other Pacific Islander</div> <div><input type="checkbox"/> Asian</div> <div><input checked="" type="checkbox"/> White</div> <div><input type="checkbox"/> Black or African American</div> <div><input type="checkbox"/> Do not wish to Provide</div>	
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
Stoney Trent	Stoney Trent	President	The Bulls Run Group, Inc	06/01/2024

Audit Information

Summary:

Has your Firm ever had a DCAA review?**NO**

VOL I - Proposal Summary

Summary:

Proposed Base Duration (in months):

6

Technical Abstract:

IKOS will build upon recent advancements in knowledge graph extraction to deliver software that will enable analysts to interact with complex multimodal knowledge graphs in a low-effort, intuitive manner. We will do so by leveraging a government off-the-shelf platform, Information Operations Network (ION), for realistic multimodal data. We will orient the functionality of our knowledge graph interactivity on insights gleaned through a campaign of Mission Engineering for Open-Source Intelligence Analysts. These insights will allow us to determine the technical feasibility, acceptability, and suitability of IKOS software that can be delivered within existing and/or emerging analytic support platforms.

IKOS will improve upon RESIN and RESIN-EDITOR to better support knowledge graph interaction. We will integrate advanced capabilities that allow for automatic fine-grained knowledge completion and real-time knowledge propagation based on user inputs. We will develop the capability to flag logical inconsistencies and conflicts as the analyst interacts with the graph, identify, and localize inconsistencies with analyst-specified domain knowledge. We will develop a new feature that automates schema matching and knowledge graph refinement, to include the ability to suggest additional edges between related nodes, infer new relationships, highlight information gaps, and automatically flag inconsistencies as users interact with the graph. We will implement a natural language interface, allowing analysts to input modifications more intuitively, which will be automatically propagated throughout the graph, ensuring coherence and consistency across all connected data points. We will implement sensitivity analysis that directs analysts' attention to aspects of the graph that require their revisions. Finally, we will adapt approaches to translating expertise into probabilistic parameters to afford analysts the ability to introduce their expert intuition into knowledge graphs.

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

IKOS will significantly enhance current approaches to knowledge graph interaction by enabling dynamic multimodal knowledge graph refinement. Rather than relying solely on the analyst's manual input or requiring extensive retraining of underlying models, the system will integrate user corrections, identify contradictions, and adjust the graph's fine-grained

structure in real-time. As a result, analysts will be able to rapidly complete and refine large-scale, multimodal knowledge graphs without needing to manually track every update and change throughout the graph.

IKOS will significantly reduce the cognitive burden on analysts, enabling them to focus on higher-order reasoning tasks instead of labor-intensive knowledge management. IKOS will provide numerous benefits for intelligence analysis applications, including more accurate and up-to-date situational awareness, improved detection of emerging patterns, and improved decision-making in time-sensitive contexts. IKOS will enable analysts to easily refine large complex knowledge structures while enhancing their completeness and consistency. IKOS knowledge structures will become an even more practical and robust tool that supports a wide range of intelligence applications. IKOS will address key limitations in current state-of-the-art systems, leading to a new generation of dynamic, multimodal knowledge graphs that are not only comprehensive but also agile enough to meet the evolving demands of real-world intelligence analysis environments.

The initial focus for IKOS is to deliver capabilities to military intelligence programs of record, beginning with the Command and Control of the Information Environment (C2IE), an analytic platform that provides open-source data as a means of real-time situational awareness. IKOS’s multimodal data integration will strengthen C2IE’s capacity to deliver actionable intelligence. In the private sector, IKOS will be positioned as a business intelligence (BI) solution for critical infrastructure firms (e.g., energy providers, financial firms), where real-time insights drive essential decision-making.

Attention:

Disclaimer: For any purpose other than to evaluate the proposal, this data except proposal cover sheets shall not be disclosed outside the Government and shall not be duplicated, used or disclosed in whole or in part, provided that if a contract is awarded to this proposer as a result of or in connection with the submission of this data, the Government shall have the right to duplicate, use or disclose the data to the extent provided in the funding agreement. This restriction does not limit the Government's right to use information contained in the data if it is obtained from another source without restriction. This restriction does not apply to routine handling of proposals for administrative purposes by Government support contractors. The data subject to this restriction is contained on the pages of the proposal listed on the line below.

Addition:

Enter the page numbers separated by a space of the pages in the proposal that are considered proprietary:

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

Knowledge Graphs, Artificial Intelligence, Cognitive Engineering, Collaborative analysis, Mission Engineering, Intelligence Analysis

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

YES

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.

Firm POW	100%
Subcontractor POW	0%

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.	YES
--	------------

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 included in Volume V: Supporting Documents of this proposal.	YES
---	------------

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?	NO
--	-----------

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?	10%
---	------------

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
--	------------

VOL I - Contact Information

Principal Investigator

Name: **Dr. Stoney Trent**

Phone: **(301) 456-4237**

Email: **stoney@bullsrungroup.com**

Address: **9207 Bulls Run Pkwy, Bethesda, MD 20817 - 2403**

Corporate Official

Name: **Dr. Stoney Trent**

Phone: **(301) 456-4237**

Email: **stoney@bullsrungroup.com**

Address: **9207 Bulls Run Pkwy, Bethesda, MD 20817 - 2403**

Authorized Contract Negotiator

Name: **Dr. Stoney Trent**

Phone: **(301) 456-4237**

Email: **stoney@bullsrungroup.com**

Address: **9207 Bulls Run Pkwy, Bethesda, MD 20817 - 2403**

IKOS: Interactive Knowledge Graphs for Open-Source Intelligence Volume 2: Technical Volume

1. Identification and Significance of the Problem

a. Current state of Open-Source Intelligence (OSINT) analysis.

OSINT has traditionally been the province of federal military and intelligence organizations (Williams & Blum, 2018). However, as an increasingly large share of human behavior becomes technologically mediated and made accessible on social media and other platforms, the tools, techniques, and procedures of government OSINT have been adopted and expanded by non-government organizations (NGOs) and the private sector, including journalists, human rights advocates, and private investigators (Dubberley et al., 2019). Although traditional intelligence analysis has been dominated by individual efforts with a strong sense of ownership, and intelligence agencies have been criticized for a lack of cooperation, these new forms of OSINT are increasingly social in nature (Belghith et al., 2022). Some even leverage expertise from both public and private entities.

Contemporary OSINT often emerges as “expert-led crowdsourcing,” with lead analysts managing a framework for contributions from professionals and hobbyists. The CrowdSolve project brought together detectives, forensics experts, and over 200 true crime enthusiasts to spend a weekend reviewing case files to make progress in two cold cases (Venkatagiri et al., 2021). Sedition Hunters, an OSINT community organized on Twitter, has contributed evidence cited in several FBI arrest warrants for suspects in the January 6th US Capitol riot (Yu et al., 2023). Trace Labs, a Canadian non-profit, partners with law enforcement to conduct crowdsourced OSINT investigations of missing persons, structured as a Capture-the-Flag competition where teams (including US military personnel (Stover, 2019)) earn points for collecting relevant evidence. These successful examples, where officials define investigation targets and/or directly supervise the crowd’s efforts, contrast with prominent failures of “digilantes” that result in misidentified suspects, as in the 2013 Boston Marathon Bombing (Nhan et al., 2017) and the 2017 Charlottesville white nationalist rally (Victor, 2017).

OSINT Challenges

- *Massive amount of multi-modal data*
- *Increasingly crowd-sourced and performed by non-government groups*
- *Team efforts involving variety of expertise*
- *Information about process is important for accurate judgments about rigor*
- *Includes a variety of objectives, approaches, and mindsets*

In the last twenty years, cognitive engineering research has documented challenges for intelligence analysis. Some of these challenges apply to IKOS. For example, Zelik, Paterson, and Woods (2007a), found that professional intelligence analysts make better judgments about analytic rigor of reports when provided with information about the analytic process. Trent, Patterson, and Woods (2007) observed that intelligence analysis is typically performed by teams comprising modestly skilled analysts who are supervised by a more experienced analyst. This conveys the importance of designing to support teamwork and supervision rather than individual work. Straus, Parker, and Bruce (2011) found that analysts struggle with polarization, overconfidence, and pressures toward uniformity. Research in Team Cognition has determined that introducing new technologies does not necessarily improve performance on complex cognitive tasks. Collectively, the corpus of research suggests that a one-size-fits-all AI model might be unhelpful or counterproductive. Knowledge graphs offer ways to accommodate information sharing across a variety of analytic objectives, approaches, and mindsets.

b. Current state of knowledge graphs supporting collaborative analysis.

Knowledge graphs represent concepts about the world in an interpretable structure (Hogan et al. 2021). Because of their ability to capture higher-level real-world knowledge in a human and machine-understandable representation, knowledge graphs have found broad downstream applications, including question answering (Ding et al. 2022), financial forecasting (Liu et al. 2019), image understanding (Aditya et al. 2019), and future event prediction (Li et al. 2021). In our own work, we have leveraged knowledge graphs for intelligence-related tasks such as identifying anomalous human movement (Bavikadi et al., 2024), predicting cyber attacks

(Sarkar et al., 2019), and detecting, attributing, and characterizing disinformation (Fung et al., 2021). Ideally, knowledge graphs are derived from reliable data or an overarching theoretical framework, which determine the conditions and hypotheses under which the model is valid or “true”. However, these models are only useful insofar as analysts can effectively interpret the model outputs, identify their limitations, and apply new knowledge or human expertise to improve on the models’ capabilities and representations.

Providing analysts with information about the limitations of these models and the knowledge graphs they produce is challenging. A model may have incorrect parameters, the chosen values may not be good approximations of reality, or the knowledge graphs may be derived from noisy or incomplete data.

Researchers often refer to this challenge as “The Black Box Menace,” where the behavior and impact of model inputs or parameters are unknown and can propagate uncertainty during analysis. Begoli, Bhattacharya, and Kusnezov (2019) highlight the issue that the “absence of theory” in purely data-driven models often makes proper uncertainty quantification an essential ingredient—even if the model behavior cannot be explicitly understood, the propagation of uncertainty and sources of uncertainty in the model inputs and outputs can be understood. However, simply highlighting the sources of uncertainty in a model only provides insight into how reliable or useful the model may be in certain contexts. It does not provide analysts with the tools necessary to improve and update the model to increase its overall efficacy.

Knowledge Graph Challenges

- *Difficult to interpret and update*
- *KG limitations are often hidden*
- *Most extraction methods are limited to text*
- *Multimodal knowledge representations rely on visual detectors trained offline on supervised data*

Even capturing relevant knowledge and encoding it in a graph remains challenging. Several human-curated knowledge graph resources, such as Wikidata (Vrandečić et al. 2014) and Freebase (Bollacker et al. 2008), have been created, but such approaches are very expensive to create, may not translate well to new domains, and often suffer from incompleteness (Zhang et al. 2018). Automated knowledge induction methods relying on text information extraction have also been proposed (Mausam et al. 2016), but such methods may be prone to many types of errors, such as incorrect information extraction output and failures in entity resolution (Gupta et al. 2019). Seo et al. (2021) leverage active learning techniques to expand and refine a knowledge graph from text, but do not attempt to refine a multimodal knowledge graph.

c. Current state of Multimodal Knowledge Graph Creation and Interaction Capabilities.

Most recent approaches to knowledge graph creation involve deep learning methods to learn graph embeddings that capture the meaning of relations (e.g. semantic relations, temporal ordering, and others) between entities within the graph (Fung et al. 2021; Zhang et al. 2021). Such embeddings are typically used to initialize a graph neural network, which is then trained to perform the downstream task of interest (Kato et al. 2018; Fung et al. 2021; Marino et al. 2017).

While there has been some recent work in representing multimodal knowledge about activities in structured graphs (Li et al. 2020; Li et al. 2022; Wen et al. 2021; Chen et al. 2021), most knowledge graph extraction methods work purely on text. Extracting multimodal knowledge graphs remains a very challenging problem because it requires a fine-grained understanding of how knowledge from one modality relates to data from a different modality (e.g. what is the relationship between a particular entity in an image with a word in text or a location on a map). At a minimum, knowledge graphs incorporating visual data require being able to detect relevant visual content (e.g., objects or events) and then ground it to appropriate points within the graph. Automated visual understanding capabilities are limited, however, and if the appropriate objects are not able to be detected, the relevant multimodal links will not be discovered. However, most existing methods for multimodal knowledge representation rely on visual detectors trained offline on supervised data (Li et al. 2020, Wen et al. 2021). Thus, knowledge graphs extracted by such systems are often incomplete and have some amount of noise.

Recent advances in artificial intelligence and information extraction have enabled the construction of sophisticated multimodal knowledge graphs (KGs) that capture structured representations of entities, events, and their relationships across diverse data sources. A general depiction of the functions comprising KG extraction can be found in Figure 1. Today, graphs go beyond traditional text-based representations, integrating information from multiple modalities such as visual, audio, geospatial, and temporal sources (Ma et al. 2022, Liang et al. 2024, Wilcke et al. 2023). By linking entities and events across these various data sources, such knowledge graph extraction systems provide a comprehensive and contextualized view of complex real-world scenarios (Wen et al. 2021, Du et al. 2022). This multimodal fusion supports a range of applications, including situational awareness, event prediction (Li et al. 2021), and intelligence analysis, by offering decision-makers a detailed and interpretable structure of interconnected facts and relationships (Rajabi et al. 2024).

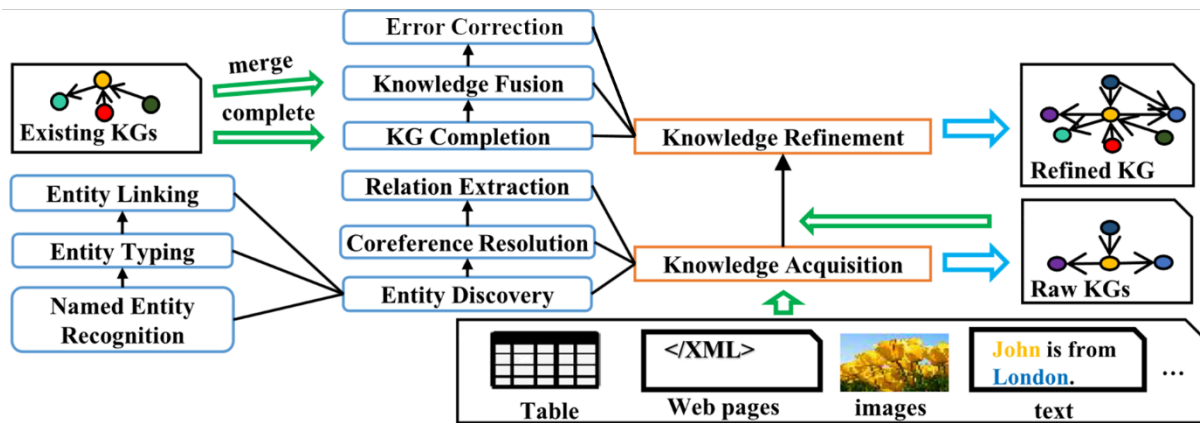


Figure 1 – Multimedia, Entity-Centric Knowledge Extraction. Adapted from Zhong et al. (2023). A suite of techniques to learn the types, links, and relations between entities.

Despite their promise, these existing systems face significant challenges when dealing with dynamically evolving and incomplete knowledge. Multimodal knowledge graphs often contain gaps or errors due to limitations in the initial extraction process, the inherent uncertainty of multimodal data, or changes in the real-world environment that render previously extracted information outdated (Wang et al. 2024, Cheng et al. 2024). Addressing these issues through manual corrections is not only labor-intensive but also impractical at scale, especially for large, complex knowledge graphs used in operational environments. Moreover, many existing systems (Wen et al. 2021, Du et al. 2022) lack the ability to automatically detect non-obvious inconsistencies within the graph and adjust their representations based on new human input. This creates a need for methods that allow analysts to refine and update knowledge structures efficiently and in so doing, minimizing the cognitive and time burden on users.

Knowledge graph refinement is a broad term that encompasses a range of methods aimed at improving the accuracy, completeness, and quality of knowledge graphs. This process includes various sub-tasks such as error correction, schema alignment, and knowledge graph completion. Refinement involves adding, modifying, or removing facts and relationships to keep the graph up-to-date, correct inaccuracies, and resolve conflicts between entities (Subagdja et al., 2024). While some early refinement methods relied heavily on manual adjustments, modern approaches increasingly leverage automated techniques, such as machine learning models and rule-based algorithms, enabling large-scale, dynamic graphs to adapt rapidly to new information (Subagdja et al., 2024). This automation is particularly useful in fast-paced contexts like military intelligence, where rapid adjustments and high accuracy are crucial.

Within the broader process of KG refinement, **knowledge graph completion** specifically focuses on predicting and filling in missing information within the graph (Shen et al. 2022). Completion techniques use the existing structure and relationships in the KG to infer new connections between entities (Wang et al. 2022). By employing methods like graph embeddings, completion models can identify plausible links based on

the graph's topology (Lin et al, 2023). While completion enhances the comprehensiveness of the graph by addressing gaps, it is primarily geared toward symbolic graphs (Chen et al. 2022) and faces challenges when applied to multimodal data (Chen et al. 2022). The presence of diverse data types and complex cross-modal relationships makes it difficult to generate accurate inferences consistently without introducing noise or errors.

