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:	**F244-0001-0006**
Proposal Title:	**Interactive Knowledge Graphs for Situational Awareness Using Nodal Function Generation**

Agency Information

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

Firm Information

Firm Name:	**Perseides Corp**
Address:	**2309 E 5th St, Austin, TX 78702-4635**
Website:	**nomadik.ai**
UEI:	**KZN9PEK8WW27**
CAGE:	**03A83**
SBA SBC Identification Number:	**002659525**

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)	**3**
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_002659525**

Supporting Documentation:

- SBC_002659525.pdf

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

Signature:

Printed Name	Signature	Title	Business Name	Date
Trevor Sorrells	Trevor Sorrells	CEO	Perseides Corp	09/05/2024

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:

Nodal Function Generation is a noval architecture for building dynamic knowledge graphs at a deeper level than traditional semantic knowledge graph generation. At a fundamental level, information can be thought of as a statement regarding some observed process(es) relating some practical quanta of knowledge. These statements cannot offer elucidation about the objective systems that causally govern the related quanta on their own; rather they must logically support the claims made by sibling informative constructs which reference those same quanta, in order to reinforce their validity, or more precisely, their proximity to an objective truth. Information that is proven to be sufficiently valid within some "objective" frame of reference can then be considered as a form of Intelligence.

Because of the implicit inclusion of uncertainty W.R.T. causal objectivity in both Information and Intelligence, the distinction between the two is more accurately described by the causal bounds in which their configuration is proposed, and the orientation of causal propagation. This is to say, the Information represents some quantized input to a hypothesized system, from which Intelligence can be derived. Furthermore, the end state of some configuration representing Intelligence can act as a quanta of information, or 'Type', within another system and frame of reference as a source of information.

With this in mind, it becomes apparent that any number of Informative quanta, or Nodes, there exists a set of valid Types that represent the attainable (W.R.T. the frame of reference) systems through which Intelligence can be derived. These Types comprise the superposition of configurations that result in novel Nodes that when distilled, can be evaluated for objective efficacy. For example, a proposed configuration of a Type may indicate a causal discrepancy by the degree with which it approximates for other systems, relative to other sibling configurations.

This specific implementation is required to meet the requirements set forth in the topic, namely: updating the graph's underlying ontology, inferring additional edges between nodes, highlighting conflicting information and information gaps, and suggesting additional changes to the graph. The nodal function generation architecture allows users to interact with the dynamic knowledge graph at a granular level and trigger expected responses from the graph that would traditional fail in a graph built on top of a semantic implementation.

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

If the project is successful there will be a resulting paradigm shift not just in how users can interact with dynamic knowledge graphs, but in the intelligence process as a whole. The nodal function generation architecture removes the burden of needing to know exact data types, data structures, and expected values. Beyond that, the automatic correlation and pattern recognition between disparate datasets, more sophisticated than the current standard of semantic matching, will enable analysts to uncover hidden patterns critical to modern conflicts. Automatically identifying conflicting information and pointing out information gaps is a critical capability due to the evolving nature of 5th generation warfare and the downstream information and psywar.

In terms of commercial application Perseides is already in the process of offering a commercial solution integrated with the Nomadik civil intelligence platform being developed for cities and states to better coordinate their city services and associated data - starting with the homeless crisis. Cities currently face a myriad of crises that require coordinated data and resources to properly address but cities currently lack a solution to achieve the level of coordination required. The Nomadik civil intelligence platform functions as a common operating picture for cities to put every stakeholder on the same page to allow them to more efficiently and effectively carry out their responsibilities. The nodal architecture allows cities to uncover hidden patterns and correlations that allow them to utilize their limited resources to produce outsized impacts.

While the initial focus of our commercial application is homelessness Perseides Corp has plans to expand into the entire civil intelligence realm: human trafficking, drug trafficking, gang activity, public safety, etc. Outside of civil intelligence this architecture has applications across every industry that uses knowledge graphs, or will use knowledge graphs, to achieve their goals. From finance, to real estate, to insurance, to retail and more. The research and development enabled by this SBIR allows Perseides Corp to dedicate time to developing a noval capability with the potential to transform how industries and organizations understand their data.