While refinement and completion methods are helpful in creating richer and more accurate multimodal knowledge graphs, most existing systems are not designed to integrate human input with these automated processes. Current methods either rely on *specialized interfaces* that require users to manually manipulate the graph's structure (Ilievski et al. 2020), depend on *models that are inflexible* to new types of multimodal knowledge, either requiring few-shot examples (Zhang et al. 2022, Li et al. 2024), or process different modalities independently (Chen et al. 2023). These sources of rigidity create bottlenecks that are incompatible with work that requires rapid, real-time updating. Further, these methods *lack mechanisms for effectively surfacing deeper, cross-modal logical inconsistencies or resolving internal contradictions* that arise from newly incorporated data or human edits. This reduces the trust analysts place in these systems and diminishes their purpose of providing human-interpretable structures in operational scenarios that require high-confidence reasoning and decision support. Further, current methods for refinement and completion lack the ability to allow domain knowledge from intelligence analysts to be directly entered into the system. For example, an analyst may see a connection made between rival criminal groups in a given area of interest, and the analyst want to prevent such connections from being automatically inferred, or at least have them flagged as requiring further examination. Along those lines, certain results with respect to knowledge graph

KG Interaction Shortcomings

- *Extensive manual effort*
- *Rely on specialized tools and interfaces*
- *Low flexibility for analysts*
- *Poor generalization to multimodal data*
- *Rely on inflexible pre-trained models*
- *Inefficient integration of human inputs*
- *lack mechanisms to identify or resolve logical inconsistencies and contradictions*

completion may require explanation to the analyst team – as the team may debate the conclusion drawn by the algorithm. Here, the team gains benefit from knowing the algorithmic provenance of a system-performed completion. In our experience in deploying real-world cyber security systems as a result of the IARPA CAUSE program (e.g., see Almukaynizi et al., 2017 which resulted in "CYR3CON" - a commercial offering used by over 80 paying customers) we found that intelligence analysts working at multiple Fortune 500 companies required product features that provided provenance of algorithmically-concluded evidence.

d. RESIN - A Multimodal Knowledge Graph Extractor

Recent DARPA projects have advanced the field of multimodal knowledge extraction and knowledge graph creation through an evolving series of systems. The GAIA knowledge extraction system (Li et al. 2020), developed as part of the DARPA AIDA program, marked an initial step toward creating structured multimodal knowledge graphs. GAIA integrates data from multiple modalities—text, visual, and audio—using a human-built ontology to link entities mentioned in text to visual regions, extract event information, cluster coreferential entities, and incorporate external knowledge from sources like Wikipedia. Importantly, GAIA supports the generation of alternative hypotheses, allowing analysts to explore different perspectives on complex topics by interacting with the automatically extracted graph (Li et al. 2020).

However, GAIA is limited by its reliance on conventionally trained, static object detectors and grounding methods. While it excels at integrating multimodal data, its components are less adaptable to dynamically evolving knowledge and real-world complexity. Recognizing these challenges, researchers in the DARPA KAIROS program introduced RESIN (Du et al. 2022), an extension of GAIA designed to enhance its ability to process and understand complex events more effectively. RESIN refines GAIA's components and incorporates advanced techniques to handle open-world data and addresses several GAIA's shortcomings. Consequently, knowledge graphs constructed by RESIN support more robust event representations and afford deeper cross-modal reasoning.

As depicted in Figure 2, RESIN preprocesses and stores multimedia data to extract knowledge graphs. Because there are potentially many pieces of media about the same event or subject, connections between media are specified in a simple TSV file format to associate multiple documents, images, video, with one another. Once the data has been preprocessed, a single docker command is issued which points the system to the folder from which to extract a knowledge graph. This command begins by running several natural language processing dockers for preprocessing the text, running entity and event extraction, linking coreferential events and entities, and predicting the relationships between entities and events. These results are formulated into a text-only knowledge graph. Next, the system invokes the dockers for visual knowledge extraction. These systems extract visual knowledge, such as objects and faces, in addition to recognizing buildings, flags, and well-known public figures, as well as event-centric visual knowledge, such as the event shown, its arguments, and semantic roles. These results are stored in a visual knowledge base. Finally, cross-media fusion links the two knowledge graphs together to form a multimodal knowledge graph.

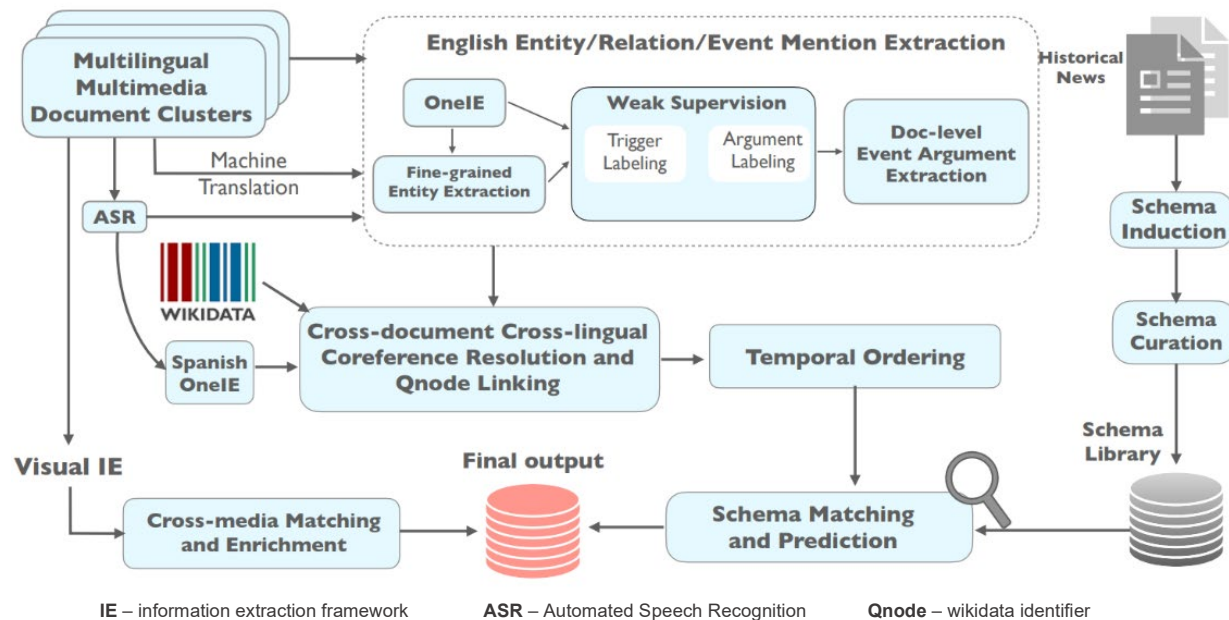


Figure 2 - RESIN – a multimodal knowledge graph extraction system from Du et al. (2022). The RESIN system constructs a multimodal knowledge graph by integrating data from text, speech, images, and video.

RESIN-EDITOR, depicted in Figure 3, provides an interactive platform that allows users to explore hierarchical knowledge structures extracted from multimedia and multi-document data (Nguyen et al. 2023). Through this interface, analysts can visually track knowledge and their relationships, inspect temporal links, and trace the provenance of each piece of knowledge back to its source document. The tool further supports schema-guided event prediction, so users can understand both observed events and potential next steps based on hierarchical schemas. Additionally, RESIN-Editor supports editing features which enable users to modify nodes, relationships, and temporal structures directly within the graph. This level of interaction allows users to not only correct AI/ML-extracted data but also to refine the graph in real-time for a more accurate and detailed situational analysis. Unfortunately, RESIN-Editor does not support natural language inputs or automatic fine-grained graph refinement, knowledge graph completion or inconsistency detection, which makes it less suited for OSINT analysis.

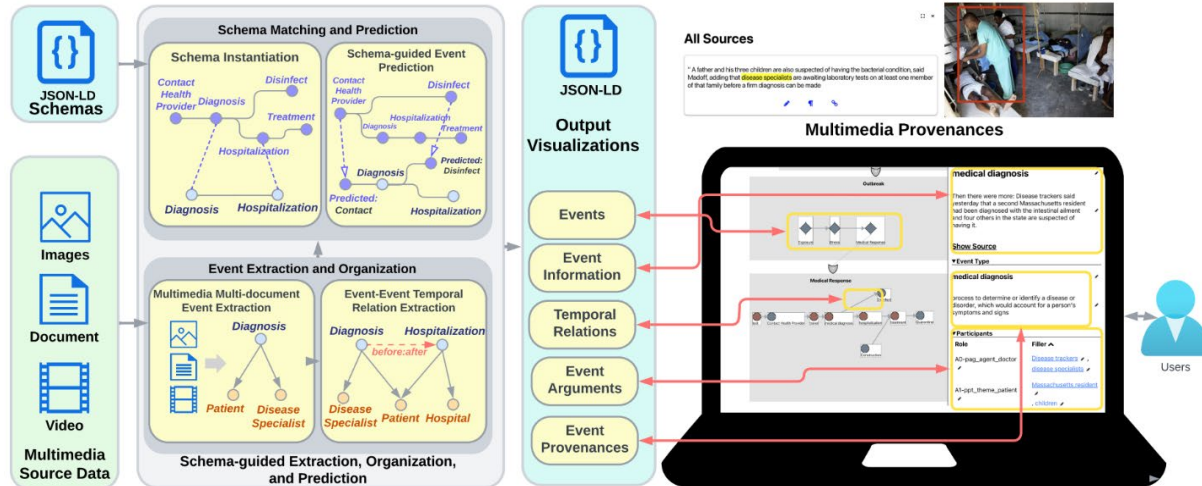


Figure 3 - The RESIN-Editor interface (Nguyen et al. 2023) - an interactive platform that allows users to explore hierarchical knowledge structures extracted from multimedia and multi-document data

IKOS co-PI, Dr. Thomas has extended the foundational technologies in RESIN by integrating more powerful multimodal large language models (LLMs) in place of several key RESIN components to support open-world fine-grained reasoning within complex knowledge structures (Tang et al. 2024). A key strength of this recent work is its ability to extract and organize multimodal knowledge from diverse sources—including text, images, video, and audio—into a structured knowledge graph which supports fine-grained logical inferencing (See Figure 4). These capabilities are essential for constructing fine-grained and accurate knowledge graphs and provide a solid foundation for IKOS. The extraction process leverages a suite of cross-document and cross-media co-reference resolution and advanced entity linking methods to merge disparate sources into a unified, fine-grained representation. Having such well-integrated representations is critical when analysts introduce new inputs or modify existing nodes, as it enables the system to accurately propagate that new information throughout the graph to maintain a coherent structure as the graph is refined. In addition, Thomas has already developed and deployed interactive visualizations of the multimodal knowledge graphs extracted using their system on the web as a feature of an article in *Columbia Engineering Magazine*, “When Trust is a Factor”. Magazine readers were able to interact and explore a knowledge graph extracted from a set of multimedia using Thomas’s system (Currin et al. 2023).

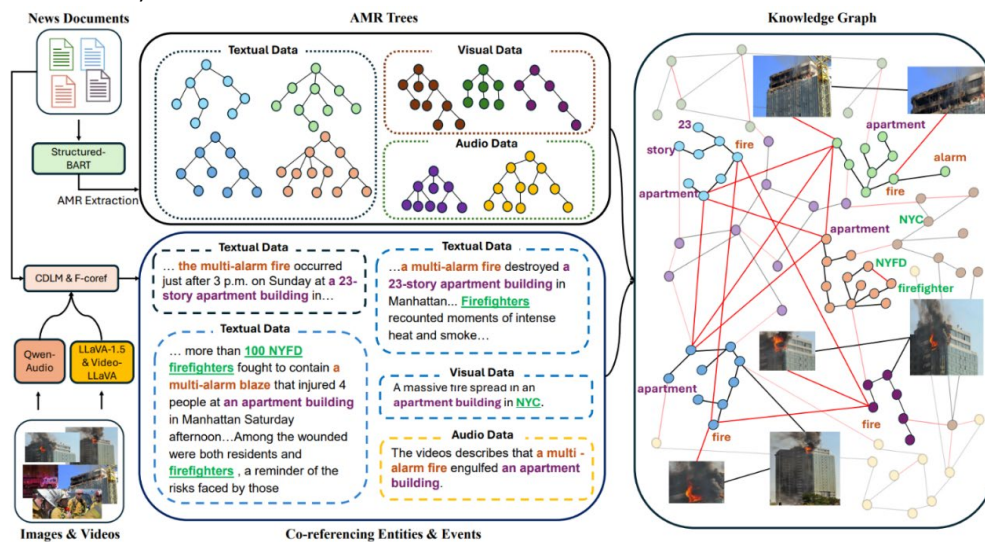


Figure 4 - Extracting multimodal knowledge structures from collections of multimedia data about an event (Tang et al. 2024)

Building on these extraction capabilities, Dr. Thomas has developed a system for multimodal fine-grained logical reasoning and consistency checking (Tang et al. 2024, Chen et al. 2024). This system facilitates detailed logical inferences between elements of the knowledge graph and a graph extracted from a claim (See Figure 5). Although initially designed to generate and verify entailment or contradiction labels extracted for claims for fact-checking, these core techniques can be adapted to support the proposal's objective of dynamically updating interactive knowledge graphs. Specifically, the fine-grained reasoning models can be employed to automatically infer new knowledge, identify, and resolve contradictions, and complete missing logical links in response to user modifications *within the knowledge graph*. Thus, user edits to the knowledge graph would trigger automated updates across related nodes and edges, ensuring logical consistency, highlighting information gaps, and suggesting additional changes to maintain coherence. For example, the current fine-grained logical generation approach presented in Tang et al. (2024)—originally developed to synthesize logically consistent or contradictory claims based on multimodal evidence—can be adapted to support fine-grained knowledge graph completion. This would enable the system to proactively suggest additional logical connections or detect inconsistencies when new nodes or edges are introduced. The ability to handle fine-grained logical relationships within a graph, combined with the capacity to extract cross-modal knowledge, supports the development of our interactive knowledge environment.

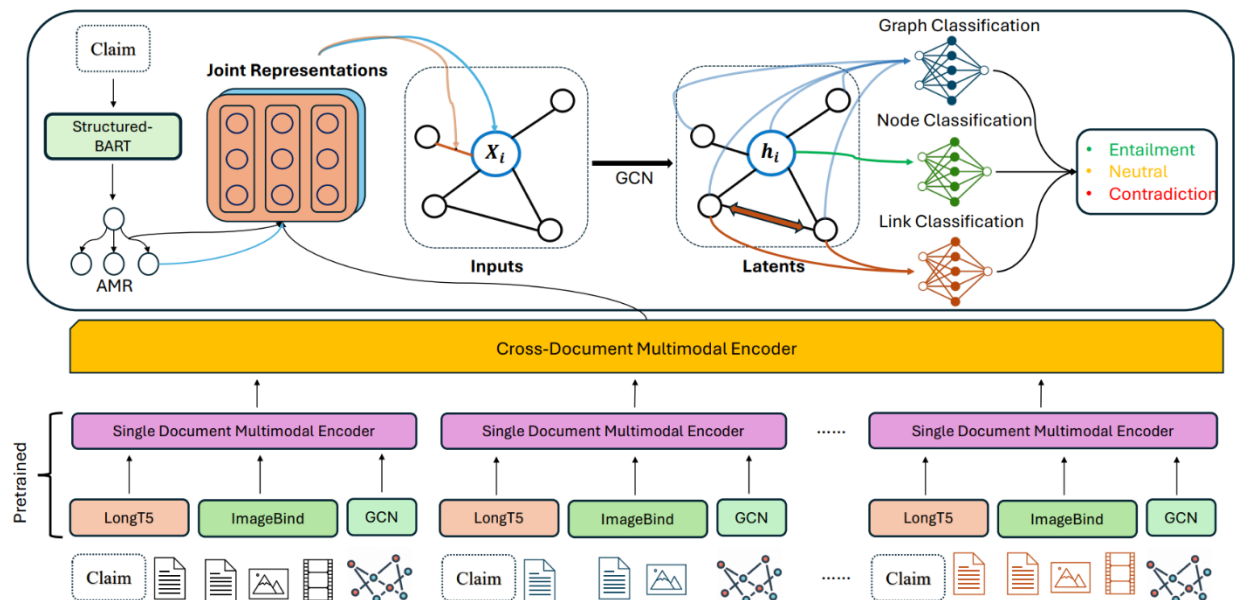


Figure 5 – Fine-grained logical consistency checking (Tang et al. 2024) Reasoning models are employed to automatically infer new knowledge, identify, and resolve contradictions, and complete missing logical links in response to user modifications

Another relevant open-source knowledge-graph capability is PyReason. The PyReason framework was developed by IKOS key person (Shakarian) and allows for reasoning on large knowledge graphs based on logic programming (Aditya et al., 2023). With the PyReason framework, an analyst can express domain knowledge in a PROLOG-like logical language (which can also be created via LLM conversion from natural language). This enables efficient identification of inconsistencies of algorithmic-created extensions with domain knowledge (Shakarian and Simari, 2022) in a manner that provides the user a detailed account as to why the algorithmic-induced inconsistency contradicts the domain knowledge. Likewise, the PyReason framework will be able to infer conclusions as combinations of domain-knowledge and algorithmically-induced knowledge. This way the analyst can gain an understanding as to what are the outcomes that a given piece of algorithmically-extended knowledge implies.

In short, there have been great advancements in knowledge graphs and knowledge graph interactions recently, and the IKOS team includes leading experts who have delivered these innovations. IKOS will build on these recent advancements to deliver on several operationally meaningful technical objectives.

2. Phase I Technical Objectives

Intelligence analysts require better capabilities for building, inspecting, editing, and sharing multimodal knowledge graphs. Working in teams, with time sensitive workflows, these analysts require intuitive mechanisms to process and leverage very large relational datasets. IKOS will enable intelligence analysts to interact with complex multimodal knowledge graphs in a low-effort, intuitive manner.

IKOS will significantly enhance the current approaches to knowledge graph interaction by enabling dynamic multimodal knowledge graph refinement. Rather than relying solely on the analyst's manual input or requiring extensive retraining of underlying models, the system will integrate user corrections, identify contradictions, and adjust the graph's fine-grained structure in real-time. As a result, analysts will be able to rapidly complete and refine large-scale, multimodal knowledge graphs without needing to manually track every update and change throughout the graph.

IKOS will **support natural language inputs**, allowing analysts to add or update information using simple textual commands or, if they so choose, by visually modifying relevant parts of the graph. *The semantic and logical effects of these updates will propagate throughout the graph structure, enabling **real-time refinement of the graph's content and schema** without extensive manual intervention.* Further, we will allow the analyst to directly input domain knowledge into the system not only in the form of graphical components (e.g., nodes, edges, and their labels) but also logical relationships among entities exhibiting certain characteristics (e.g., two members of rival groups cannot have "friend" relationship between them). IKOS will **infer new relationships** based on the analyst's modifications and entered domain knowledge, automatically deriving appropriate conclusions, and **identifying potential logical inconsistencies or conflicts** with the newly added information. For the case in which inconsistencies are identified with domain knowledge, IKOS shall provide an explainable trace to localize the source of the inconsistency in a manner understandable to the user.

IKOS will **significantly reduce the cognitive burden on analysts**, enabling them to focus on higher-order reasoning tasks instead of labor-intensive knowledge management. IKOS will provide numerous benefits for intelligence analysis applications, including **more accurate and up-to-date situational awareness, improved detection of emerging patterns**, and **improved decision-making in time-sensitive contexts**. IKOS will enable analysts to easily refine large complex knowledge structures while enhancing their completeness and consistency. IKOS knowledge structures will become an even more practical and robust tool that supports a wide range of intelligence applications. IKOS will address key limitations in current state-of-the-art systems, leading to a **new generation of dynamic, multimodal knowledge graphs** that are not only comprehensive but also agile enough to meet the evolving demands of real-world intelligence analysis environments.

In Phase 1, we will obtain baseline performance metrics for our use case with a convenience sample experienced OSINT analysts. We will then evaluate the effects of our proposed approaches to answer the following questions and select the most promising features to develop in Phase 2.

- 1) Which IKOS features (See Figure 6 below) to RESIN can be implemented to support OSINT analysis?
- 2) Which IKOS features will improve OSINT analysis involving large, complex knowledge graphs?

3. Phase I Statement of Work

- a. **Scope.** As shown in Figure 6, IKOS will build upon recent advancements in knowledge graph extraction to deliver a prototype for an interactive knowledge graph. In Phase 1, we will:
 - 1) Initiate a campaign of Mission Engineering for Open-Source Intelligence Analysis to orient the design and feasibility assessments of IKOS;
 - 2) leverage a government off-the-shelf platform, Information Operations Network (ION), to acquire realistic multimodal data for an OSINT use case;

- 3) employ an open-source knowledge extraction system (RESIN) to generate an initial knowledge graph for our analytic use case; and
- 4) enhance RESIN and RESIN-EDITOR with interactivity features to support team curation of knowledge graphs.

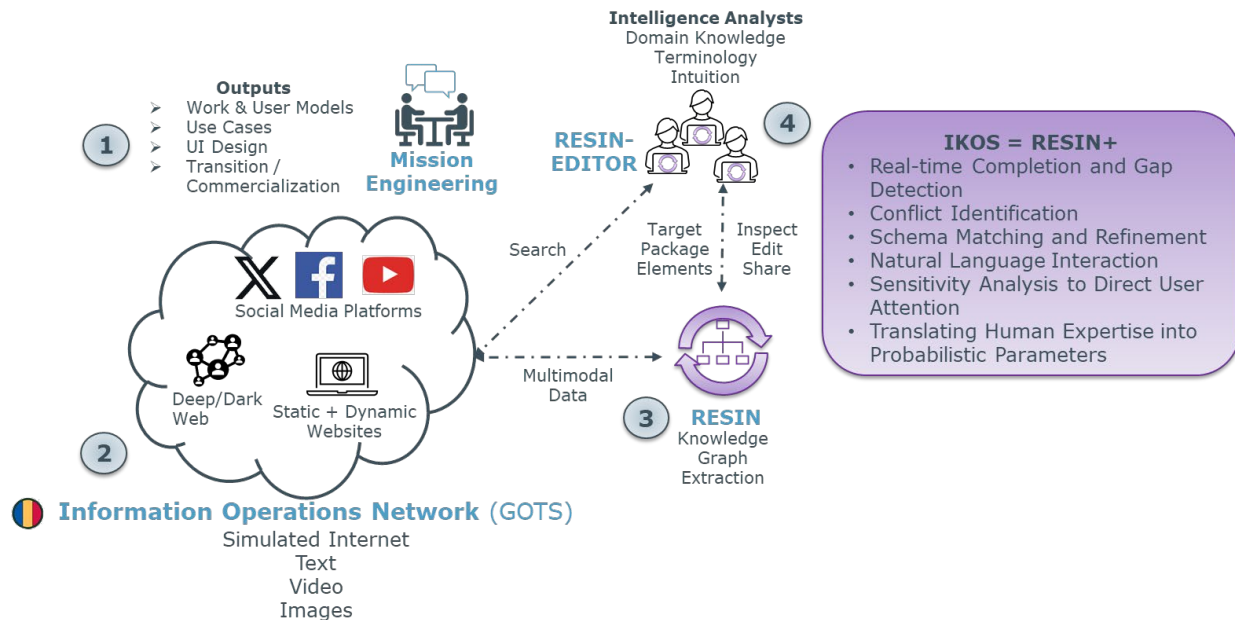


Figure 6 - IKOS Concept. We will leverage a GOTS data repository (ION) and extend upon an open-source application (RESIN) to deliver an interactive support tool for OSINT analysts.

IKOS will deliver novel support for knowledge graph interaction. We will integrate advanced capabilities that allow for automatic **fine-grained knowledge completion** and **real-time knowledge propagation** based on user inputs. We will develop new features that **automate schema matching and knowledge graph refinement**, to include the ability to suggest additional edges between related nodes, infer new relationships, highlight information gaps, and **automatically flag inconsistencies** as users interact with the graph. We will implement a **natural language interface**, allowing analysts to input modifications more intuitively, which will be automatically propagated throughout the graph, ensuring coherence and consistency across all connected data points. We will implement **sensitivity analysis that directs analysts' attention** to aspects of the graph that require their revisions. Finally, we will adapt approaches to **translating expertise into probabilistic parameters** to afford analysts the ability to introduce their expert intuition into knowledge graphs.

- b. **Task Outline.** We propose the following five tasks to deliver IKOS and set the conditions for a successful Phase 2 effort to build upon our Phase 1 prototype.

Task 1 | Mission Engineering

We will constitute an interdisciplinary Mission Engineering team, which will employ mixed ethnographic methods, to develop a deep understanding of the current and anticipated needs for knowledge graph interaction amongst open-source intelligence analysts. This campaign, as illustrated in Figure 7, will yield user experience needs, user and work models, and use cases for the new envisioned system. The IKOS field study will directly drive design specifications by generating user interface design improvements for RESIN-Editor, commercialization plans, and a system architecture to support transition.

IKOS will build on progress and relationships from AFRL RAINFLY, which is a project to deliver a cognitive-psychology informed planning tool for information operations. For RAINFLY, BRG has established relationships with open-source intelligence workgroups, such as the 123rd Intelligence Squadron and 288th Operations Support Squadron in the 188th ISR Group (Arkansas Air National Guard). Through this and other connections to

operational OSINT workgroups, we will accelerate the knowledge elicitation required to support IKOS. BRG has also established a dialogue with the Command and Control of the Information Environment (C2IE) program office. C2IE is a platform for collecting and analyzing publicly available information and is a likely transition partner for IKOS.

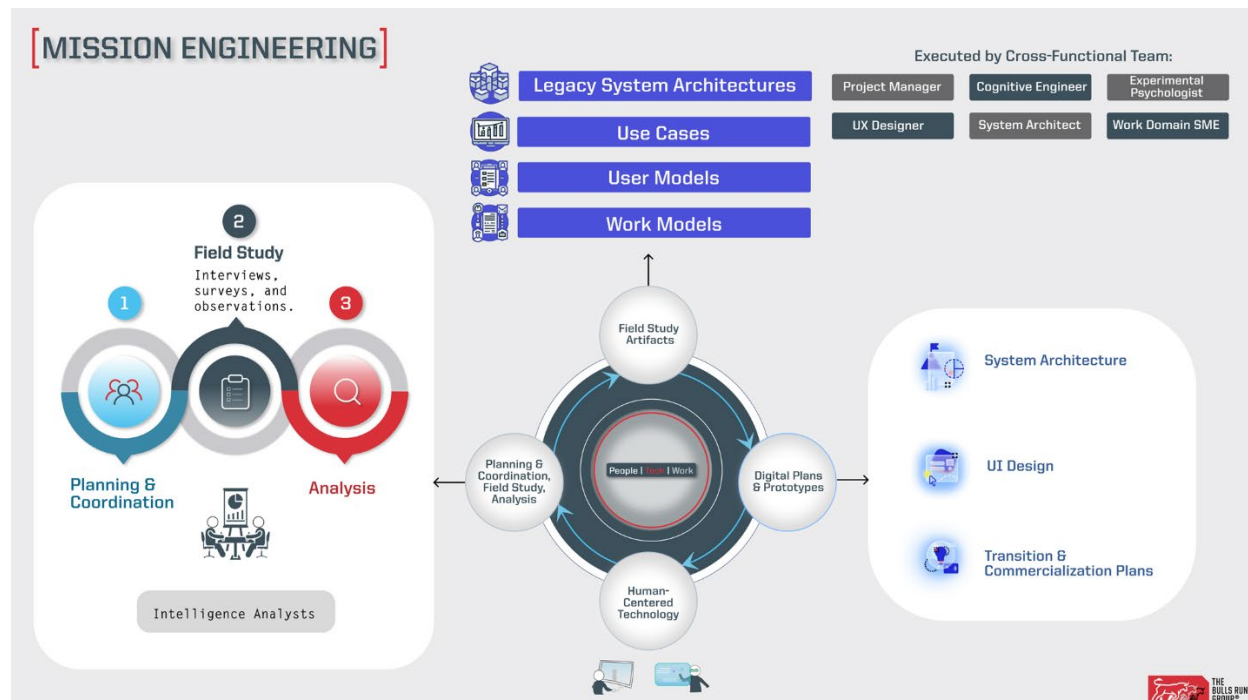


Figure 7 - Mission Engineering – a deliberate campaign engaging OSINT analysts and system designers to ensure IKOS is useable, useful, and can be transitioned to a program of record

Task 1.1 | Design UI/UX

The objective of this sub-task is to generate interactive user interface (UI) designs to enable intuitive, efficient user experiences for OSINT analysts. These designs will accelerate user experience research and inform the application development in Phase II.

We propose to extend on the RESIN-EDITOR user interface by designing visualization tools and enabling users to interactively modify knowledge graphs constructed from multimodal data. Users will be able to directly add, delete, or modify nodes and edges representing entities, events, or relationships. The interface will also include rich visualizations of the knowledge graph, with different data modalities (text, images, videos, etc.) linked to specific nodes. This will provide users with an integrated view of the data, ensuring that all modalities are represented and connected in a single knowledge structure.

In Phase I we will develop mock-ups and high-level interface designs to support the development of clickable wireframes showing the functionality of the envisioned system for the scoped set of targeted users, with special attention to accessibility options, including color blind. We will employ a user-centered design (UCD) approach (ISO 13407, 1999; Greenbaum & Kyng, 1991; Schuler & Namioka, 1993; Beyer & Holtzblatt, 1998), grounding the process in user needs, expectations, and limitations, to focus on the design of a complete end-to-end workflow. The UCD approach is founded on an iterative cycle of design, user feedback, and revision that not only requires designers to analyze and envision the way users are likely to use a product or service, but also to validate their assumptions with actual users.

We will also leverage recent work from the human-computer interaction (HCI), computer supported cooperative work (CSCW), and intelligent user interfaces (IUI) research communities, focusing on theories, frameworks, and design considerations for supporting collaborative sensemaking and analysis, including crowdsourcing, collective intelligence, and human-AI teaming approaches.

Task 1.2 | Design System Architecture

This sub-task will entail the design of the system architecture to afford scalability, reliability, and security. We will review RESIN and C2IE architectures to better understand how IKOS will need to operate if transitioned to C2IE. In parallel, we will design the architecture with requirements elicited through field study. For the back end, we envision a microservices architecture that allows for components to be developed, tested, and integrated independently. IKOS will incorporate a scalable and secure reference architecture for front-end UI application development and AI/ML services development. The initial architecture will serve as the baseline to accommodate the design and deployment of future workflows. This design will leverage the latest in compute, storage, transport, and datastore configurations — suitable for current and future AI development. Using AWS or other cloud providers, we will create a modular open-system architecture to support rapid prototyping resource-intensive AI services.

Task 1.3 | Transition and Commercialization Planning

This sub-task will set the conditions for transition and commercialization. As noted previously, we will extend on our previous work for AFRL RAINFLY to determine a transition pathway for IKOS into C2IE. This will include coordinating with C2IE requirements owners (A35), to document future IKOS capabilities as requirements for the C2IE roadmap. This will help secure funding in future years for system integration. We also anticipate IKOS will have value in non-defense industries, such as business intelligence. Business intelligence is an increasingly lucrative form of open-source intelligence. Current platforms only aggregate, sort, and display data from a variety of feeds. None incorporate knowledge graph structures to assist with more nuanced analysis. We will investigate the funding and technical pathways to deliver IKOS as a licensed feature or service within one or more existing platforms and/or developing a new business intelligence platform with interactive knowledge graphs as a core value proposition. The outputs of this sub-task will be a comprehensive, phased transition and/or commercialization plan to be pursued in Phase 2.

Task 2 | Multimodal Data Acquisition

We will leverage the Information Operations Network (ION), an Army Training and Doctrine Command (TRADOC) G-2 platform for realistic data for the development and testing of IKOS. ION is a simulated internet with replicated social media and other websites including deep and dark web. ION has provided analysts with realistic multi-modal data in over 400 intelligence and information operations exercises in the last two years (including 47 USAF events). As such, it is an ecologically valid sandbox of representative data with an ecosystem of intelligence and information operations users. This platform supports simple and complex challenge problems for individuals and teams of multiple open-source intelligence analysts. The ION program office pushes regular updates to this platform to include new data sets and has agreed participated in previous AFRL and ARL research programs. We will coordinate for free access to an unclassified instance of ION in Month 1 of this project.

Task 3 | Generate Initial Knowledge Graph

We will leverage the core components of the RESIN system (Wen et al. 2021), including relation prediction, event extraction, and argument extraction, cross-media fusion, etc. as the foundation for our knowledge graph extraction. Where newer methods—such as enhanced entity recognition, coreference resolution, and cross-modal relation prediction—offer better performance, we will integrate such components from our recent multimedia knowledge graph extraction system (Tang et al. 2024) to improve overall knowledge graph quality and minimize erroneous extractions. We will also allow for the input of user-created domain knowledge via PyReason – residing as a layer on top of RESIN to allow for explainable identification and localization of inconsistencies with respect to domain knowledge.

The textual pipeline processes both text-based media sources and transcribed audio. It starts with information extraction to extract entities, relations, and events. Next, cross-document entity and event coreference resolution links related entities and events across sources, creating a cohesive knowledge representation. This is further enriched by predicting higher-level semantic relations like temporal and causal links. For audio data, we use Qwen-Audio (Chu et al. 2023) to extract relevant

knowledge from audio signals, along with Whisper (Radford et al. 2023) to transcribe speech into text, ensuring complete integration of spoken information into the knowledge graph. For images and video, we use a state-of-the-art multimodal large language model (Ye et al. 2024) to create grounded event structures. Visual content is first converted into detailed captions, which are then transformed into dense knowledge structures that include objects, events, and relationships. Cross-modal coreference resolution and relationship prediction then integrate this visual data with the text-based knowledge graph to form a comprehensive multimodal graph. Once the multimodal graph is constructed, we use a hierarchical schema repository to align extracted information with matching schemas. This guides the final integration and structuring of the multimodal knowledge graph. See Figure 2 for an illustration of the knowledge graph extraction process in RESIN.

RESIN is fully Dockerized, publicly accessible, and readily deployable on day one. By the end of this task, we will generate comprehensive semantic multimodal knowledge graphs from unstructured intelligence media assets which will then be refined through human interaction in later tasks.

Task 4 | Develop Interactive Knowledge Graph Features

This task represents the core of the research to be performed in IKOS. In this task, we will build upon the open-source code bases for RESIN, RESIN-EDITOR, and PyReason to deliver a minimally functional version of IKOS that supports an OSINT use case. Figure 8 illustrates the six features that will 1) improve graph refinement, 2) afford natural language interaction, and 3) improve usability. The following sub-tasks describe how each of these features will be implemented.