Attention:

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:

2 3 4 5 6 7 8 9 10 11 12 13 14

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

dynamic knowledge graph, Artificial intelligence, knowledge graph, ontology, machine learning, AI/ML, nodal function generation, graph node

VOL I - Proposal Certification

Summary:

1. At a minimum, two thirds of the work in Phase I will be carried out by your small business as defined by 13 C.F.R Section 701-705. The numbers for this certification are derived from the budget template. To update these numbers, review and revise your budget data. If the minimum percentage of work numbers are not met, then a letter of explanation or written approval from the funding officer is required. Please note that some components will not accept any deviation from the Percentage of Work (POW) minimum requirements. Please check your component instructions regarding the POW requirements.	**YES**
Firm POW	**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.	**NO**
8. Has a proposal for essentially equivalent work been submitted to other US government agencies or DoD components?	**NO**
9. Has a contract been awarded for any of the proposals listed above?	**NO**
10. Firm will notify the Federal agency immediately if all or a portion of the work authorized and funded under this proposal is subsequently funded by another Federal agency.	**YES**
11. Are you submitting assertions in accordance with DFARS 252.227-7017 Identification and assertions use, release, or disclosure restriction?	**YES**

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. 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.	**YES**
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?	**100%**
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?	**NO**

VOL I - Contact Information

Principal Investigator

Name:	**Morgan Winters**
Phone:	**(470) 261-9222**
Email:	**morgan@nomadik.ai**
Address:	**2309 E 5th St, Austin, TX 78702 - 4635**

Corporate Official

Name:	**Trevor Sorrells**
Phone:	**(770) 317-1526**
Email:	**trevor@nomadik.ai**
Address:	**2309 E 5th St, Austin, TX 78702 - 4635**

Authorized Contract Negotiator

Name:	**Trevor Sorrells**
Phone:	**(770) 317-1526**
Email:	**trevor@nomadik.ai**
Address:	**2309 E 5th St, Austin, TX 78702 - 4635**

Form Generated on 11/05/2024 04:24:03 PM

Interactive Knowledge Graphs for Situational Awareness Using Nodal Function Generation

Volume 2: Technical Volume

1. Identification and Significance of the Problem or Opportunity.

Warfare is no longer about who has the best equipment, who has the most manpower, or who has the most money, warfare in the 21st century is about decision advantage - who has the best intelligence. The United States has spent decades building the organizations required to collect and analyze data from every domain imaginable, but the process to derive actionable intelligence from an ever growing ocean of information is slow, and the second most important factor in any modern conflict is time. Knowledge graphs were a stepping stone towards a solution that aimed to reduce the time and manpower required to turn actionable intelligence into decisive action. However, the current paradigm of AI powered knowledge graphs has glaring flaws that hamstring the potential power of these knowledge graphs.

Knowledge graphs are built upon the data they are fed, which results in inherent human error originating downstream. These errors manifest themselves in incomplete and inaccurate data that results in inaccurate knowledge graphs. Missing links, hidden patterns, and incorrect correlations that need to be fixed by human analysts to ensure the decisions that are made with these knowledge graphs are as well informed as possible. Current knowledge graphs however lack a way to reorient themselves as data is corrected, added, or removed. Requiring either a complete rebuilding of the graph and in a time constrained environment this leads to decision lag and the consequences that come with it.

Dynamic knowledge graphs are the next logical step in the technological development process. Providing both the interface for analysts to correct data and add or remove nodes and edges as well as the underlying infrastructure to highlight conflicting information and missing information while inferring new edges and nodes to provide a more comprehensive picture. For analysts to have the clearest situational awareness picture of their operating environment, to understand the patterns of life on the battlefield, and to enable effective threat detection and targeting operations the dynamic knowledge graph needs to be able to learn and evolve with the analyst.

The opportunity to build a system that doesn't just enable ease of user intervention but uses that input to learn and understand context, to suggest real time updates and fill gaps that the user may not have even considered is one that can change the face, and the pace, of modern conflict. The importance of a system such as this is not just about speeding up legacy workflows, but rethinking how we understand and act on the intelligence gathered by the vast array of systems and organizations throughout the DoD. By building a system that learns, adjusts, and evolves alongside its human operators, we create a force multiplier—a tool that doesn’t just store data but thinks with us, advancing our situational awareness and cutting the time to action dramatically.

2. Phase I Technical Objectives.

The technical objectives for the Phase I period of performance will focus on the foundational design and technological feasibility for Persiedes’ nodal function generation and how it provides the stated capability requirements outlined for the topic. This phase is critical in ensuring that the nodal function generation enables the interactions outlined by Air Force requirements not currently possible with traditional knowledge graph architectures. During this phase Perseides Corp will conduct comprehensive research and development sprints aiming to define and test the nodal function generation and supporting architecture.

The specific technical objectives for Phase I, along with the pivotal questions to be addressed are as follows:

Evaluate Data Ingestion Modularity
Objective: A key component to any system that relies on multiple data streams is being able to easily and efficiently parse new data types into the system.

Key Questions: How accurately does the system classify existing and new data types? How well does the system infer types from unstructured text fields?