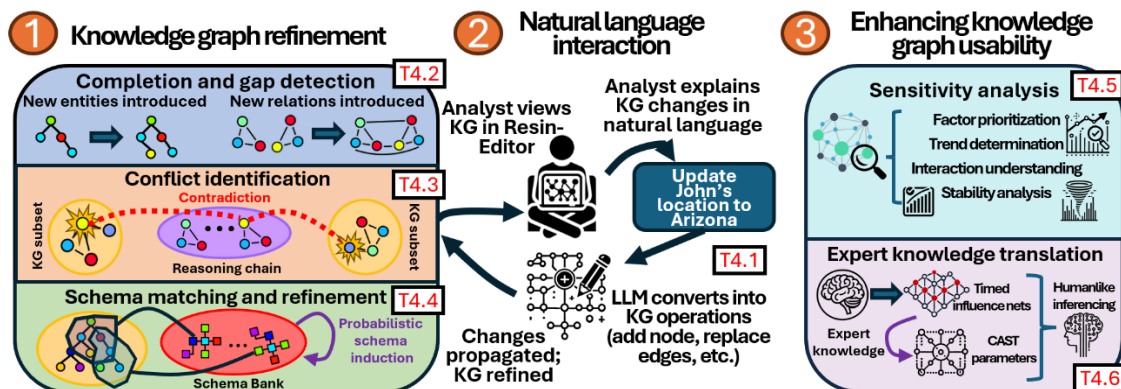


Figure 8- IKOS improvements upon RESIN

Task 4.1 | Implement natural language interaction

In this sub-task, we will develop two forms of user interaction. First, users will be able to directly query the knowledge graph, view the results, make manual edits, and augment with analyst domain knowledge. While this provides full control, manually finding and editing relevant nodes and edges can be time-consuming and cumbersome, especially when analysts need to navigate large, complex graphs. Therefore, we also propose to implement a second, more intuitive form of interaction leveraging natural language-based graph editing.

Users will input free-form text to express their intended modifications to the graph, and our system will automatically translate these instructions into corresponding graph operations. For example, a user could enter a command such as: "Update the location of Agent X to Berlin and note that they are now carrying a briefcase." This input will be processed using a state-of-the-art large language model (LLM) such as Qwen-2.5 (Yang et al., 2024) or LLaMA 3.1 (Dubey et al., 2024) which will parse the text and translate it into a series of structured actions, such as: 1) Modify Node: Update the 'location' attribute of the node representing 'Agent X' to 'Berlin'; 2) Add Attribute: Add a new attribute 'carrying' with the value 'briefcase' to the node representing 'Agent X'; etc. These proposed changes to the graph in a visual form

will be presented back to the user for confirmation before they are applied to the knowledge graph, allowing the user to review and refine the actions if necessary. The system will also allow for further clarification in natural language to allow the user to further fine-tune the changes to be made.

Technically, integration of the LLM with the knowledge graph system will involve several steps. First, the LLM will process the user's input to identify intent, entities, attributes, and relationships mentioned. Extracted elements will be mapped to existing nodes and edges within the knowledge graph, considering coreferential edges within the graph. Next, the LLM will generate a sequence of graph operations that reflect the user's intent, following the schema and constraints of the knowledge graph. Finally, before executing the actions, the system will display the interpreted operations to the user for validation.

Our proposed natural language interface will provide a more efficient and user-friendly way to interact with the knowledge graph, especially in scenarios where manually navigating and editing the graph is impractical. By providing two complementary methods for interaction—direct manipulation and natural language input—our system will allow analysts to choose the interaction mode that best suits their needs. Moreover, our design positions us to test and refine methods that can significantly reduce the time required for analytic tasks in later program phases. Collectively, these robust interaction capabilities will help ensure that our knowledge graph remains a living, up-to-date resource that can rapidly evolve with analyst input and operational needs while minimizing analyst interaction overhead and cognitive burden.

Task 4.2 | Enable real-time completion and gap detection

The information captured within a knowledge graph, whether extracted from media sources or input by analysts, may not always encompass all logical relationships or conclusions that can be derived from the available data. In such cases, additional nodes and edges may be inferred to enhance the graph's semantics and to provide higher level conclusions derived from the data which may be useful to analysts. To address this, we will apply our method for inferring new logical connections for automated knowledge graph completion. This method will derive new conclusions based on reasoning over the existing graph which are then introduced as new nodes or relationships. These updates will be triggered whenever new information is input into the graph or when users make modifications which will ensure that the knowledge graph evolves dynamically and intelligently in response to user input. Once new information is derived and integrated into the graph, we will perform logical consistency checks as outlined earlier. This process will help ensure that the newly introduced relationships maintain the semantic and logical integrity of the entire knowledge structure and prevent contradictions or inconsistencies from arising.

On a technical level, our approach leverages our existing model capable of generating logical entailments from a set of multimodal knowledge. Our system will accept a claim graph alongside a target entailment label (e.g., entailment, contradiction, or neutrality). For Phase 1, we will retrofit our fine-grained generation module to support knowledge graph completion. Specifically, once a graph is constructed or edited by a user, we will aggregate relevant parts of the graph and feed them into our generation module. The module, trained with a logical entailment critic, will then infer logically entailed relationships that are not yet present in the graph. These new semantic conclusions, derived from existing evidence, will be automatically integrated into the graph as additional nodes or edges. After these updates, our consistency checking mechanism will be applied to verify the validity of the new conclusions, ensuring that they do not introduce contradictions or undermine the coherence of the graph. In Phase I, we will focus on retrofitting our existing generator specifically for knowledge graph completion to measure the current abilities of the system to derive and integrate logically consistent relationships and conclusions into the knowledge graph.

Additionally, we will extend these capabilities to address information gaps within the graph. Specifically, we will explore multiple methods to detect where the knowledge graph lacks sufficient evidential support. One approach will involve prompting our generation module to produce neutral logical evidence with respect to the graph. Portions of the graph for which the model predicts "neutral" entailments—where no strong logical support or contradiction is available—will be flagged as potential

information gaps. These regions suggest that the graph lacks sufficient evidence for certain topics or relationships, and they will be highlighted for further review by analysts. This feature will help analysts identify where more data collection or refinement is necessary which will help improve both the completeness and reliability of the knowledge graph.

Task 4.4. | Enable conflict identification

To ensure the knowledge graph remains coherent and reliable as users interact with and modify it, we propose to perform fine-grained logical entailment during interaction to check for introduced inconsistencies. While existing knowledge graphs often include relations such as co-reference, causal links, and hierarchical structures, they typically lack the logical relations that capture how individual facts combine to entail or contradict other facts. This gap limits the graph's ability to support advanced reasoning and maintain consistency, especially when new information is added or existing information is altered. Our fine-grained logical entailment system will support fine-grained consistency checking of a claim knowledge graph. We will apply this technique in Phase 1 to detect logical inconsistencies introduced as a result of edits or from conflicting knowledge sources.

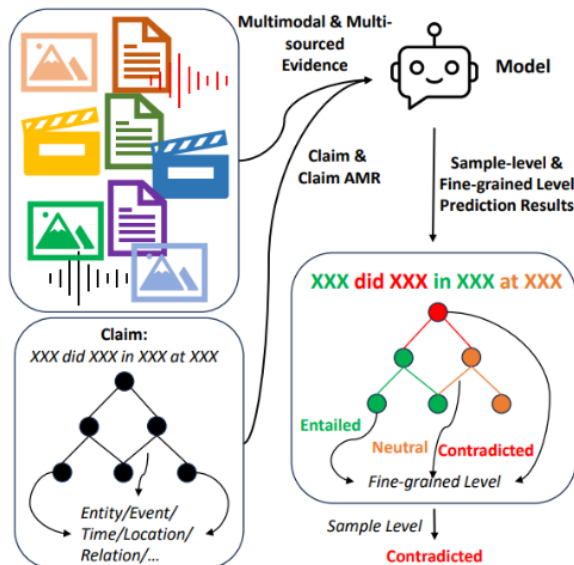


Figure 9 - Logical inconsistency and conflict identification

As shown in Figure 9, we will extend our fine-grained multimodal logical entailment mechanism (Tang et al., 2024) to operate internally within the knowledge graph. Our technique will allow the system to infer logical implications and detect contradictions resulting from user modifications. Specifically, when a user updates the graph—by adding, deleting, or modifying nodes and edges—we will identify relevant nodes not only through direct connections but also using retrieval techniques that operate over the entire graph's structure. Our entailment mechanism will then assess these nodes to determine logical relationships such as entailment, contradiction, or neutrality. This approach goes beyond simple surface-level checks or propagating changes through co-referential edges.

Traditional methods might only flag direct inconsistencies, like conflicting attributes of a single entity. In contrast, our technique will uncover deeper logical inconsistencies that may not be immediately apparent. By considering how new and existing pieces of information logically combine, our proposed system will detect contradictions that emerge from indirect relationships or inferred knowledge. Further, our extensions allowing for the addition of user-specified domain knowledge (using PyReason) will allow rapid identification and localization of inconsistencies among such knowledge, the elements of the graph, and other algorithmic-induced changes (Figure 10). For example, in our recent work (Bavikadi, Lee, et al., 2024) we were able to provide human-understandable explanation as to why our PyReason-based

systems produced likely future locations of suspicious maritime vessels.

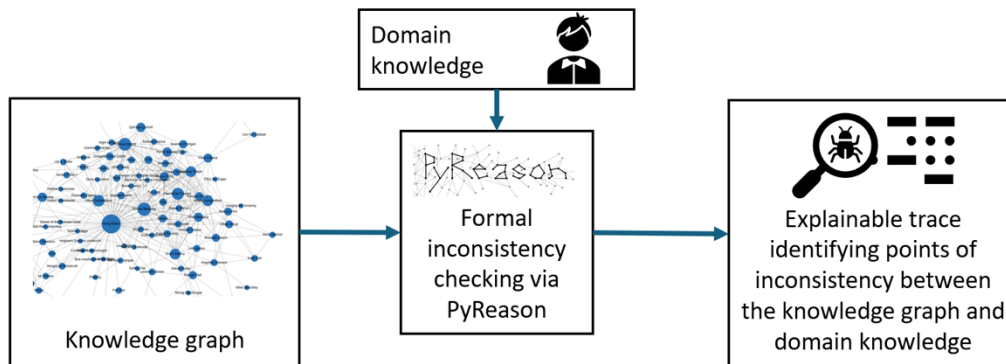


Figure 10 - PyReason framework supporting inconsistency checks

Consider a scenario within the knowledge graph involving multimodal data: Fact A (Text Data): "Agent Smith participated in a covert operation in Tokyo on September 12th, 2024." Fact B (Image Data): A photograph tagged with metadata indicating it was taken in Rio de Janeiro on September 12th, 2024, showing Agent Smith attending a diplomatic event. Fact C (Existing Knowledge): "There are no direct flights between Tokyo and Rio de Janeiro that take less than 24 hours." Each fact resides in a different modality and seems plausible individually. Moreover, no direct conflict arises from any of these facts. However, inferentially one can easily see that from Fact C, it's impossible for Agent Smith to travel from Tokyo to Rio de Janeiro within a 24-hour period resulting in a contradiction. These types of fine-grained, semantically driven inconsistency judgments are not supported by current knowledge graph techniques.

For Phase 1, to achieve this level of sophisticated inconsistency detection, our system will extend the existing claim-based entailment mechanisms to operate across the entire knowledge graph. We will use our entailment mechanism to evaluate how the new information interacts with existing data, inferring potential logical relationships (See Figure 11). Our system will highlight the detected logical inconsistency for the analyst's attention, providing contextual information such as the conflicting facts and the impossibility of rapid travel between the two locations. By identifying these inconsistencies, the system can suggest additional possible changes to the graph based on the analysts' input.

Our method will help ensure that as users interact with the graph—making necessary changes to reflect new intelligence or correct AI/ML outputs—the knowledge structure maintains a high level of consistency and trustworthiness. Our goal is to ensure that analysts can rely on the graph to provide accurate and logically coherent information. Moreover, this logical consistency checking reduces the cognitive load on analysts by automating the detection of inconsistencies that might otherwise require extensive manual review. It allows them to focus on higher-level analysis rather than verifying the basic accuracy of the data. This capability is particularly valuable in operational settings where information is rapidly evolving, and decisions must be made quickly based on the most accurate and coherent data available.

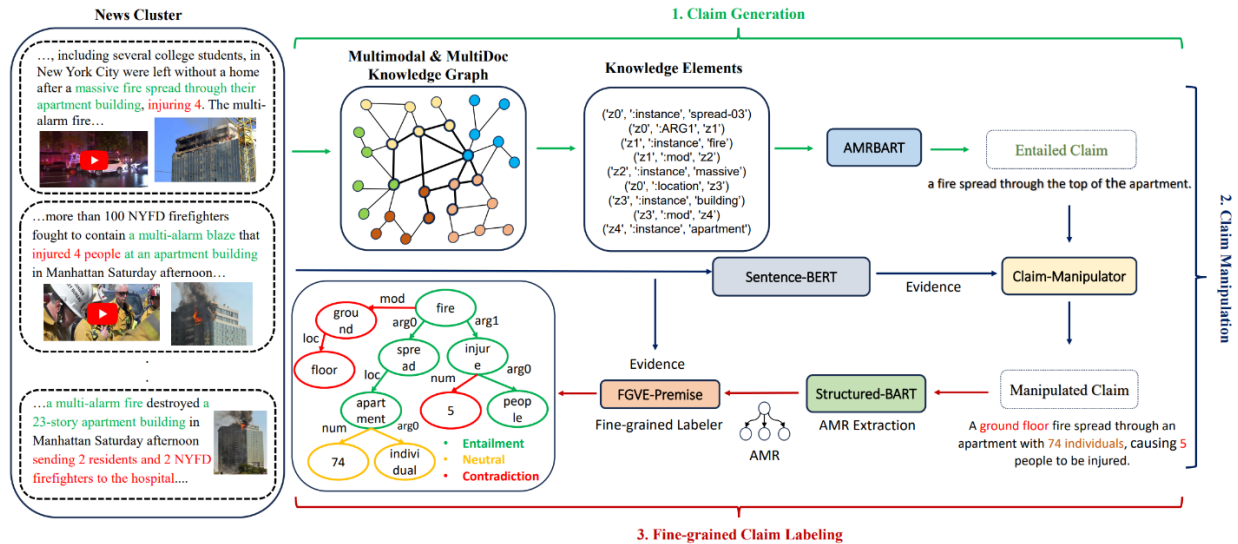


Figure 11 – Generating fine-grained relationships in knowledge graphs.

Task 4.4 | Enable graph schema updating

A key aspect of maintaining an effective and adaptable knowledge graph is ensuring that its underlying schema evolves as new data and user modifications are integrated. In this proposal, we will implement several schema updating mechanisms to ensure that the knowledge graph's ontology remains coherent and optimized for the evolving data structure. To achieve this, we propose two complementary approaches that leverage state-of-the-art schema induction techniques (Jin et al., 2022; Zhang et al., 2023; Du et al., 2022; Hao et al., 2023), each designed to automatically refine and adapt the schema based on user edits and new data.

Zero-Shot Schema Induction for Dynamic Schema Updating

When users make edits to the knowledge graph, we will trigger an automated schema update process. Specifically, our system will leverage zero-shot schema induction techniques (Dror et al., 2023; Zhang et al., 2023; Li et al., 2023) on the edited portions of the graph. Zero-shot schema induction allows us to generate new schema representations without requiring pre-labeled data to adapt the ontology to reflect the new terms used in the edits or modified knowledge. Once the zero-shot schema is induced, we will perform schema matching between the newly induced schema and the existing knowledge graph structure. This process ensures that any modifications introduced by the user are reflected in the most accurate and semantically aligned schema possible. If the newly induced schema provides a better structural fit to the knowledge graph, as determined by semantic similarity metrics and schema matching algorithms (Wang et al., 2022), the system will automatically replace the existing graph schema with the refined version.

Probabilistic Schema Induction and Divergence-Based Schema Refinement

In parallel, we will implement a probabilistic schema induction technique that generates schema representations based on the statistical analysis of a set of documents or existing intelligence-based knowledge graphs. This probabilistic approach generates schema with derived probabilities, reflecting how likely a particular schema structure represents the underlying data. As users interact with the graph and make edits, we will continuously monitor how the edited graph aligns with the originally induced schema.

To ensure that the graph's structure remains semantically valid, we will perform lifelong schema matching by matching the edited graph with the induced probabilistic schema. If the probability of the current graph's alignment with the original schema falls below a predefined threshold, indicating significant divergence from the original schema, the system will automatically re-initiate the schema

induction process. Depending on the size and complexity of the graph, the system will either employ zero-shot schema induction for smaller graphs or probabilistic schema induction for larger, more complex graphs or those with enough data to meaningfully derive a probabilistic schema. These techniques will allow the graph's schema to evolve in response to knowledge graph interaction without direct user intervention.

Task 4.5. | Sensitivity Analysis to Direct User Attention

One major challenge for human-machine interaction in knowledge graph development and refinement is identifying the aspects of the model most in need of human inspection or expertise. Manually browsing a large knowledge graph to search for areas that might be improved, edited, or updated is intractable for human analysts and a poor use of their time and expertise. In IKOS, we will develop a sensitivity and decision analysis microservice that will use proven, quantitative techniques to direct user attention to aspects of the model that would most benefit from refinement.

Sensitivity analysis and decision analysis techniques are often used to identify the most “important” features of a model and highlight the impact of variables or the stability of outcomes in a complex system. When used with decision support tools, sensitivity analysis provides a mathematically rigorous approach to model introspection and explanation, increasing transparency, enhancing interpretability, and increasing awareness of the model behavior (Guidotti et al. 2018, Saltelli 2019).

Applying sensitivity analysis techniques, especially to a complex graphical model, enables

- (1) increased awareness of model behavior;
- (2) understanding of model limitations;
- (3) understanding of areas requiring additional modeling efforts;
- (4) redirection of data collection to more important or impactful features;
- (5) proper communication of results; and
- (6) proper assessment of outcomes and effects.

We propose to adapt these tools to RESIN to help direct user attention to aspects of the model that might need updating or enhancement via human expertise or human-directed data collection. Table 1 describes the sensitivity analysis techniques that we will incorporate in IKOS.