Our proposed system will consist of a central interface and arbitration service, managing several containerized database systems. Each system will be of a unique engine and/or locality (providing distributed storage flexibility as needed), for the optimal ingestion and handling of a variety of data types. The two required database systems will consist of:

- Extended PostGreSQL Container
  - Ingestion of any/all tabular and/or ambiguously Typed (W.R.T. existing Nodal representations) data.
  - Implicit ingestion of unambiguously Typed and externally stored data (i.e. imagery, literature, or other media)
  - System utilities
- TypeDB Container
  - Ingestion and storage of ambiguous Type representations, or Knowledge Graphs.
  - Indexed storage of hypothetical Type representations and the causal ‘edges’ that relate them to other graphs.

The central system itself will arbitrate ingestion of all data, orchestration of workers, systemic testing, serving of API/HMI, and other utilities. As novel types are introduced, workers will be tasked with manipulating these representations in an effort to construct testable representations of the system that produced the ingested quanta of an implied Type.

Evaluate Nodal Generation for Edge Discovery
Objective: Determine the accuracy of edge discovery in the nodal function generation system.

Key Questions: Given the sub-nodal structure how well does the system discover edges compared to current semantic systems?

Unlike the semantic representation of relational information between nodes, our proposed system relies on causally structured and testable sets of representations of individual nodes themselves. These representations are comparable to knowledge sheaves, albeit highly contextualized and inherently ambiguous in terms of face value validity. Each node conceptually represents a hypothesized ideal ‘Type’, the causal form of which can be elucidated by mapping out the systems that (allegedly) produce quanta of said Type. These ideal Types should be thought of as simultaneously complex and axiomatic, the distinction only becoming relevant when considering a specific representation of a Type, which can be thought of as a system in which specific input and output Types are desirable and therefore immutable.

Because of these causal underpinnings in our architecture, we hypothesize that the abundance or lack of representations referring to a hypothetical Type speak to its objective validity, and therefore can be used in conjunction with the generation of novel representations of relationships, as well as the augmenting the discovery of approximator systems via morphological analysis.

Evaluate Conflicting Information and Information Gap Discovery
Objective: Determine the accuracy of the nodal architecture in detecting conflicting information and information gaps.

Key Questions: How well does the system do in detecting inaccuracies at both the semantic and sub-nodal level? At what level of information gap is the system capable of detecting (e.g. if our data consistently references crime without a crime dataset plugged in will it be able to find that gap)?

Once the infrastructure supporting the workers traversal of the data structures is completed, we anticipate the 'priming' of known Types, both pragmatically macroscopic and microscopic. These known relationships will serve to inform the exploration of supplied data (i.e. nodal generation). From similar small scale tests, we anticipate the indexing of implicit but unstable Types to become statistically apparent, the recovery of which to depend upon the specific nature of ancillary data.

Furthermore, when two or more representations of a Type conflict causally (i.e. data sources providing conflicting information), this too becomes very quickly apparent, and the Type(s) and source(s) in question are flagged for re-evaluation.

Evaluate Graph Change Suggestions Based on User Input
Objective: Utilize user input to inform graph exploration to suggest further changes informed by edges that already exist from the edited node as well as suggested edges after user input.

Key Questions: How quickly does the system utilize user input to both discover new connections as well as suggest further changes to the graph? What is the scope of a suggested change based on different types of user input?

Fundamentally, each of these operations by the user equates to the ingestion of a supposed graph into the system. As each sub-graph is architecturally synonymous with a hypothetical representation of a Type, the system will infer user defined ontologies and schemas in the same fashion as the automated processes index hypothetical Types. As the supposed graph representation is explored, the measures of its 'validity' are tested in conjunction with known Types, and the morphological distinctions as it grows / shrinks.

Design User Interface
Objective: Design a familiar and intuitive knowledge graph UI that enables the following user inputs and user feedback from the system:

- allow a user to interact with a dynamic knowledge graph by making changes and additions to the knowledge graph.
- displaying inferred additional edges between nodes
- highlighting conflicting information in the graph
- highlighting information gaps
- suggesting additional changes to the graph

Key Questions: How should information from a node or edge be displayed to users that generalizes across various data types? What is the optimal workflow for each user interaction? How should feedback from the system be displayed to the user?

Our proposed system will serve API/HMI functionality, for the exploration of the nodal representations of specific Types, akin to specific relations between user-defined axiomatic Types. Visually, this will be accomplished by building off of existing libraries, and we anticipate an interface built upon frameworks like networkx.

Architect User Interface to Nodal Graph API
Objective: Connect the user interface with the nodal graph backend via API.

Key Questions: How should additions, removals, and edits be passed from the UI to the nodal graph backend to enable rapid processing and user feedback?

Our proposed system will continuously and automatically assign workers to explore ingested data to construct new nodal representations, as well as flagging causal ‘dead-ends’, until such a time when morphological analysis implies that a change in criticality has occurred. As all incoming data will be indexed by these workers, the API will support a priority scheme where defined, allowing users to request priority for the exploration of a subset of data, or request an immediate re-evaluation of a flagged sub-graph.