Table 1 - Sensitivity Analysis Techniques for IKOS

Technique	Purpose
Factor-prioritization	<ul style="list-style-type: none"> determine which features drive the results of a model prediction, and whether it is worth it to collect further data identify factors that may cause the predicted outcomes or optimal COA to change
Trend determination	<ul style="list-style-type: none"> determine whether the model responds positively or negatively to input variations and can enable comparison of model results under different starting conditions model debugging (e.g., is the response coherent with intuition? Is there incorrect or missing information?) managerial action (if a variable has a positive impact, action must be taken to guarantee the variable does not decrease) policy implementation
Understanding interactions	<ul style="list-style-type: none"> determine whether two or more variables produce synergistic or antagonistic effects
Understanding stability	<ul style="list-style-type: none"> determine whether outcome is stable if inputs vary determine if a preferred strategy or observed effect is robust to variations in external factors
Frequency analysis	<ul style="list-style-type: none"> assess frequency with which each COA or outcome is selected as optimal through objective optimization analysis

Expected Value of Perfect Information	<ul style="list-style-type: none"> quantify the degree to which the model results may be improved by gathering additional targeted data on specific variables or parameters
Generalized Tornado Analysis	<ul style="list-style-type: none"> assess the interaction effects of combinations of model variables and identifying which ones may jointly require attention to ensure model effectiveness

Figure 12 illustrates the output of a Generalized Tornado Analysis on a knowledge graph representing the socioeconomics of agricultural production in Gambella, Ethiopia. This analysis indicates for each node in the graph the individual effects of the variable (i.e., how “influential” is this variable in terms of the changes observed in predictive outcomes or optimal courses of action due to changes in the variable value), the interaction effects (i.e., how influential is this variable in concert with other variables), and the total effects. The graph is sorted by total effect, with the most influential variables at the top. Such analysis can help direct users to key aspects of the model to provide expertise and refinement. For example, if the variables identified as most/least influential do not comport with the expert knowledge of the human analyst, then this indicates a case for perhaps updating and modifying model parameters. Alternatively, highly influential variables may indicate places where it is more critical to have accurate and sufficient data, indicating new data collection efforts may be needed.

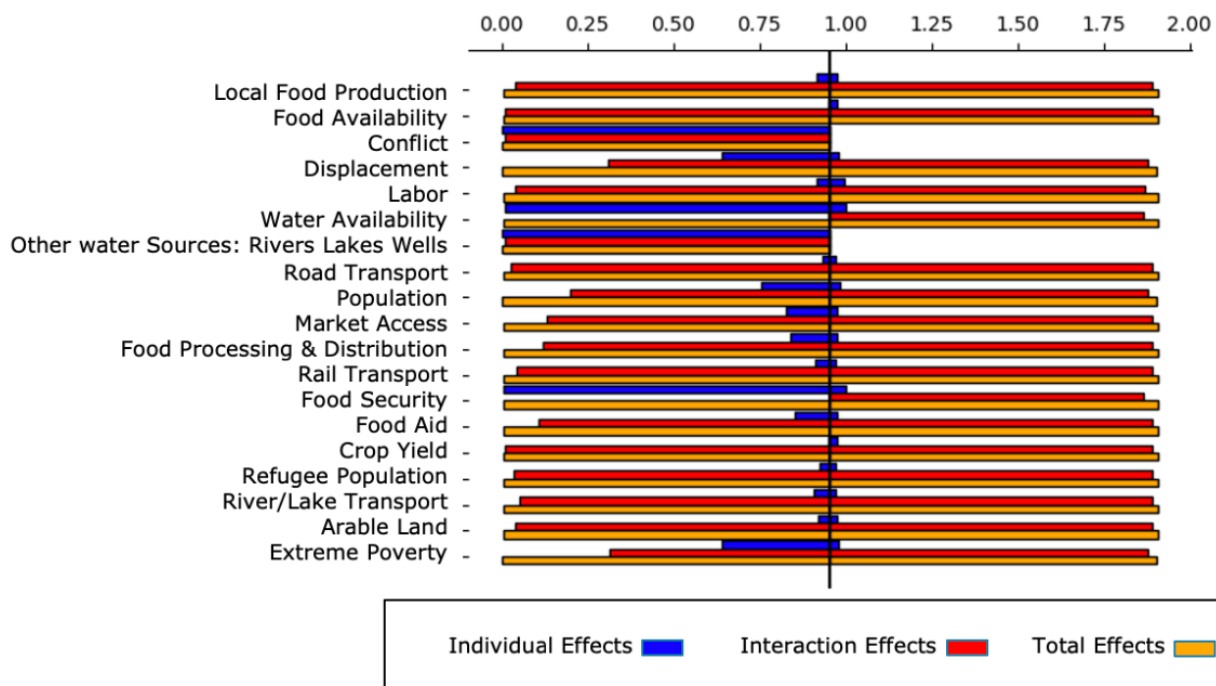


Figure 12 – Example Output of Generalized Tornado Analysis

In the IKOS Phase I, we will develop prototype implementations of sensitivity analysis techniques to assess the feasibility of applying these techniques to knowledge graphs in the RESIN framework to identify key aspects for human expert attention.

Task 4.6. | Translating Human Expertise into Probabilistic Model Parameters

Knowledge graphs can provide representational richness in terms of relationships between entities and events, but for these frameworks to achieve their potential as decision support tools, they must also be computational, that is they must be able to perform some inference to answer key questions that may be posed by analysts. Such questions may be in the form of hypothetical “what if?” scenarios to determine likely outcomes, identifying the optimal course of action (COA) for a real or hypothetical

scenario, or identifying factors most influential in achieving a desired outcome. A wide variety of modeling frameworks exist to add computational capabilities into graphical knowledge representations to address the issues of stochastic relationships and influences—probabilistic graphical models (e.g., Bayesian networks (Pearl, 1988), influence diagrams (Howard and Matheson, 2005), probabilistic relational models, (Getoor, et al., 2007)), probabilistic soft logic (Kimming et al., 2012), Markov logic networks (Richardson and Domingos, 2006), and many others. However, humans, regardless of the depth of their domain expertise, are notoriously bad at estimating probabilities, and even worse at estimating the complex probabilistic relationships and interactions required by these computational modeling frameworks.

To address this issue, we propose adapting an approach originally developed for timed influence nets (TIN) (Haider and Levis, 2008) a probabilistic graphical model that is a variation on a standard influence diagram (Howard and Matheson, 2005) that captures probabilistic and uncertain influences in dynamic environments. TINs were originally developed for use in intelligence analysis as a means of providing analysts with an intuitive knowledge representation that could also be used computationally for prediction and inference. Like classic influence diagrams, a TIN contains different types of nodes—those that represent random variables and those that represent decisions—and can be transformed to a Bayesian network (Pearl, 1988) for analysis and inference. However, in TINs, the model parameters (i.e., the influence parameters) are described in the Causal Strength (CAST) (Haider & Levis, 2008) format indicating the conditional influence of the parent nodes on their children. (An example TIN is shown in Figure 13.) The CAST representation was originally developed for TINs to provide domain experts with a more intuitive method for specifying probabilistic influence relationships, requiring fewer parameters than standard Bayesian conditional probability tables (CPTs). In CAST:

- A base probability is specified for each node, indicating the independent, unconditional likelihood of an event.
- Each edge contains two influence values indicating the strength of the influence when the parent is true or false
- Using these values, the CAST algorithm computes the remaining CPT parameters, enabling the TIN to be interpreted as a fully specified Bayesian network or influence diagram.

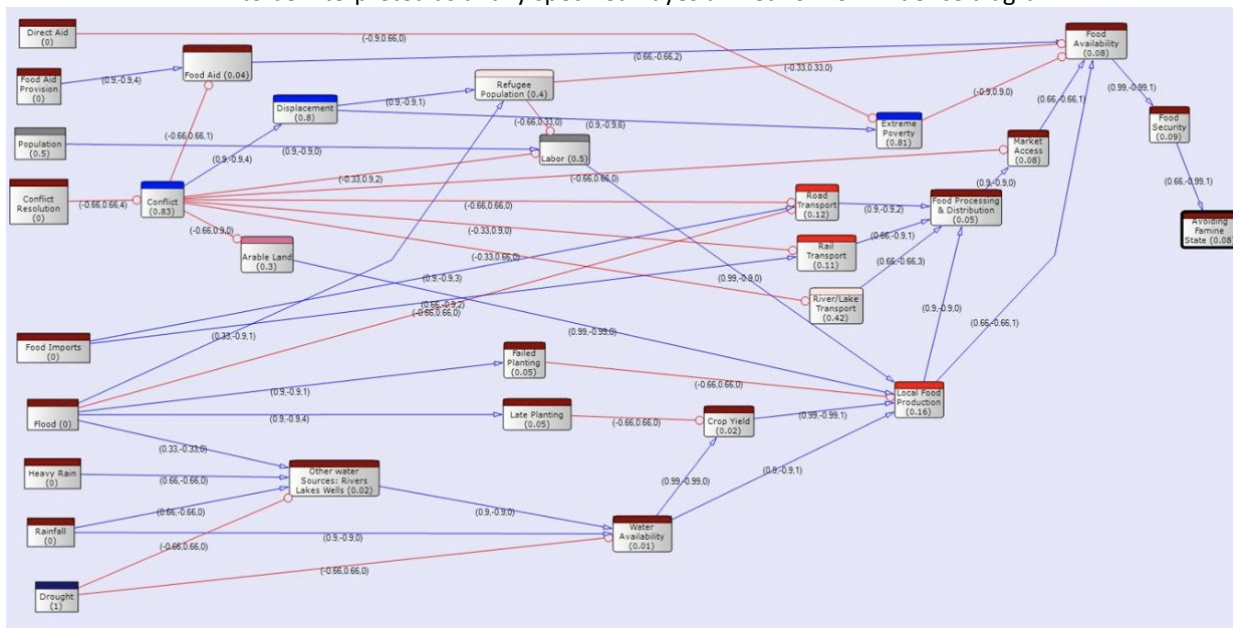


Figure 13 – Timed Influence Network (TIN) of the agroeconomic dynamics in Gambella, Ethiopia. Base probabilities are specified by food security experts.

Using this intuitive TIN framework, it is possible for users to modify not only the schema and relationships in a complex knowledge graph, but also apply their expertise to the initialization or tuning of

quantitative parameters that enable using the knowledge graph as a computational model for purposes of predictive or explanatory analytics.

The output of the TIN will work as part of our overall strategy for augmenting graphical information with expert domain knowledge using PyReason. In short, we can use a TIN to provide an output of a posterior probability that can be then directly represented in PyReason and used in downstream inference tasks on RESIN. Further, the ability to identify inconsistencies of such knowledge with the graphical constructs in the resin system can allow feedback to the user to adjust causal strength parameters when attempting to rectify inconsistencies.

In Phase I, we will demonstrate the feasibility of representing a knowledge graph computationally as a TIN and leveraging the output for downstream analytic modeling tasks using PyReason and the RESIN system.

Task 5 | Project Management

BRG's agile project management approach will be used to organize the division of labor, monitor, and manage technical implementation tasks, and set clear timelines and benchmarks for IKOS. The IKOS team will employ proven methods for distributed work and agile software development. The PI and PM are former military officers with over 20 years' experience with planning and directing team projects. The key persons on IKOS have extensive experience leading research and technology programs. The project manager has a PMP certification and has experience with managing agile software development projects. The PI or PM will organize weekly team meetings via Zoom or Microsoft Teams in addition to the bi-monthly meetings with the Sponsors. Documents, tasking, data, and code will be shared through Teams and Git. The deliverables from this task include monthly financial reports and progress reports as described in paragraph d below.

c. Milestone Schedule

The Schedule for the proposed project is shown below in Table 2.

Table 2 - Project or Milestone Schedule

Tasks	M1	M2	M3	M4	M5	M6	Objectives
1: Mission Engineering	X	X	X	X	X	X	<ul style="list-style-type: none"> • User and workflow models • Use cases • Performance measures • Transition/Commercialization Plan
2: Data Acquisition	X	X					<ul style="list-style-type: none"> • Access to ION data
3: Generate Initial Knowledge Graph		X					<ul style="list-style-type: none"> • Initial multi-modal KG for use by OSINT analysts
4: Develop Interactive Knowledge Graph Features	X	X	X	X	X	X	<ul style="list-style-type: none"> • Feasibility assessment of our approaches to allow inspecting, editing, sharing, and consistency checking of KGs
5: Project Management	X	X	X	X	X	X	<ul style="list-style-type: none"> • Project reports and Phase 2 proposal plan

d. Deliverables

We will schedule a kickoff briefing within 30 days after contract award to review project objectives, activities, and milestones and discuss feedback with the Sponsor. The deliverable for this effort will be a final report of background, anticipated functionality, technical challenges for software development and implementation, and recommendations for prototype development.

Deliverable	Date
Refined execution plan	Kickoff meeting
Progress Reports	Month 2, 4
Final Report with SF 298	Month 6

4. Related Work

Founded in 2020, BRG is a human-machine systems company, specializing in designing, planning, and developing trustworthy, fieldable systems for high-risk domains and high-performance organizations. BRG comprises experts from across military operations, engineering, data science, and cybersecurity domains who understand the “Full Capability Stack”, which integrates hardware, software, and people as a cognitive work system. The team proposed for this project brings extensive individual experience. Specific related projects of key persons on this project are shown Table 3.

Table 3 - Past related projects

Project Name	Sponsor	Contract	POC	PoP	Description
TechSuite	USAF SBIR/STTR Office (Subk to Mile 2)	FA8649-19-9-9038	J. Graley	Aug – Oct 2020	Developing a cloud compute environment to enable AI in support of SBIR proposals, subcontract to Mile Two.
RAINFLY (Phase 1)	AFRL (Subk to Mile 2)	GSA 47QFLA23D0001/TO 47QFLA23F0061	E. Thompson	Oct 2023 – Sep 2024	Employ Mission Engineering to design a system to support Information Warfare through the application of human performance findings
ReSCIND (Phase 1)	IARPA (Subk to Raytheon)	N66001-24-C-4504	K. Ferguson-Walter	Jan 2024 – July 2025	Serve as Lead Systems Integrator to deliver a novel, containerized cybersecurity application. Employ Mission Engineering to inform system design and development and testing experiments. Execute a campaign of human subjects experiments to provide data for model training and system testing.
SemaFor	DARPA	HR001120C0123	W. Corvey	July 2020 – Jan 2025	Represented Columbia developing semantic defenses for detecting generated media. Developed a multimodal knowledge graph-based generated content detection system and components for fine-grained logical reasoning over multimedia.
AIDA	DARPA	FA8750-18-2-0014	B. Onyshkevych	Jan 2018 – Jan 2023	Contributed to GAIA knowledge extraction system by developing methods for jointly extracting knowledge from videos and articles and extended system for downstream applications.
KAIROS	DARPA	FA8750 19-2-1004	B. Onyshkevych	July 2019 – Dec 2023	Contributed components for multimedia knowledge graph extraction, such as cross-modal event relation prediction, multimedia information extraction, and cross-modal temporal multimodal alignment for knowledge graph construction.
MCAI	DARPA	Cooperative agreement HR00112420370	N. Bastian	July 2024 – June 2025	Augmenting the PyReason framework to support ISR-reasoning task.
Reasoning about Global Supply Networks	ONR	N00014-23-1-2580	D. Phillips	May 2023 – May 2026	Augmenting PyReason for reasoning about supply networks and maritime vessel movement.

The IKOS Team has extensive experience using and extending RESIN multimodal knowledge graph extraction systems. Dr. Thomas has worked on several components that were integrated into the system. Specifically, Dr. Thomas has worked on enhancing the ability of RESIN systems to extract cross-modal event

structures from video and news articles with limited supervision. Together with fellow researchers, Dr. Thomas created models which can jointly extract events from video and articles (Chen et al., 2021). The model developed from this research was integrated into the RESIN system (Du et al., 2022). In follow-up work, he developed enhanced video-text alignment models that rely on temporal constraints to achieve better alignment over the prior approach to enhance the ability of the system to learn the connections between modalities without strong supervision (Chen et al., 2022). Dr. Thomas has also worked as part of the DARPA KAIROS program on learning new types of cross-modal relationships to enhance knowledge graph's representations. Dr. Thomas and collaborators developed a multimodal relationship prediction model to detect indirect relationships between text and visual content, whereas prior systems only handled cases where the visual and textual content were directly aligned (Ayyubi et al., 2022). Dr. Thomas has also deployed the knowledge graph extraction systems described for downstream reasoning applications. Dr. Thomas and colleagues developed the InfoSurgeon fine-grained knowledge inconsistency detection system for detecting machine-generated multimedia news articles (Fung et al., 2021). The system makes use of the RESIN knowledge graph extraction system to extract a multimodal knowledge graph from multimedia news articles. Then, the knowledge graph is instantiated in a graph neural network, and predictions are made for each relationship within the graph. Not only did the system outperform the state-of-the-art machine-generated detection system at the time substantially, but it was also capable of producing fine-grained explanations of its predictions by identifying the specific inconsistencies within the knowledge structure.

5. Relationship with Future Research and Development

a. Anticipated Results

In Phase I, our primary objective is to demonstrate a proof-of-concept system that effectively supports user interaction with a dynamic multimodal knowledge graph. We will leverage the existing RESIN framework, repurposing its components alongside more recent multimedia knowledge graph extraction techniques. The process will begin with generating a base knowledge graph from multimedia sources, integrating a number of refinement techniques to enhance accuracy and completeness as well as better align with analyst domain knowledge. Key components as described above include knowledge graph completion, logical inconsistency detection, information gap detection, and schema matching and refinement. These components will collectively contribute to refining the initial knowledge graph extracted by our system and will also refine the graph after interactions.

To establish a functional interactive interface, we plan to extend the existing RESIN Editor and our PyReason framework to facilitate direct user interaction with the knowledge graph. Analysts will have the option to interact with the graph using both manual edits and natural language commands, which will be translated into a series of structured graph operations. After any user modifications, the system will reapply the refinement techniques to ensure semantic consistency and coherence across the updated knowledge graph.

While we expect Phase I to deliver a foundational proof-of-concept for knowledge graph refinement, there will be notable limitations. For instance, the proposed knowledge graph completion approach relies on fine-grained logical entailment generation to infer new relationships or entities based on existing evidence. A core challenge is determining the appropriate set of evidence to provide as input for this inference process. In this phase, we will use a graph traversal technique to aggregate information within a limited radius of selected nodes, as outlined in Tang et al. (2024), or use our recent multimedia evidence summarization approach to retrieve relevant evidence (Chen et al. 2024). However, both approaches have scalability challenges; as the knowledge graph expands, traversal becomes computationally intensive and dependent on the relevance of the selected evidence. Inaccurate evidence selection can lead to erroneous inferences, potentially reducing the quality of knowledge graph completion.

Consequently, in Phase I, we expect to achieve primarily shallow inferences, introducing straightforward relationships and new entities from localized multimedia data. The synthesis of more complex inferences, particularly those requiring the integration of distantly related evidence, will almost certainly remain beyond the capabilities of the Phase I prototype. As such, the proposed system's capacity for deeper, multimodal knowledge synthesis is anticipated to be limited and will be explored further in Phase II.

Our approach for logical inconsistency and gap detection uses our fine-grained entailment classifier (Tang et al. 2024) as well as the use of deductive inference via PyReason (Aditya et al., 2024) applied within the

multimedia knowledge graph. Unlike the knowledge graph completion technique, which depends on the logical entailment generator, this component leverages a classifier to identify inconsistencies by evaluating the logical alignment between knowledge graph elements and supporting multimedia evidence. We anticipate that this approach will effectively detect semantic inconsistencies when provided with relevant, semantically related evidence from the knowledge graph.

However, like the challenges faced in knowledge graph completion, this technique requires identifying the correct set of supporting evidence for each knowledge element under review. Determining the most relevant evidence for logical entailment is complex, as the logical relationships within the knowledge graph can be indirect or involve semantically distant facts. In Phase I, we will detect and explain logical inconsistencies arising from analyst domain knowledge – which will provide evidence back to the analyst on the source of the inconsistency. Additionally, for classifier-based inconsistency (Tang et al., 2024), we plan to utilize evidence retrieval techniques, incorporating graph traversal algorithms to select evidence within a specified distance from the target node. This strategy will allow us to conduct fine-grained logical entailment verification when evidence is within reasonable semantic proximity to the hypothesis being tested.

Similarly, our gap detection approach will identify portions of the knowledge graph where entailment confidence is low or where the classifier predicts neutrality, indicating weak or insufficient evidence for nodes or relationships. These nodes and edges will be flagged as areas with weaker evidential support, suggesting potential information gaps that may require further refinement or additional data integration. In Phase I, we anticipate that gap detection will be feasible primarily for cases involving semantically related evidence, with more complex gap identification being deferred to Phase II.

Our schema matching and refinement components will leverage the existing schema matching and schema induction techniques embedded within RESIN. After generating initial schemas, we will align the knowledge graph to one or more of these schemas, providing analysts with a high-level structural overview that aids in understanding the relationships within the graph. As analysts interact with and refine the knowledge graph through our updated RESIN-Editor, we will initiate a new round of knowledge graph refinement, followed by another schema matching step. This process will measure whether the schema continues to accurately represent the updated knowledge graph. If the fit between the graph and the schema falls below a predefined threshold, the system will trigger schema induction to generate an updated schema.

However, there are several limitations to this baseline approach. The current design does not explicitly track user-made updates in the knowledge graph to inform the schema induction process. As a result, it does not assign greater weight to user edits compared to pre-existing data which reduces the accuracy of induced schemas. In Phase I, we anticipate that our system will match the multimodal knowledge graph to a set of schemas that can adapt to user edits. However, the refinement process will offer limited control over schema induction, with no direct user input for schema manipulation.

In addition, while the RESIN Editor's extensions will facilitate user interaction with the knowledge graph itself, they will not support direct manipulation of the schema. Any schema changes will occur indirectly, reflecting the user's modifications to the instance graph rather than allowing direct schema edits. In future program phases, we plan to introduce more advanced schema controls that enable finer adjustments to both the schema induction and matching processes, allowing for improved alignment with operational use cases as required by analysts.

Lastly, our multimedia knowledge graph extraction and refinement components rely on static vision and language models, which will not be updated through user interaction. While we expect this initial design to function reasonably as a proof-of-concept in Phase I, it overlooks a potentially useful supervisory signal inherent in the user edits to the knowledge graph. As user modifications are not used to refine the underlying models, there is a potential risk of user frustration, as they may need to repeatedly correct similar errors in different operational contexts. For example, if the system consistently fails to extract a specific type of multimodal information, users will have to insert this data repeatedly manually or correct the same error repeatedly across various use cases, leading to increased cognitive load and reduced operational efficiency.

Ideally, all components of IKOS—including knowledge graph extraction, vision feature detection, and refinement modules—would dynamically adapt in response to user edits, ensuring that user feedback informs and improves the system's performance. In Phase II, we plan to address this gap by leveraging user interactions as an additional supervisory signal. This will enable adaptive refinement of all components

involved in knowledge graph extraction and refinement which will ultimately improve the system's responsiveness and reduce the need for repeated corrections of the same issues by analysts.

b. Significance of Phase 1 as foundation for Phase 2

In Phase 1, we will develop a prototype of the IKOS system capable of extracting a multimedia knowledge graph, performing basic refinement, and enabling user interaction. While this initial version will have limitations, as outlined in the anticipated results section, it will serve as a critical foundation for Phase 2. The Phase 1 prototype will expose us to several crucial challenges inherent in applying the RESIN framework to intelligence analysis use cases, offering valuable insights that will inform larger-scale development efforts in Phase 2.

This initial system will act as a testbed, allowing us to evaluate both the strengths and shortcomings of the existing RESIN extraction capabilities and the newer components we integrate. In particular, Phase 1 will provide a clear understanding of how well our knowledge completion, inconsistency detection, and schema refinement modules perform in real-world intelligence scenarios. For knowledge completion specifically, we aim to study the types of missing information introduced by users and the extent to which this missing knowledge could have been automatically inferred by the system. This will help motivate the need, if it exists, for automatically refining IKOS' multimodal information extraction capabilities on the fly. Further, by tracking the provenance of user-added knowledge, we will gain insights into the degree of indirect knowledge relationships required for effective completion. The extent of such indirect, complex multimodal knowledge synthesis remains uncertain in operational intelligence settings, making the Phase 1 system a valuable testbed to explore these dynamics.

Additionally, our approach for logical inconsistency and gap detection will help identify the extent to which the system can detect semantic inconsistencies requiring multi-hop and multi-modal inferences as well as inconsistencies with analyst-specified domain knowledge. We will explore the limitations of the current system, particularly its ability to handle complex inconsistencies that involve intermediate conclusions across different modalities. This includes determining whether more sophisticated evidence retrieval techniques, beyond surface-level semantic similarity, are needed to identify logically relevant evidence. We will also evaluate the effectiveness of gap detection in identifying evidential gaps or perspective gaps that could have operational significance, such as missing evidence that might support or refute critical hypotheses. By leveraging the RESIN system's disparate hypothesis generation capabilities, we will test the system's potential to propose missing hypotheses of interest. This exploration will help us understand the technical limitations of our current gap detection capabilities, which can then be refined in Phase 2 to better address analyst requirements and reduce cognitive burden.

Lastly, for schema induction and refinement, Phase 1 will explore how well customized schemas can align with operational intelligence workflows. We will investigate the degree to which schemas generated in Phase 1 support specific analytical tasks and how effectively these schemas can be adapted to different intelligence analyst use cases. Our results from this phase will directly inform the development of more advanced schema induction techniques in Phase 2, such as dynamic schema customization based on completed workflows, user queries, and analyst feedback during interactions. By analyzing various use cases, we will assess the feasibility of task-driven schema adaptation and the potential for schema learning from past interactions or training data. These insights will shape Phase 2's design to ensure that the knowledge graph system can better meet diverse operational demands.

c. Requirements for Phase 2

IKOS is envisioned to delivery software that can be integrated into a government program of record that supports intelligence analysis. Accordingly, IKOS will require an architecture and system security plan that will support such integration. The lead security engineer for BRG has experience with managing the Risk Management Framework to achieve an Authority to Operate on other classified information systems.

As described in our Statement of Work above, our mission engineering efforts will connect us with at least one applicable program office (i.e., C2IE), and orient our design to meet their system architecture requirements. To execute our proposed mission engineering effort, some of our team will require TS security clearances. This will allow us to visit and interact with contemporary OSINT analysts and transition program

officers. BRG has a TS facility clearance, and key persons on this project have active TS-SCI clearances. We will request a DD-254 for this project to support the clearances necessary for our proposed work.

6. Commercialization Strategy

a. First Targeted Product

The initial focus for IKOS is to deliver capabilities to military intelligence programs of record, beginning with the Command and Control of the Information Environment (C2IE), an analytic platform that provides open-source data as a means of real-time situational awareness. IKOS's multimodal data integration will strengthen C2IE's capacity to deliver actionable intelligence. In the private sector, IKOS will be positioned as a business intelligence (BI) solution for critical infrastructure firms (e.g., energy providers, financial firms), where real-time insights drive essential decision-making (U.S. Department of Homeland Security, 2024).

b. Customers and estimated market size.

The DoD has allocated \$1.4 billion in 2024 for information dominance and is likely to increase these investments in future years as it postures for conflicts in the Pacific (U.S. Department of Defense, 2023).

Accordingly, DoD programs such as C2IE represent significant customers for IKOS. In the private sector, IKOS will offer great value to companies seeking intelligence to protect critical infrastructure. Current business intelligence platforms fail to meet the decision support demands for these companies. As shown in figure 14, the global business intelligence market in 2024 is estimated at \$21.1 billion, with a CAGR of 10.8% through 2033 (Persistence Market Research, 2023). Within this landscape, knowledge graphs offer an unmet need for a \$1.1 billion segment that has a 21.8% CAGR (Markets & Markets, 2024). IKOS will foreseeably provide value to the 10,000 U.S. critical infrastructure entities, which represent a total \$500 million opportunity based on average implementation costs for open-source intelligence.

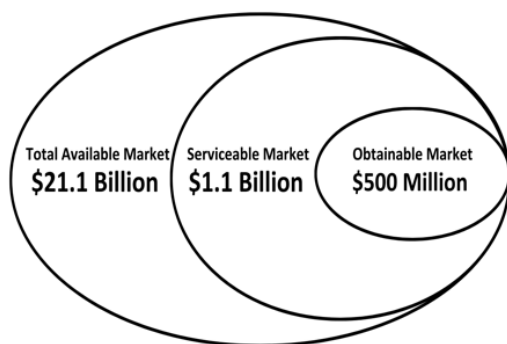


Figure 14 - IKOS target market size

c. Funding required to bring the technology to market and how will that money be raised

Beyond the research and development funding from Phase 2 SBIR award, we will seek additional venture capital to bring IKOS to market. As we near completion of Phase 2, we plan to seek \$3 million in venture capital for a Series A investment. This funding will allow us to decide whether to license the capability to existing platforms or seek additional funding (Series B) to bring a new platform to market. We anticipate Series B funding to be approximately \$10 million for our \$500 million market valuation described above. To raise these investments, BRG will leverage its established networks of investment bankers. In particular, we will target venture capital firms known for successfully scaling and exiting positions in business intelligence platforms, such as Baird Capital, Viking Global Investors, and Grotech Ventures. Such partnerships will enable strong market penetration and create additional value for users of IKOS.

d. Corporate marketing expertise or plan to bring that into the company

BRG's current marketing strengths are in the defense and cybersecurity sectors. We will bring business intelligence marketing expertise on board with Series A investment. This will initially be contracted marketing from established firms specializing in marketing novel innovative solutions. Our preferred advertising firm is Robot House. Robot House offers boutique advertising out of Oklahoma City, and benefits from its proximity to critical infrastructure providers in the energy sector.

e. Competitors and BRG Competitive Advantage over the competitors

Competing intelligence platforms, such as the commercially available AlphaSense and Tellius, incorporate knowledge graphs as ancillary features, yet lack the specialized configurations essential to drive informed

decisions in government and critical infrastructure (Owler, 2024). BRG distinguishes itself from these platform providers with its focus on mission engineering and advanced technology development.

f. Schedule

Table 4 provides a high-level description of our phased commercialization approach for IKOS.

Table 4 - Commercialization Schedule

Phase	Duration	Objectives
I. Concept and Feasibility	6 months	<ul style="list-style-type: none"> Establishing feasibility and aligning with market needs; refining commercialization strategy
II. Research and Development	12 months	<ul style="list-style-type: none"> Refining product and testing in real-world settings; preparing venture capital pitch Transition roadmap with DoD program of record
III. Commercialization	24 months	<ul style="list-style-type: none"> Transition IKOS to a DoD program of record Series A (\$3M) fund raise – year 3 Series B (\$10M) fund raise – year 4

7. Key Personnel

BRG is a non-traditional defense contractor that only employs U.S. Citizens. Additionally, no subcontractors or consultants will be employed in the proposed work. Specific publications for key persons are included in bold in the references section.

Stoney Trent – Principal Investigator	
Education:	University of Notre Dame, B.S. – Mechanical Engineering, 1995 The Ohio State University, M.S. Systems Engineering, 2005 The Ohio State University, Ph.D. Systems Engineering, 2005
Relevant Experience:	Cognitive Engineer and Military Intelligence and Cyber Warfare veteran with over 15 years of experience leading mission engineering for the defense and intelligence communities. Dr. Trent is the President of BRG and a research professor at the Virginia Tech National Security Institute, where he focuses on human-machine teaming and trustworthy AI systems. While on active duty, he designed the Joint Artificial Intelligence Center (JAIC) for the Department of Defense, where he established product lines to deliver human-centered AI to improve warfighting and business functions. Dr. Trent is a co-PI on IARPA ReSCIND, where he is leading mission engineering, experimentation, and system integration for a project to deliver AI-enabled cybersecurity capabilities. Dr. Trent is also a PI on AFRL RAINFLY, where he is leading the design and system integration for a cognitive psychology-based Information Warfare capability. Dr. Trent is a graduate of the Army War College and former Cyber Fellow at the National Security Agency.
Relevant Awards / Patents:	Defense Superior Service Medal – Establishing mission engineering methods that contributed to the successful delivery of AI-enabled capabilities at the Joint AI Center
Relevant Publications:	Trent, S., Patterson, E., Woods, D. (2007). Challenges for cognition in intelligence analysis. Journal for Cognitive Engineering and Decision making.

Chris Thomas – Lead Computer Scientist (co-PI)	
Education:	University of Pittsburgh, B.S. - Computer Science, 2013 University of Pittsburgh, PhD - Computer Science, 2013 Postdoctoral Researcher, Columbia University, 2020-2022
Relevant Experience:	

Chris Thomas, Ph.D., is an Assistant Professor in the Department of Computer Science at Virginia Tech. His research interests lie at the intersection of computer vision and natural language processing. A primary focus of his research is to build robust AI systems that can understand and reason about real-world multimedia content (images, text, video, and audio) as humans do. He previously represented Columbia University in the DAPRA SemaFor program focused on disinformation detection. He has also worked as part of other DARPA programs focused on multimedia understanding and knowledge graphs including AIDA and KAIROS. His work has appeared in top conferences and journals, including CVPR, ACL, NeurIPS, ECCV, IEEE PAMI, and IJCV. He has served on program committees for over twenty conferences and journals and has won six outstanding reviewer awards at top vision conferences (CVPR, ICCV, ECCV). He has diverse ongoing interdisciplinary collaborations with faculty from other fields and institutions, including Columbia, UIUC, UCLA, and the University of Pennsylvania.

Relevant Publications:

Tang, C.-W., Chen, T.-C., Ishmam, A. M., Nguyen, K. A., Mehrab, K. S., & Thomas, C. (2024). M3D: MultiModal MultiDocument Fine-Grained Inconsistency Detection. To appear, In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. Long papers.

Thomas, C., Zhang, Y., & Chang, S. F. (2022, October). Fine-grained visual entailment. In European Conference on Computer Vision (pp. 398-416). Cham: Springer Nature Switzerland.

Fung, Y., Thomas, C., Reddy, R. G., Polisetty, S., Ji, H., Chang, S. F., ... & Sil, A. (2021, August). Infosurgeon: Cross-media fine-grained information consistency checking for fake news detection. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers) (pp. 1683-1698).

Chen, T. C., Tang, C. W., & Thomas, C. (2024, August). MetaSumPerceiver: Multimodal Multi-Document Evidence Summarization for Fact-Checking. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 8742-8757).

Ayyubi, H., Thomas, C., Chum, L., Lokesh, R., Chen, L., Niu, Y., ... & Chang, S. F. (2024, March). Beyond Grounding: Extracting Fine-Grained Event Hierarchies across Modalities. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 38, No. 16, pp. 17664-17672).

Chen, L., Niu, Y., Chen, B., Lin, X., Han, G., Thomas, C., ... & Chang, S. F. (2022). Weakly-Supervised Temporal Article Grounding. In 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022.

Chen, B., Lin, X., Thomas, C., Li, M., Yoshida, S., Chum, L., ... & Chang, S. F. (2021, November). Joint Multimedia Event Extraction from Video and Article. In Findings of the Association for Computational Linguistics: EMNLP 2021 (pp. 74-88).