Build Test Suite to Evaluate Performance Metrics
Objective: Build out a test suite to evaluate performance metrics such as accuracy, graph completeness, edge inference accuracy, information gap accuracy, and change suggestion accuracy.

Key Questions: What sort of graph structure is required to accurately test the aforementioned metrics? How do data sources, data types, and data relevance distance affect the various metrics? How do we measure information gap and change suggestion accuracy? What size truth graph is required to achieve 80%, 90%, 95%, and 99% coverage of edge case testing?

Our proposed system will handle continuous system wide testing and evaluation, providing real-time reporting for system diagnostics, graph validation, and data type ‘asks’ suggested by the system itself. Each evaluation pipeline will adhere to a documented schema, from which additional tests can be defined and incorporated.

In addition, where the condensation/substitution of Types is performant (i.e. a containerized machine learning service would pragmatically optimize a particular causal transformation or system of transformations), the system will present these findings.

3. Phase I Statement of Work

The work will be broken into two focus areas, the front-end user interface (UI) to interact with the knowledge graph itself, and the back-end knowledge graph generation and maintenance facilitated through our nodal function generation approach. The overall UI will take the form of a traditional knowledge graph interface, with nodes and edges representing discrete information and their relationship to other pieces of information.

Task Outline:

1. **Task 1: Data Ingestion Design and Development**
    a. ***Objective***: Design a system that gracefully handles data input to create sub-nodal representations of data fields for sub-nodal edge detection.
    b. ***Activities***:
        i. Establish and Implement Data Storage Infrastructure
            1. Containerized Storage Engines
                a. PostGreSQL
                    i. Temporary storage of ingested raw and tabular data
                    ii. Storage of intermediately typed ingested data
                    iii. Storage of any/all tabular data concerning Type composition
                    iv. Worker queuing / logging
                    v. System Operations
                b. TypeDB
                    i. Storage of Typed Nodal Structures, i.e. Knowledge Graphs
                    ii. Storage of sub-graphs representing the assumed superpositions of a complex Type
            2. Containerized Management and Storage Interface System
                a. Managing workers, logging, and API/HMI
                b. Responsible for ingestion pipelines and distributed storage interfacing
        ii. Implement Framework for Nodal Ingestion and Type Approximation
            1. Establish basic ingestion of existing nodal structures of inherently ambiguous axiomatic granularity
            2. Development, testing, and restructuring of heuristic ontologies of algorithms for approximating Types **
            3. Initial implementation of categorical Type compression/decompression indicators **

iii. Prune and Refine Type Approximation
1. Implement/refactor worker processes and management infrastructure where necessary
2. Develop scheduling for systemic testing and benchmarking

2. Task 2: Nodal Edge Discovery Design and Development

a. ***Objective***: Design the edge discovery based on nodal and sub-nodal connections that performs at least as well as traditional semantic edge discovery.

b. ***Activities***:

i. Implement Categorical Ranking of Nodal Structures
1. Develop framework for relating superpositions of potentially related Types via causal morphology
2. Develop indexing and ranking of testable graphs, modulating 'compression' to optimize for end-node performance

ii. Prune and Refine Categorical Ranking
1. Implement/refactor worker processes and management infrastructure where necessary
2. Develop scheduling for systemic testing and benchmarking
3. Implement initial HMI for visualizing and interacting with ranked graphs
4. Large-scale testing of labeled causal system representations **

iii. Deploy Categorical Causal Graph Association
1. Continue large-scale testing of labeled causal system representations, masking testable sub-graphs
2. Rank and vetting of proposed substitutions **
3. Indexing of macroscopic causal ranking (composite Types, and supposed 'edges')
4. Initial testing of proposed edge elucidation methods for HMI **

3. Task 3: Conflicting Information and Gap Discovery Design and Development

a. ***Objective***: Design the processes required to find conflicting information within the high level data nodes of the graph and highlight gaps in information based on the current understanding of the graph.

b. ***Activities***:

i. Initiate Knowledge ‘Priming’
1. Large scale ingestion of known and engineered Types, in order to prime the morphological and causal indices
2. Execute tests across multiple deployed instances operating with seeded subsets of such Types.

ii. Develop and Test Conflict Awareness
1. Repeat priming procedures across multiple instances with various degrees of conflicted data
2. Record morphological behavior and subsequent implications of cascaded conflict

3. Formalize metrics for identification of conflicting data sources among intentionally noised data.

iii. Refine Morphological Gap Detection Techniques
1. Repeat priming procedures with distinctly masked or conflicted data, and optimize exploratory algorithms for disambiguation
2. Develop initial pipelines for flagged Types

4. **Task 4: Graph Change Suggestions from User Input Design and Development.**
   a. ***Objective***: Develop the API to handle user input and the processes required to explore the graph and suggest additional changes based on the input.
   b. ***Activities***:
      i. Develop Composite Operation Pipelines
         1. Implement 'usher' workers for the execution of tasks and retrieval of information
         2. Implement queueing for parallelized task execution
      ii. Define API
         1. Formalize all composite tasks to be made from the host, and assign task schedules and endpoints
         2. *Conventional API Development
         3. Develop tests, and execute Type ingestion

5. **Task 5: User Interface Design and Development**
   a. ***Objective***: Design an intuitive and user-friendly interface for the dynamic knowledge graph to enable editing and highlighting of information gaps and discrepancies.
   b. ***Activities***: Develop the graphical user interface; connect the user interface to the back-end dynamic knowledge graph architecture via API; ensure user experience editing data and knowledge graph feedback is intuitive via use case testing.

6. **Task 6: Comprehensive System Testing and Validation**
   a. ***Objective***: Conduct extensive testing of the dynamic knowledge graph to ensure accuracy, reliability, and ease of use.
   b. ***Activities***: Develop a truth graph to then test a graph with artificial gaps, discrepancies, and changes; Develop test system to apply the set of changes required to get to the truth graph in random order to ensure edge case coverage; develop scoring metrics to measure accuracy of graph after edits compare to truth graph.

   With the infrastructure, API, and UI constructed, we will continue evaluating the performance of aforementioned testing and benchmarking routines, as well as develop and formalize the breadth of priming and associated effects. In addition, we hope to distill these components, such that iterative computation and data requirements to achieve validation are minimized.

Milestone Schedule:
The following are the key milestones in Phase I:

Milestone	Expected Delivery	Deliverable	Acceptance Criteria	Performer
Data Ingestion Design and Development	Award + 1 Month	Delivery of written summary of data ingestion pipeline.	DAF TPOC receives the written summary via email.	Perseides Corp
Nodal Edge Discovery Design and Development	Award + 2 Month	Delivery of written summary of nodal edge discovery.	DAF TPOC receives the written summary via email.	Perseides Corp
Conflicting Information and Gap Discovery Design and Development	Award + 3 Month	Delivery of written summary of conflicting information and graph discovery process.	DAF TPOC receives the written summary via email.	Perseides Corp
Graph Change Suggestions from User Input API Design and Development.	Award + 4 Month	Delivery of written summary of user input API change suggestions.	DAF TPOC receives the written summary via email.	Perseides Corp
User Interface Design and Development	Award + 5 Month	Delivery of written summary of user interface.	DAF TPOC receives the written summary via email.	Perseides Corp
Comprehensive System Testing and Validation	Award + 6 Month	Delivery of written summary of system testing and validation along with prototype of system.	DAF TPOC receives the written summary via email.	Perseides Corp

Deliverables:

The project deliverables for Phase I are:

- Monthly progress reports.

- Kickoff and review meeting presentations.
- Media (pictures and videos) of the dynamic knowledge graph demonstrations.
- Prototype of system.
- Final report with SF 298.

4. Related Work.

The principal investigator is currently working on the nodal function generation architecture in support of the commercial offering by Perseides Corp, namely the Nomadik civil intelligence platform. Current efforts are focused on ingesting disparate and noisy data sources, and using purpose built transformation schemes to uncover knowledge gaps, opaque relationships, and conflicting information. As these processes work to construct a traversable causal graph, they also indicate exploitable morphisms between seemingly disparate and conventionally incoherent causal chains. These indications then are used to direct algorithmic strategy when approximating for problematic nodes.

Our commercial data specifically concerns homelessness and its interactions with a host society, and a prime example of this process in action concerns the reporting of illicit drug activity. What constitutes a valid indication of substance abuse within an encampment community or transient group, and the methods of its detection, are inherently complex and dynamic systems in and of themselves. Our users, who are able to report upon such activity, provide raw data in the form of unstructured and structured data alike, as well as the relevant imagery and video providing supposed ‘proof’ of such activity; each of these pipelines naturally incurring unique information transformations, predicated by phenomena such as individual bias / experience, network robustness, encampment proximity and visibility, and the dynamism of substance identification and classification. Each of these nodal adjacencies and their contribution in coalescing an empirical function with which to gauge the likelihood of substance abuse can be elucidated by observing the causal morphology of other systems, which indicate the nature of a knowledge gap can be ascertained (this is to say, the Nodes related by a noisy relationship) by evaluating similar systems through causal and semantic inference.

In the case of deriving effective metrics for quantifying the aforementioned relationships between user-reporting and the objectivity of incoming data, the semantic distinctions between observations can be traced, and furthermore the causation such distinction can be modeled; in effect, novel nodes can be generated, and their utility serves to reinforce or conflict other hypotheses. This can be seen when the observation of novel drugs within a region commences, from which existing semantic distinctions are reinforced, and the causal relationships that define them are linked to new gaps within the macroscopic graph (general substance abuse within a populated region).