Paulo Shakarian – Senior Computer Scientist

Education: United States Military Academy, B.S. - Computer Science, 2002
University of Maryland, College Park, M.S. – Computer science, 2009
University of Maryland, College Park, Ph.D. – Computer science, 2011

Relevant Experience:

A Senior Computer Scientist at BRG and a tenured Associate Professor in the School of Computing and AI at the Fulton Schools of Engineering at Arizona State University. He also holds the additional position as Research Director for the School of Computing and AI at ASU. He specializes in the fusion of symbolic artificial intelligence and machine learning – publishing numerous scientific books and papers. Shakarian was named a “KDD Rising Star,” received the Air Force Young Investigator award, received multiple “best paper” awards and has been featured in major news media outlets such as CNN and The Economist. Paulo has been funded by various organizations including IARPA (HAYSTAC, CAUSE, ICARUS), ARO (4x), ONR (5x), AF/AFOSR (2x), and DARPA as well as various industry partners. Paulo also co-founded a startup company that used machine learning to predict future exploits; the company was acquired after raising \$8 million in venture capital and having obtained over 80 customers. Paulo also founded and manages the Neuro Symbolic Channel on YouTube which has over 2,500 subscribers. Earlier in his career, Paulo was an officer in the U.S. Army where he served two combat tours in Iraq, earning a Bronze Star and the Army Commendation Medal for Valor. During his military career, Paulo also served as a DARPA Fellow and as an advisor to IARPA.

Relevant Awards / Patents:

Holds 10 issued US patents, prior experience commercializing AI technology used by cybersecurity analysts, prior training as an intelligence analyst (Army Military Intelligence Officer Basic and Captain Career course graduate)

Relevant Publications:

D. Bavikadi, D. Aditya, D. Parkar, P. Shakarian, G. Mueller, C. Parvis, G. Simari, Geospatial Trajectory Generation via Efficient Abduction: Deployment for Independent Testing, 40th Intl. Conference on Logic Programming (ICLP-24).

D. Aditya, K. Mukherji, S. Balasubramanian, A. Chaudhary, P. Shakarian, PyReason: Software for Open World Temporal Logic, AAAI Spring Symposium (Mar. 2023).

Amy Sliva – Senior Computer Scientist

Education: Georgetown University, B.S. - Computer Science, 2005
University of Maryland, College Park, M.S. – Computer Science, 2007
University of Maryland, College Park, M.P.P. – International Security and Economic Policy, 2010
University of Maryland, College Park, Ph.D. – Computer Science, 2011

Relevant Experience:

A computational social scientist with expertise in artificial intelligence, cyber security, and international security and economics. Amy is an Assistant Professor and Program Director of Computer Science at King's College and has also served on the faculty of Northeastern University in both Computer Science and Political Science. She has over 15 years' experience developing large-scale data analytics and artificial intelligence models of human behavior to support decision making in the National Security and Intelligence Communities. As an experienced Principal Investigator and Project Manager for DARPA and IARPA Prime contracts and DoD SBIRs, Amy has led commercial and research software development to include the development and use of APIs.

Relevant Awards / Patents:

Text

Relevant Publications:

"Decision Analysis in Stochastic Sociocultural Systems" (with Emanuele Borgonovo, Alexander Levis, Christopher Pawlenok, and Nathaniel Plaspohl). *Proceedings of 27th International Conference on Analytical & Stochastic Modelling Techniques & Applications (ASMTA)*, 2023.

"Causal Analysis Graph Modeling for Strategic Decisions" (with Alexander Levis). In *Proceedings of European Council for Modelling and Simulation (ECMS) 36th International Conference on Modelling and Simulation*, 2022,

"Food Security in Developing Economies: A Decision Analysis Approach" (with Alexander Levis, Emanuele Borgonovo, and Glenn Takata). *Institute for Operations Research and the Management Sciences (INFORMS) Annual Meeting*. 2020.

Brandon Beltz – Senior User Experience Researcher

Education: University of Utah, B.S. Philosophy, 2000
George Mason University, M.A Experimental Psychology, 2004
George Mason University, Ph.D. Experimental Psychology, 2009

Relevant Experience:

Brandon Beltz is a Principal User Experience Researcher at the Bulls Run Group with 16 years of experience enhancing software product usability and adoption. He has led multimethod research projects across several industries including digital health, finance, and intelligence analysis tools. His professional roles have included Senior User Research Manager at Sage, Director of User Experience Research at Vibrent Health, and user experience leadership roles at MITRE and Northrop Grumman.

Relevant Awards / Patents:

None

Relevant Publications:

None

James Doty III – Project Manager and Intelligence Analysis SME

Education:	USMA, B.S. History, 1995 The Ohio State University, M.A. History, 2005 The Ohio State University, Ph.D. History, 2010
Relevant Experience:	A Military Intelligence veteran, historian, and is the Chief of Project Management for the BRG. Dr. Doty specializes in translating operational experiences for technology programs. He has over 20 years' experience leading analytic teams and preparing assessments for senior leaders. As the Senior Intelligence Officer for the National Training Center (NTC) in Fort Irwin, CA, Jim coordinated the planning and execution of intelligence training and wargames for Army units preparing to deploy and collaborated with the Defense Advanced Research Projects Agency's INSIGHT program. As a Professor of Military Science, he directed leadership and military technical training for cadets preparing for commissions in the US Army. Jim has served as a Strategic Policy Research Fellow with the RAND Arroyo Center and is a former Assistant Professor of Military History at the United States Military Academy. He has deployed to combat and other operations in Iraq, Saudi Arabia, Germany, and Kosovo. Dr. Doty will serve as the Project Manager and Intelligence SME.
Relevant Awards / Patents:	None
Relevant Publications:	None

Van Truong – System Architect / Senior Software Engineer	
Education:	NC A&T State University, M.S. in Computer Science, 2009 NC A&T State University, Computer Science, 2007
Relevant Experience:	Senior Software Engineer with over 15 years of experience in defense and commercial projects involving software development, testing, and automation. He has excelled in roles that include tester, developer, system administrator, and tech lead for the various systems, including radar systems, satellite systems, transportation systems, MES systems, and smart technical manuals. Mr. Truong is currently the Lead System Integrator on IARPA ReSCIND, where he is integrating AI modules and other software components to deliver a containerized cybersecurity application. Previously, Mr. Truong was in the Raytheon Rotation Leadership Development program, where he went through cross-functional training and leadership development. Van has Security+ and Raytheon Six Sigma certifications. Van has a bachelor's and master's degree in computer science with a master's thesis in natural language processing (NLP) for assisting semi-auto tutoring systems. He will serve as the Lead Software Engineer.
Relevant Awards / Patents:	None
Relevant Publications:	Developing an Algebra Sublanguage for Tutorial Dialogues, Twentieth Midwest AI And Cognitive Science Conference (MAICS 2009), Fort Wayne

8. Foreign Citizens. None**9. Facilities/Equipment.**

BRG operates a distributed software factory with all team members using AWS GovCloud and Microsoft Government Cloud Compute High for development and storage of Controlled Unclassified Information. BRG is FEDRAMP and NIST compliant. All workspaces meet local, state, and federal environmental and safety regulations. BRG will develop the objective system in accordance with ISO/IEEE 12207 Systems and software engineering software life cycle processes and CMMI best practices. Using CI/CD and Agile development processes, BRG delivers rapid prototypes that are ready for rapid transition. In general, our software teams employ two-week sprints consisting of planning, development, review, and feedback every month. As a Cybersecurity Maturity Model Certification Registered Practitioner Organization, BRG will ensure that system development incorporates cybersecurity throughout the development lifecycle. Cybersecurity engineers will be embedded within the

software development team to ensure design and development activities are aligned with the Risk Management Framework (RMF). Should a government information system be desired, our security engineers will initiate the documents necessary for an Authority to Operate (ATO) in Phase II.

10. Subcontractors/Consultants. None

11. Prior, Current or Pending Support of Similar Proposals or Awards. No prior, current, or pending support has been provided for proposed work.

12. Identification and Assertion of Restrictions on the Government's Use, Release, or Disclosure of Technical Data or Computer Software. None

References

- Aditya, S., Yang, Y., & Baral, C. (2019, August). Integrating Knowledge and Reasoning in Image Understanding. In Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence.
- Aditya, K. Mukherji, S. Balasubramanian, A. Chaudhary, P. Shakarian, [PyReason: Software for Open World Temporal Logic](#), *AAAI Spring Symposium* (Mar. 2023).
- Almukaynizi, E. Nunes, K. Dharaiya, M. Senguttuvan, J. Shakarian, P. Shakarian, [Proactive Identification of Exploits in the Wild Through Vulnerability Mentions Online](#), *2017 IEEE International Conference on Cyber Conflict (CyCon-US)* (Nov. 2017).
- Bavikadi, N. Lee, P. Shakarian, C. Parvis, [Sea-cret Agents: Maritime Abduction for Region Generation to Expose Dark Vessel Trajectories](#), submitted (2024).
- Bavikadi, D. Aditya, D. Parkar, P. Shakarian, G. Mueller, C. Parvis, G. Simari, [Geospatial Trajectory Generation via Efficient Abduction: Deployment for Independent Testing](#), *40th Intl. Conference on Logic Programming (ICLP-24)*.
- Bollacker, K., Evans, C., Paritosh, P., Sturge, T., & Taylor, J. (2008, June). Freebase: a collaboratively created graph database for structuring human knowledge. In Proceedings of the 2008 ACM SIGMOD international conference on Management of data (pp. 1247-1250).
- Chen, B., Lin, X., Thomas, C., Li, M., Yoshida, S., Chum, L., ... & Chang, S. F. (2021, November). Joint Multimedia Event Extraction from Video and Article. In Findings of the Association for Computational Linguistics: EMNLP 2021 (pp. 74-88).
- Chen, L., Niu, Y., Chen, B., Lin, X., Han, G., Thomas, C., ... & Chang, S. F. (2022, December). Weakly-Supervised Temporal Article Grounding. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (pp. 9402-9413).
- Chen, T. C., Tang, C. W., & Thomas, C. (2024, August). MetaSumPerceiver: Multimodal Multi-Document Evidence Summarization for Fact-Checking. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 8742-8757).
- Chen, X., Zhang, N., Li, L., Deng, S., Tan, C., Xu, C., ... & Chen, H. (2022, July). Hybrid transformer with multi-level fusion for multimodal knowledge graph completion. In Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval (pp. 904-915).
- Chen, Z., Chen, J., Zhang, W., Guo, L., Fang, Y., Huang, Y., ... & Chen, H. (2023, October). Meaformer: Multi-modal entity alignment transformer for meta modality hybrid. In Proceedings of the 31st ACM International Conference on Multimedia (pp. 3317-3327).
- Cheng, S., Zhang, N., Tian, B., Chen, X., Liu, Q., & Chen, H. (2024, March). Editing language model-based knowledge graph embeddings. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 38, No. 16, pp. 17835-17843).
- Chu, Y., Xu, J., Zhou, X., Yang, Q., Zhang, S., Yan, Z., ... & Zhou, J. (2023). Qwen-audio: Advancing universal audio understanding via unified large-scale audio-language models. arXiv preprint arXiv:2311.07919.
- Currin, G. (2023). When trust is a factor. Columbia Magazine. From <https://topics.engineering.columbia.edu/when-trust-is-a-factor/>
- Ding, Y., Yu, J., Liu, B., Hu, Y., Cui, M., & Wu, Q. (2022). MuKEA: Multimodal Knowledge Extraction and Accumulation for Knowledge-based Visual Question Answering. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5089-5098).
- Dror, R., Wang, H., & Roth, D. (2023, May). Zero-Shot On-the-Fly Event Schema Induction. In Findings of the Association for Computational Linguistics: EACL 2023 (pp. 705-725).
- Du, X., Zhang, Z., Li, S., Yu, P., Wang, H., Lai, T., ... & Ji, H. (2022, July). RESIN-11: Schema-guided event prediction for 11 newsworthy scenarios. In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: System Demonstrations (pp. 54-63).
- Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., ... & Ganapathy, R. (2024). The llama 3 herd of models. arXiv preprint arXiv:2407.21783.
- Fung, Y., Thomas, C., Reddy, R. G., Polisetty, S., Ji, H., Chang, S. F., ... & Sil, A. (2021, August). Infosurgeon: Cross-media fine-grained information consistency checking for fake news detection. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers) (pp. 1683-1698).

- Fung, Y., Wang, H., Wang, T., Kebarighotbi, A., Bansal, M., Ji, H., & Natarajan, P. (2023, May). DeepMaven: Deep question answering on long-distance movie/TV show videos with multimedia knowledge extraction and synthesis. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 3041-3051).
- Getoor, L., Friedman, N., Koller, D., Pfeffer, A., Taskar, B. (2007). Probabilistic relational models. *Introduction to statistical relational learning* 8.
- Gupta, S., Kenkre, S., & Talukdar, P. (2019, November). Care: Open knowledge graph embeddings. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 378-388).
- Haider, S., Levis, A.H. (2008). Modeling time-varying uncertain situations using dynamic influence nets. *International Journal of Approximate Reasoning* 49(2), 488–502.
- Hao, Y., Cao, P., Chen, Y., Liu, K., Xu, J., Li, H., ... & Zhao, J. (2023, December). Complex Event Schema Induction with Knowledge-Enriched Diffusion Model. In *Findings of the Association for Computational Linguistics: EMNLP 2023* (pp. 4809-4825).
- Hogan, A., Blomqvist, E., Cochez, M., d’Amato, C., Melo, G. D., Gutierrez, C., ... & Zimmermann, A. (2021). Knowledge graphs. *ACM Computing Surveys (Csur)*, 54(4), 1-37.
- Howard, R.A., Matheson, J.E. (2005). Influence diagrams. *Decision Analysis* 2(3), 127–143.
- Ilievski, F., Garijo, D., Chalupsky, H., Divvala, N. T., Yao, Y., Rogers, C., ... & Szekely, P. (2020). KGTK: a toolkit for large knowledge graph manipulation and analysis. In *The Semantic Web–ISWC 2020: 19th International Semantic Web Conference, Athens, Greece, November 2–6, 2020, Proceedings, Part II* 19 (pp. 278-293). Springer International Publishing.
- Jin, X., Li, M., & Ji, H. (2022, July). Event schema induction with double graph autoencoders. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 2013-2025).
- Kato, K., Li, Y., & Gupta, A. (2018). Compositional learning for human object interaction. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 234-251).
- Kimmig, A., Bach, S., Broecheler, M., Huang, B., Getoor, L. (2012). A short introduction to probabilistic soft logic. In: *Proceedings of the NIPS workshop on probabilistic programming: foundations and applications*. pp. 1–4.
- Li, M., Li, S., Wang, Z., Huang, L., Cho, K., Ji, H., ... & Voss, C. (2021, November). The Future is not One-dimensional: Complex Event Schema Induction by Graph Modeling for Event Prediction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (pp. 5203-5215).
- Li, Q., Chen, Z., Ji, C., Jiang, S., & Li, J. (2024). LLM-based Multi-Level Knowledge Generation for Few-shot Knowledge Graph Completion. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (2024)*. <https://api.semanticscholar.org/CorpusID/271494703>.
- Li, M., Zareian, A., Lin, Y., Pan, X., Whitehead, S., Chen, B., & Freedman, M. (2020, July). Gaia: A fine-grained multimedia knowledge extraction system. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 77-86).
- Li, M., Zareian, A., Zeng, Q., Whitehead, S., Lu, D., Ji, H., & Chang, S. F. (2020, July). Cross-media Structured Common Space for Multimedia Event Extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 2557-2568).
- Li, S., Zhao, R., Li, M., Ji, H., Callison-Burch, C., & Han, J. (2023, January). Open-Domain Hierarchical Event Schema Induction by Incremental Prompting and Verification. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Vol. 1)*.
- Liang, W., Meo, P. D., Tang, Y., & Zhu, J. (2024). A survey of multi-modal knowledge graphs: Technologies and trends. *ACM Computing Surveys*, 56(11), 1-41.
- Lin, Q., Mao, R., Liu, J., Xu, F., & Cambria, E. (2023). Fusing topology contexts and logical rules in language models for knowledge graph completion. *Information Fusion*, 90, 253-264.
- Liu, Y., Zeng, Q., Ordieres Meré, J., & Yang, H. (2019). Anticipating stock market of the renowned companies: A knowledge graph approach. *Complexity*, 2019.
- Ma, Y., Wang, Z., Li, M., Cao, Y., Chen, M., Li, X., ... & Shao, J. (2022). MMEKG: Multi-modal event knowledge graph towards universal representation across modalities. *Association for Computational Linguistics*.