While the primary use case for the nodal architecture is to look for useful patterns and correlations for organizations to utilize in their piece of addressing homelessness, the dynamic knowledge graph user interface allows users to further explore the data. Data surrounding homelessness is inherently noisy, thus users will need to make corrections and have the supporting architecture that reorients patterns and correlations to effectively apply their limited resources. This is further enabled with the architectures ability to point out knowledge gaps and make further suggestions that refine the information landscape for all organizations involved.

5. Relationship with Future Research or Research and Development.

(a) *State the anticipated results of the proposed approach if the project is successful.*

If the project is successful there will be a resulting paradigm shift not just in how users can interact with dynamic knowledge graphs, but in the intelligence process as a whole. The nodal function generation architecture removes the burden of needing to know exact data types, data structures, and expected values. Beyond that, the automatic correlation and pattern recognition between disparate datasets, more sophisticated than the current standard of semantic matching, will enable analysts to uncover hidden patterns critical to modern conflicts. Automatically identifying conflicting information and pointing out information gaps is a critical capability due to the evolving nature of 5th generation warfare and the downstream information and psywar.

(b) Discuss the significance of the Phase I effort in providing a foundation for a Phase II research or research and development effort.

Phase I is critical in allowing Perseides Corp to conduct the research and development required to prove our novel thesis on nodal function generation. The Phase I effort provides the time, and non-commercial capital, to prove out a critical capability that serves both the Air Force end user as well as our commercial customers. At Perseides we think about information, and intelligence, in a fundamentally different way than previous attempts at constructing knowledge graphs.

At a fundamental level, information can be thought of as a statement regarding some observed process(es) relating some practical quanta of knowledge. These statements cannot offer elucidation about the objective systems that causally govern the related quanta on their own; rather they must logically support the claims made by sibling informative constructs which reference those same quanta, in order to reinforce their validity, or more precisely, their proximity to an objective truth. Information that is proven to be sufficiently valid within some "objective" frame of reference can then be considered as a form of Intelligence.

Because of the implicit inclusion of uncertainty W.R.T. causal objectivity in both Information and Intelligence, the distinction between the two is more accurately described by the causal bounds in which their configuration is proposed, and the orientation of causal propagation. This is to say, the Information represents some quantized input to a hypothesized system, from which Intelligence can be derived. Furthermore, the end state of some configuration representing Intelligence can act as a quanta of information, or ‘Type’, within another system and frame of reference as a source of information.

With this in mind, it becomes apparent that any number of Informative quanta, or Nodes, there exists a set of valid Types that represent the attainable (W.R.T. the frame of reference) systems through which Intelligence can be derived. These Types comprise the superposition of configurations that result in novel Nodes that when distilled, can be evaluated for objective efficacy. For example, a proposed configuration of a Type may indicate a causal discrepancy by the degree with which it approximates for other systems, relative to other sibling configurations.

Phase II will allow Perseides to take the research and development from Phase I and build a system that satisfies all of the Air Force end users requirements from both the UI/UX experience and the underlying architecture that provides the most accurate dynamic knowledge graph by expanding well beyond traditional semantic connections.

(c) Identify the applicable clearances, certifications and approvals required to conduct Phase II testing and outline the plan for ensuring timely completion of said authorizations in support of Phase II research or research and development effort.

Given that both Phase I and Phase II require participants to utilize their own datasets no clearances, certifications, or approvals are required for the Phase II effort. Trevor Sorrells, one of the key personnel, currently holds a Secret clearance. Should that be required Perseides Corp will work to find a suitable partner to hold his clearance as Perseides Corp does not have a facilities clearance.

6. Commercialization Strategy.

Perseides is already in the process of offering a commercial solution integrated with the Nomadik civil intelligence platform being developed for cities and states to better coordinate their city services and associated data - starting with the homeless crisis. Cities currently face a myriad of crises that require coordinated data and resources to properly address but cities currently lack a solution to achieve the level of coordination required. The Nomadik civil intelligence platform functions as a common operating picture for cities to put every stakeholder on the same page to allow them to more efficiently and effectively carry out their responsibilities.

The key enabler of this coordination of data and services is the nodal function generation outlined in section 4. By ingesting data across organizations and standardizing formats through the use of sub-nodal structures, Perseides is able to more effectively build knowledge graphs that uncover hidden correlations and patterns. These correlations and patterns can then be utilized by participating organizations to more effectively address their piece of the puzzle, in our initial case, homelessness.

The issue identified by Perseides of a lack of coordination among aligned organizations due to data silos and fragmented information is not exclusive to cities. The ability to build coordinated knowledge graphs among any group of organizations that need to share data and ultimately derive actionable intelligence is a capability that can be exported to any industry, especially among DoD and IC organizations.

7. Key Personnel.

Morgan Winters - Principal Investigator
Kennesaw State University, BS Electrical Engineering, 2021, ML Engineer Modern Intelligence, PLC Engineer Knapp

Relevant Experience
At Perseides, Morgan is the CTO in charge of building the aforementioned nodal structure to support the meta pattern matching functionality utilized by city and state customers to find

correlations, patterns, and predictions between participating organizations data, expanded on in section 4 relevant work. Previously at Modern Intelligence, as a senior machine learning engineer, Morgan was in charge of implementing the optics library required to place objects detected with computer vision models in a monocular EO/IR camera system in real space. Alongside the optics library Morgan led research for novel self-supervised machine learning models to turn 2D images into 3D objects in support of the target re-identification efforts.

Relevant Awards or Patents

- Morgan is a co-author, along with Trevor Sorrells, on a patent currently in the submission process at Modern Intelligence titled “Object recognition, geolocating, and tracking using optical systems”.

Trevor Sorrells
Georgia Institute of Technology, BS Computer Science, 2019, Lead Engineer Modern Intelligence, Research Engineer GTRI, Embedded Engineer Sandia National Labs

Relevant Experience
At Perseides, Trevor is the CEO and engineer in charge of the mobile component of the Nomadik civil intelligence platform, providing users with an efficient and coordinated way to submit information to the platform and access relevant intelligence in the field. Previously at Modern Intelligence, Trevor was the lead engineer on the maritime domain awareness product in charge of building and deploying object recognition and re-identification machine learning models. Alongside the machine learning models Trevor was responsible for the multi-target tracker sensor fusion system to associate tracks from EO/IR camera systems, radars, AIS transponders, satellite data, and other sensors providing fused target location data to operators.

Relevant Awards or Patents

- Trevor is a co-author, along with Morgan Winters, on a patent currently in the submission process at Modern Intelligence titled “Object recognition, geolocating, and tracking using optical systems”.
- GTRI TEN Division Adaptable Entrepreneur Award

Austin White
Wichita State University, BS Computer Science, 2018, Software Engineer Modern Intelligence, Lockheed Martin, & Microsoft

Relevant Experience
At Perseides, Austin is the lead engineer on the Nomadik civil intelligence platform, building out the UI/UX experience for customers within cities to coordinate and understand patterns within combinations of their datasets. Previously at Modern Intelligence, Austin built the UI/UX for Modern’s AI maritime domain awareness product for US Navy and Airforce users. Prior to Modern, Austin led the development of a full stack web interface for satellite command and operation at Lockheed Martin. Before that Austin led development for a full stack web dashboard for viewing and distributing data results from internal testing applications at Microsoft. Austin has extensive experience building intuitive UI/UX experiences for customers ensuring they have a fast and easy to use interface to work with complex software products.

Relevant Awards or Patents

- Boeing Award of Excellence
- IEEE Wichita Chapter Best Engineering Project
- Koch Industries Most Entrepreneurial Engineering Project
- Millennium Concepts Best Project
- Shocker New Venture Trade Show 1st Place

8. Foreign Citizens.

None

9. Facilities/Equipment.

Perseides works out of their primary office space with both local and cloud compute available to support the research and development of the project.

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.

DSIP Defense SBIR/STTR Innovation Portal

Cost Volume

SBIR Phase I Proposal

Proposal Number	F244-0001-0006
Topic Number	AF244-0001
Proposal Title	Interactive Knowledge Graphs for Situational Awareness Using Nodal Function Generation
Date Submitted	11/05/2024 04:24:00 PM

Firm Information

Firm Name	Perseides Corp
Mail Address	2309 E 5th St, Austin, Texas, 78702
Website Address	nomadik.ai
UEI	KZN9PEK8WW27
Cage	03A83

Total Dollar Amount for this Proposal	$139,736.00
Base Year	$109,808.00
Year 2	$29,928.00
Technical and Business Assistance(TABA)- Base	$0.00
TABA- Year 2	$0.00

Base Year Summary

Total Direct Labor (TDL)	$129,808.00
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 10%) x Base ()	$0.00
Total Firm Costs	$129,808.00
Subcontractor Costs	
Total Subcontractor Costs (TSC)	$0.00
Cost Sharing	-$20,000.00
Profit Rate (0%)	$0.00
Total Estimated Cost	$109,808.00
TABA	$0.00

Year 2 Summary

Total Direct Labor (TDL)	$54,928.00
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 10%) x Base ()	$0.00
Total Firm Costs	$54,928.00
Subcontractor Costs	
Total Subcontractor Costs (TSC)	$0.00
Cost Sharing	-$25,000.00
Profit Rate (0%)	$0.00
Total Estimated Cost	$29,928.00
TABA	$0.00