- Marino, K., Salakhutdinov, R., & Gupta, A. (2017). The More You Know: Using Knowledge Graphs for Image Classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2673-2681).
- Mausam, M. (2016, July). Open information extraction systems and downstream applications. In Proceedings of the twenty-fifth international joint conference on artificial intelligence (pp. 4074-4077).
- Markets & Markets. (2024). *Knowledge graph market*. <https://www.marketsandmarkets.com/Market-Reports/knowledge-graph-market-217920811.html>
- Nguyen, K. D., Zhang, Z., Suchocki, R., Li, S., Palmer, M., Brown, S., ... & Ji, H. (2023). RESIN-EDITOR: A Schema-guided Hierarchical Event Graph Visualizer and Editor. In 2023 Conference on Empirical Methods in Natural Language Processing, EMNLP 2023.
- Owler. (2024). *Tellius competitors and alternatives*. <https://www.owler.com/company/tellius/competitors>
- Pearl, J. (1988). Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann.
- Persistence Market Research. (2023). *Business intelligence platform market*. <https://www.persistencemarketresearch.com/market-research/business-intelligence-platform-market.asp>
- Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2023, July). Robust speech recognition via large-scale weak supervision. In International conference on machine learning (pp. 28492-28518). PMLR.
- Rajabi, E., & Etmnani, K. (2024). Knowledge-graph-based explainable AI: A systematic review. Journal of Information Science, 50(4), 1019-1029.
- Richardson, M., Domingos, P. (2006). Markov logic networks. Machine learning 62, 107– 136 (2006).
- Almukaynizi, J. Shakarian, P. Shakarian, *Mining user interaction patterns in the darkweb to predict enterprise cyber incidents*, *Social Network Analysis and Mining*, 9(1), 2019.
- Seo, S., Oh, B., Jo, E., Lee, S., Lee, D., Lee, K. H., ... & Lee, Y. (2021). Active learning for knowledge graph schema expansion. IEEE Transactions on Knowledge and Data Engineering, 34(12), 5610-5620.
- Shakarian, G.I. Simari, *Extensions to Generalized Annotated Logic and an Equivalent Neural Architecture*, *IEEE TransAI*, 2022.
- Shen, T., Zhang, F., & Cheng, J. (2022). A comprehensive overview of knowledge graph completion. Knowledge-Based Systems, 255, 109597.
- Subagdja, B., Shanthoshigaa, D., Wang, Z., & Tan, A. H. (2024). Machine Learning for Refining Knowledge Graphs: A Survey. ACM Computing Surveys, 56(6), 1-38.
- Tang, C.-W., Chen, T.-C., Ishmam, A. M., Nguyen, K. A., Mehrab, K. S., & Thomas, C. (2024). M3D: MultiModal MultiDocument Fine-Grained Inconsistency Detection. To appear, In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. Long papers.
- Thomas, C., Zhang, Y., & Chang, S. F. (2022, October). Fine-grained visual entailment. In European Conference on Computer Vision (pp. 398-416). Cham: Springer Nature Switzerland.
- U.S. Department of Homeland Security. (2024). *Critical infrastructure*. <https://www.dhs.gov/archive/science-and-technology/critical-infrastructure>
- U.S. Department of Defense. (2023, March). *Department of Defense releases the President's fiscal year 2024 defense budget*. <https://www.defense.gov/news/releases/release/article/3326875/department-of-defense-releases-the-presidents-fiscal-year-2024-defense-budget/>
- Vrandečić, D., & Krötzsch, M. (2014). Wikidata: a free collaborative knowledgebase. Communications of the ACM, 57(10), 78-85.
- Wang, C., Zhou, X., Pan, S., Dong, L., Song, Z., & Sha, Y. (2022, June). Exploring relational semantics for inductive knowledge graph completion. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 36, No. 4, pp. 4184-4192).
- Wang, D., Tang, K., Zeng, J., Pan, Y., Dai, Y., Li, H., & Han, B. (2024). MM-Transformer: A Transformer-Based Knowledge Graph Link Prediction Model That Fuses Multimodal Features. Symmetry, 16(8), 961.
- Wang, H., Zhang, Z., Li, S., Han, J., Sun, Y., Tong, H., Olive, J. P., & Ji, H. (2022). Proceedings of the 4th Conference on Automated Knowledge Base Construction (AKBC 2022).
- Wang, Q., & Li, M. (2021, June). COVID-19 Literature Knowledge Graph Construction and Drug Repurposing Report Generation. In Proc. The 2021 Conference of the North American Chapter of the Association for Computational Linguistics-Human Language Technologies (NAACL-HLT2021) Demo Track.
- Wen, H., Lin, Y., Lai, T., Pan, X., Li, S., Lin, X., ... & Ji, H. (2021, June). Resin: A dockerized schema-guided cross-document cross-lingual cross-media information extraction and event tracking system. In Proceedings of the

- 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Demonstrations (pp. 133-143).
- Wilcke, W. X., Bloem, P., de Boer, V., & van t Veer, R. H. (2023). End-to-end learning on multimodal knowledge graphs. arXiv preprint arXiv:2309.01169.
- Wu, X., Huang, K. H., Fung, Y., & Ji, H. (2022, July). Cross-document misinformation detection based on event graph reasoning. In Proceedings of the 2022 conference of the north american chapter of the association for computational linguistics: Human language technologies (pp. 543-558).
- Yang, A., Yang, B., Hui, B., Zheng, B., Yu, B., Zhou, C., Li, C., Li, C., Liu, D., Huang, F., & others. (2024). Qwen2 technical report. arXiv Preprint arXiv:2407.10671.
- Ye, J., Xu, H., Liu, H., Hu, A., Yan, M., Qian, Q., ... & Zhou, J. (2024). mplug-owl3: Towards long image-sequence understanding in multi-modal large language models. arXiv preprint arXiv:2408.04840.
- Zhang, M., Xia, Y., Liu, Q., Wu, S., & Wang, L. (2023, July). Learning latent relations for temporal knowledge graph reasoning. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 12617-12631).
- construction. ACM Computing Surveys, 56(4), 1-62.
- Zhang, T., Tham, I., Hou, Z., Ren, J., Zhou, L., Xu, H., ... & Callison-Burch, C. (2023). Human-in-The-loop schema induction. In 61st Annual Meeting of the Association for Computational Linguistics, ACL-DEMO 2023 (pp. 1-10). Association for Computational Linguistics (ACL).
- Zhang, T., Tham, I., Hou, Z., Ren, J., Zhou, L., Xu, H., ... & Callison-Burch, C. (2023, July). Human-in-the-loop Schema Induction. In The 61st Annual Meeting Of The Association For Computational Linguistics.
- Zhang, X., Liang, X., Zheng, X., Wu, B., & Guo, Y. (2022, September). MULTIFORM: few-shot knowledge graph completion via multi-modal contexts. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases (pp. 172-187). Cham: Springer International Publishing.
- Zhang, Z., Zhuang, F., Qu, M., Lin, F., & He, Q. (2018). Knowledge graph embedding with hierarchical relation structure. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (pp. 3198-3207).
- Zhong, L., Wu, J., Li, Q., Peng, H., & Wu, X. (2023). A comprehensive survey on automatic knowledge graph



SBIR Phase I Proposal

Proposal Number	F244-0001-0048
Topic Number	AF244-0001
Proposal Title	IKOS: Interactive Knowledge Graphs for Open-Source Intelligence Analysis
Date Submitted	11/05/2024 05:16:50 PM

Firm Information

Firm Name	The Bulls Run Group, Inc
Mail Address	9207 Bulls Run Pkwy, Bethesda, Maryland, 20817
Website Address	https://www.bulls.run
UEI	KDH9SC281WL8
Duns	129524731
Cage	8LSH4

Total Dollar Amount for this Proposal	\$139,992.88
---------------------------------------	--------------

Base Year	\$139,992.88
Year 2	\$0.00
Technical and Business Assistance(TABA)- Base	\$0.00
TABA- Year 2	\$0.00

Base Year Summary

Total Direct Labor (TDL)	\$108,453.31
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)	\$2,213.00
Total Other Direct Costs (TODC)	\$0.00
G&A (rate 15%) x Base (TDL+TOH+TDT)	\$16,599.95
Total Firm Costs	\$127,266.26
Subcontractor Costs	
Total Subcontractor Costs (TSC)	\$0.00
Cost Sharing	-\$0.00
Profit Rate (10%)	\$12,726.63
Total Estimated Cost	\$139,992.88
TABA	\$0.00

Year 2 Summary

Total Direct Labor (TDL)	\$0.00
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)	\$0.00
G&A (rate 15%) x Base (TDL+TOH+TDT)	\$0.00
Total Firm Costs	\$0.00
Subcontractor Costs	
Total Subcontractor Costs (TSC)	\$0.00
Cost Sharing	-\$0.00
Profit Rate (10%)	\$0.00
Total Estimated Cost	\$0.00
TABA	\$0.00

Base Year

Direct Labor Costs						
Category / Individual-TR	Rate/Hour	Estimated Hours	Fringe Rate (%)	Fringe Cost	Cost	
Chief Executive/ Principal Investigator (Stoney Trent)	\$240.00	90	0	\$0.00	\$21,600.00	
Computer and Information Research Scientist/ coPI (Chris Thomas)	\$150.00	90			\$13,500.00	
Computer and Information Research Scientist/ Senior Computer Scientist (Paulo Shakarian)	\$150.00	70			\$10,500.00	
Computer and Information Research Scientist/ Senior Computer Scientist (Amy Sliva)	\$150.00	70			\$10,500.00	
General and Operations Manager/ Project Manager/ Intel SME (James Doty)	\$81.73	80	15	\$980.76	\$7,519.16	
Computer and Information Research Scientist/ Senior UX Researcher (Brandon Beltz)	\$91.25	80	24	\$1752.00	\$9,052.00	
Software Developer/ System Architect/ Software Engineer (Van Truong)	\$89.09	43	38	\$1455.73	\$5,286.60	
Computer and Information Research Scientist/ Software Developer	\$45.00	240	15	\$1620.00	\$12,420.00	
Subtotal Direct Labor (DL)					\$90,377.76	
Labor Overhead (rate 20%) x (DL)					\$18,075.55	
Total Direct Labor (TDL)					\$108,453.31	

Direct Travel Costs

Project kickoff	\$2,213.00
Total Direct Travel Costs (TDT)	\$2,213.00
G&A (rate 15%) x Base (TDL+TOH+TDT)	\$16,599.95
Cost Sharing	-\$0.00
Profit Rate (10%)	\$12,726.63
Total Estimated Cost	\$139,992.88
TABA	\$0.00

Year 2

Direct Labor Costs						
	Category / Individual-TR	Rate/Hour	Estimated Hours	Fringe Rate (%)	Fringe Cost	Cost
	Chief Executive/ Principal Investigator (Stoney Trent)	\$250.00	0			\$0.00
Subtotal Direct Labor (DL)						\$0.00
Labor Overhead (rate 20%) x (DL)						\$0.00
Total Direct Labor (TDL)						\$0.00

Direct Travel Costs

none	\$0.00
Total Direct Travel Costs (TDT)	\$0.00
G&A (rate 15%) x Base (TDL+TOH+TDT)	\$0.00
Cost Sharing	-\$0.00
Profit Rate (10%)	\$0.00
Total Estimated Cost	\$0.00
TABA	\$0.00

Explanatory Material Relating to the Cost Volume
The Official From the Firm that is responsible for the cost breakdown
Name: Stoney Trent
Phone: (301) 456-4237
Phone: stoney@bullsrungroup.com
Title: Proposal Owner

If the Defence Contracting Audit Agency has performed a review of your projects within the past 12 months, please provide: No
Select the Type of Payment Desired: Partial payments

Cost Volume Details

Direct Labor
Base

Category	Description	Education	Yrs Experience	Hours	Rate	Fringe Rate	Total
Chief Executive	Principal Investigator	PhD	15	90	\$240.00	0	\$21,600.00
Computer and Information Research Scientist	coPI	PhD	10	90	\$150.00		\$13,500.00
Computer and Information Research Scientist	Senior Computer Scientist	PhD	15	70	\$150.00		\$10,500.00
Computer and Information Research Scientist	Senior Computer Scientist	PhD	15	70	\$150.00		\$10,500.00
General and Operations Manager	Project Manager/ Intel SME	PhD	25	80	\$81.73	15	\$7,519.16
Computer and Information Research Scientist	Senior UX Researcher	PhD	20	80	\$91.25	24	\$9,052.00
Software Developer	System Architect/ Software Engineer	Master's Degree	20	43	\$89.09	38	\$5,286.60
Computer and Information Research Scientist	Software Developer	Master's Degree	2	240	\$45.00	15	\$12,420.00

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.
Dr. Trent is the Owner, President, and CEO of The Bulls Run Group. Dr. Trent is a uniquely skilled researcher with 25+ years of relevant operational expertise.

Labor rate Documentation:

- [Trent rate justification.docx](#)

Direct Labor Cost (\$):

\$90,377.76

Year2

Category	Description	Education	Yrs Experience	Hours	Rate	Fringe Rate	Total
Chief Executive	Principal Investigator	PhD	15	0	\$250.00		\$0.00

Are the labor rates detailed below fully loaded?

YES

Please explain any costs that apply.
There are no costs proposed in Year 2 because this is a 6 month proposal.

Provide any additional information and cost support data related to the nature of the direct labor detailed above.
NA

Direct Labor Cost (\$):

\$0.00

Sum of all Direct Labor Costs is(\$):

\$90,377.76

Overhead
Base

Labor Cost Overhead Rate (%)

20

Apply Overhead to Direct Travel Cost?

NO

Overhead Comments:
IT equipment, software, internet, training, legal and contracting, office expenses

Overhead Cost (\$):

\$18,075.55

Year2

Labor Cost Overhead Rate (%)

20

Apply Overhead to Direct Travel Cost?

NO

Overhead Comments:

Overhead Cost (\$):

\$0.00

Sum of all Overhead Costs is (\$):	\$18,075.55
------------------------------------	-------------

General and Administration Cost Base

G&A Rate (%):	15
---------------	----

Apply G&A Rate to Overhead Costs?	YES
-----------------------------------	-----

Apply G&A Rate to Direct Labor Costs?	YES
---------------------------------------	-----

Apply G&A Rate to ODC- Travel?	YES
--------------------------------	-----

Please specify the different cost sources below from which your company's General and Administrative costs are calculated.

Business development, finance, contract administration, insurance, certifications, accounting software, wireless data plans, collaboration software, HR management system, advertising and marketing, tax preparation, video conferencing, IT equipment

G&A Cost (\$):	\$16,599.95
----------------	-------------

Year2

G&A Rate (%):	15
---------------	----

Apply G&A Rate to Overhead Costs?	YES
-----------------------------------	-----

Apply G&A Rate to Direct Labor Costs?	YES
---------------------------------------	-----

Apply G&A Rate to ODC- Travel?	YES
--------------------------------	-----

Please specify the different cost sources below from which your company's General and Administrative costs are calculated.

G&A Cost (\$):	\$0.00
----------------	--------

Sum of all G&A Costs is (\$):	\$16,599.95
-------------------------------	-------------

ODC-Travel Base

Description: Project kickoff

Location From: Blacksburg, VA

Location To: Arlington, VA

Number of People: 3

Number of Days: 1

Purpose of Trip:

Project Kickoff meeting

Total Airfare Costs (\$): \$0.00

Total Car Rental Costs (\$): \$0.00

Total Per Diem Costs (\$): \$1,138.00

Total Other Costs (\$): \$1,075.00

Total Costs (\$): \$2,213.00

Purpose of Trip:

Project Kickoff meeting

Sources of Estimates:

GSA per diem and mileage rates

Explanation/Justifications: 3 key persons driving in POV for kickoff meeting

Year2

Description: none

Location From: NA

Location To: NA

Number of People: 0

Number of Days: 0

Purpose of Trip:

NA

Total Airfare Costs (\$): \$0.00

Total Car Rental Costs (\$): \$0.00

Total Per Diem Costs (\$): \$0.00

Total Other Costs (\$): \$0.00

Total Costs (\$): \$0.00

Purpose of Trip:

NA

Sources of Estimates:

NA

Explanation/Justifications: NA

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 (%):	10
Profit Explanation:	
Total Profit Cost (\$):	\$12,726.63

Year2

Cost Sharing (\$):	-
Cost Sharing Explanation:	
Profit Rate (%):	10
Profit Explanation:	
Total Profit Cost (\$):	\$12,726.63
Total Proposed Amount (\$):	\$139,992.88

BULLS RUN GROUP, LLC, THE

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.

Privileged and confidential and not subject to disclosure pursuant to 15 U.S.C. 638 (k)(4) and 5 U.S.C. 552.

Total Investments:	Total Sales:	Total Patents:	Government Designated Phase III Funding:
\$250,000.00	\$2,000,000.00	0	\$0.00

Company Information			
Address:	9207 BULLS RUN PKWY, BETHESDA, MD 20817-2403		
SBC Control ID:	SBC_001972933	Company URL:	https://www.bulls.run/

Additional Company Information	
% Revenue for last fiscal year from SBIR/STTR funding:	Total revenue for last fiscal year:
0%	\$100,000 - \$499,999
Year Founded:	# Employees Currently:
2020	10
Year first Phase I award received:	# SBIR/STTR Phase I Awards:
N/A	N/A
Year first Phase II award received:	# SBIR/STTR Phase II Awards:
N/A	N/A
# Employees at first Phase II award:	Mergers and Acquisition within past 2 years:
N/A	No
Spin-offs resulting from SBIR/STTR:	IPO resulting from SBIR/STTR Year of IPO:
No	No
Patents resulting from SBIR/STTR #Patents:	List of Patents:
No	
Woman-Owned:	Socially and Economically Disadvantaged:
No	No
HUBZone-Certified:	SBC majority-owned by multiple VCOC, HF, PE firms By what percent (%):
Yes	No

Additional Investment From		
	Last Submitted Version (2021-05-10 09:48:09)	Current Version
DoD contracts/DoD subcontracts	\$0.00	\$0.00
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

Privileged and confidential and not subject to disclosure pursuant to 15 U.S.C. 638 (k)(4) and 5 U.S.C. 552.

Additional Investment	\$0.00	\$0.00
Total Investment	\$0.00	\$0.00

Phase III Sales To

	Last Submitted Version (2021-05-10 09:48:09)	Current Version
DoD or DoD prime contractors	\$0.00	\$0.00
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	\$0.00	\$0.00

Government Phase III Contracts

	Last Submitted Version (2021-05-10 09:48:09)	Current Version
Funding Obligated	\$0.00	\$0.00

Commercialization Narrative

Commercialized Awards

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

##

CERTIFICATE OF COMPLETION

THIS CERTIFICATE IS PRESENTED TO

Stoney Trent, The Bulls Run Group, Inc

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



Oct 11, 2024

COMPLETION DATE

Oct 11, 2025

EXPIRATION DATE