Base Year

Direct Labor Costs

Category / Individual-TR	Rate/Hour	Estimated Hours	Fringe Rate (%)	Fringe Cost	Cost
Software Developer/ Principal Investigator (Morgan Winters)	$72.00	960	30	$20736.00	$89,856.00
Software Developer/ Front end developer (Austin White)	$72.00	320	30	$6912.00	$29,952.00

Subtotal Direct Labor (DL)	$119,808.00
Labor Overhead Cost	$10,000.00
Total Direct Labor (TDL)	**$129,808.00**

G&A (rate 10%) x Base ()	$0.00
Cost Sharing	-$20,000.00
Profit Rate (0%)	$0.00
Total Estimated Cost	$109,808.00
TABA	$0.00

Year 2

Direct Labor Costs

Category / Individual-TR	Rate/Hour	Estimated Hours	Fringe Rate (%)	Fringe Cost	Cost
Software Developer/ Principal Investigator (Morgan Winters)	$72.00	480	30	$10368.00	$44,928.00

Subtotal Direct Labor (DL)	$44,928.00
Labor Overhead Cost	$10,000.00
Total Direct Labor (TDL)	**$54,928.00**

G&A (rate 10%) x Base ()	$0.00
Cost Sharing	-$25,000.00
Profit Rate (0%)	$0.00
Total Estimated Cost	$29,928.00
TABA	$0.00

Explanatory Material Relating to the Cost Volume

The Official From the Firm that is responsible for the cost breakdown

Name: Trevor Sorrells
Phone: (770) 317-1526
Phone: trevor@nomadik.ai
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
Software Developer	Principal Investigator	Bachelor's Degree	5	960	$72.00	30	$89,856.00
Software Developer	Front end developer	Bachelor's Degree	9	320	$72.00	30	$29,952.00

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.
The rate and hours covers the salary for the expected work required to deliver the solution within 6 months.

Direct Labor Cost ($): $119,808.00

Year2

Category	Description	Education	Yrs Experience	Hours	Rate	Fringe Rate	Total
Software Developer	Principal Investigator	Bachelor's Degree	5	480	$72.00	30	$44,928.00

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.
The hours and rate cover the follow-on work likely required to support and improve the solution deliverable.

Direct Labor Cost ($): $44,928.00

Sum of all Direct Labor Costs is($): $164,736.00

Overhead

Base

Labor Cost Overhead Cost **$10,000.00**

Overhead Comments:
Overhead for this contract will primarily be compute costs associated with the deliverable.

Overhead Cost ($): **$10,000.00**

Year2

Labor Cost Overhead Cost **$10,000.00**

Overhead Comments:
Overhead for this contract will primarily be compute costs associated with the improvement of the deliverable.

Overhead Cost ($): **$10,000.00**

Sum of all Overhead Costs is ($): **$20,000.00**

General and Administration Cost

Base

G&A Rate (%): **10**

Apply G&A Rate to Overhead Costs? **NO**

Apply G&A Rate to Direct Labor Costs? **NO**

Please specify the different cost sources below from which your company's General and Administrative costs are calculated.
G&A costs primarily come from software and rent: - RIppling payroll - Google suites - Slack

---

G&A Cost ($): **$0.00**

Year2

G&A Rate (%): **10**

Apply G&A Rate to Overhead Costs? **NO**

Apply G&A Rate to Direct Labor Costs? **NO**

Please specify the different cost sources below from which your company's General and Administrative costs are calculated.
G&A costs primarily come from software and rent: - RIppling payroll - Google suites - Slack

---

G&A Cost ($): **$0.00**

Sum of all G&A Costs is ($): **$0.00**

Profit Rate/Cost Sharing

Base

Cost Sharing ($): **-$20,000.00**

Cost Sharing Explanation:
The development of the solution works towards our own feature road map for our civil intelligence platform. The knowledge graph is UI is more specific to the contract, but the back end logic has overlap with features we want to develop for customers.

Profit Rate (%): **0**

Profit Explanation:
Currently in contract negotiations with initial customers, but no revenue to report.

Total Profit Cost ($): **$0.00**

Year2

Cost Sharing ($): **-$25,000.00**

Cost Sharing Explanation:
The development of the solution works towards our own feature road map for our civil intelligence platform. The knowledge graph is UI is more specific to the contract, but the back end logic has overlap with features we want to develop for customers.

Profit Rate (%): **0**

Profit Explanation:
Currently in contract negotiations with initial customers, but no revenue to report.

Total Profit Cost ($): **$0.00**

Total Proposed Amount ($): **$139,736.00**

CERTIFICATE OF COMPLETION

THIS CERTIFICATE IS PRESENTED TO

Trevor Sorrells, Perseides Corp

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

Nov 05, 2024
COMPLETION DATE

Nov 05, 2025
EXPIRATION DATE