Topic: DON26BX05-NP003 - NAVSEA Open Topic for Model-Based Systems Engineering (MBSE) Acceleration via Artificial Intelligence or Machine Learning Data Ingestion

1) The topic requires direct conversion to MBSE models in Cameo MagicDraw. Would it be acceptable for an external application to generate a SysML model that is then imported into Cameo, or is the intent for the conversion capability itself to operate directly within Cameo? Also, is there a specific Cameo/MagicDraw version and SysML version the NAVY expects the solution to support? Understanding there are multiple tools that can auto generate sysML models which; therefore, would make the topic not innovative if it were the sole objective of the SBIR. With that in mind, the goal of this SBIR is for tool / tech developed to be cognizant of existing models and model structure (if available), the target model schema (or have the ability to ingest and model model in multiple schemas), and be able to generate updated model components, or new models in the desired schema that work with the latest versions of MagicDraw (not just one). Ideally the product should be able to update models with a high degree of accuracy, alignment to existing model style, data structures, and schema, with minimal human intervention based on data inputs from system scans, documentation, requirements, PCAP / NetFlow, and other disparate design and operational data sets.
2) **Cameo integration:** Does "direct conversion" require model creation through a specific Cameo/MagicDraw API, plugin, import/export format, or SysML interchange mechanism, or may offerors propose the most appropriate integration architecture? Offerors are invited to propose efficient / appropriate creation and ingestion mechanisms; however, to the extent possible should avoid re-developing interfaces or capabilities native to the Cameo / MagicDraw tool. Additionally, proposed methods should align to guidance in 1 above and require minimal human interaction to produce accurate results.
3) **SysML version:** Is NAVSEA currently targeting a particular SysML version or Cameo/MagicDraw baseline for generated models? Capabilities should work with multiple versions of Cameo / MagicDraw
4) **Phase I data:** May Phase I feasibility be demonstrated using representative unclassified or synthetic legacy technical documentation when actual AEGIS technical data is unavailable? Demonstrations may be proposed to be executed with legacy data provided it is of a similar complexity but should be representative of the problem to be resolved.
5) **Model accuracy:** Are there specific accuracy, completeness, or validation metrics NAVSEA intends to use to assess automatically generated SysML models? We have not published a set of assessment metrics for the tool and product quality assessment, but will likely assess based on accuracy, alignment to desired model structures, compatibility / interoperability, flexibility, ease of use / LOE required to produce quality products, etc.
6) **Human involvement:** The topic says both that the system should automate model development with "no human input" and that finished models may require human modification. Can you please help explain the difference? Development (use of tool) and integration of finished products should require minimal human input.
7) **Commercial AI use:** Can a commercial AI model be used during Phase I with public, synthetic, or other non-sensitive data, with the ability to switch to a Navy-approved AI model in Phase II if required? Commercial AI can be used for proof of concept with

representative data from similarly complex systems; however, system design data should not be provided to commercial AI due to compilation security concerns.

8) **Phase II AI environment:** In Phase II, does the AI need to run inside the Navy's environment or can we use a Government-approved external AI service? No determination has been made yet. Use of other secure AI services will be assessed based on risk profile and sensitivity of data.
9) **Government test data:** Will the Navy provide representative legacy technical documents or other system data for Phase I or Phase II testing? If not, may we use synthetic data that represents the types of Navy documents listed in the topic? The government will NOT provide AEGIS, or other GFI data for the phase 1 portion of this SBIR, nor licenses for CAMEO Magic Draw. Synthetic data, or data from another system of similar complexity may be used for phase 1 demonstrations. The Government reserves the option to provide or request offerors utilize synthetic data again in phase 2; however, no decision has been made yet.
10) What specific SysML version, Cameo MagicDraw version, and model exchange format (MDZIP, XMI, API, etc.) are expected for prototype delivery? Capabilities should work with multiple versions of Cameo / MagicDraw
11) What specific SysML model elements, views, diagrams, profiles, and levels of structural/behavioral representation constitute a sufficient automatically generated model for Phase II Government evaluation? This is phase 1 SBIR
12) The topic requires model generation with a "high degree of accuracy." What quantitative metrics and acceptance thresholds will NAVSEA use to measure extraction accuracy, relationship accuracy, model completeness, and overall model correctness? We have not published a set of assessment metrics for the tool / product quality assessment, but will likely assess based on accuracy, alignment to desired model structures, compatibility / interoperability, flexibility, ease of use / LOE required to produce functional products, etc.
13) Does "automate development of models with no human input" refer to zero-touch initial model generation, while allowing subsequent SME review of low-confidence or conflicting information, or is fully autonomous validated model production expected? Development (use of tool) and integration of finished products should require minimal human input. Human interaction to resolve errors or conflicts is acceptable but should be minimized in design.
14) Will NAVSEA provide representative legacy artifacts and/or corresponding reference SysML models during Phase I for feasibility evaluation, and, if so, which document/data types are expected to be available? For the phase 1 any similarly complex system can be used to demonstrate the capability. Specific GFI will not be provided. The government will NOT provide AEGIS, or other GFI data for the phase 1 portion of this SBIR, nor licenses for CAMEO Magic Draw. Synthetic data, or data from another system of similar complexity may be used for phase 1 demonstrations. The Government reserves the option to provide or request offerors utilize synthetic data again in phase 2; however, no decision has been made yet.
15) What versions of SysML and Cameo/MagicDraw, Navy-specific profiles/metamodels, repository technologies, and integration mechanisms must the solution support to satisfy the requirement that generated models work seamlessly within the Government's modeling environment? Capabilities should be compatible with multiple versions of Cameo / MagicDraw, including the latest.
16) What representative AEGIS datasets or document types will be made available for Phase I development and evaluation? The government will NOT provide AEGIS, or other GFI data for

the phase 1 portion of this SBIR, nor licenses for CAMEO Magic Draw. Synthetic data, or data from another system of similar complexity may be used for phase 1 demonstrations. The Government reserves the option to provide or request offerors utilize synthetic data again in phase 2; however, no decision has been made yet.

17) Does NAVSEA prefer or require the AI/ML capability to run fully within a Government-controlled or on-premises environment, or may the solution use a Government-approved externally hosted commercial AI service? For Phase II planning, should offerors assume the solution must support fully local AI inference without dependence on external AI services? Commercial AI can be used for proof of concept with representative data from similarly complex systems; however, system design data should not be provided to commercial AI due to compilation security concerns. Use of other secure AI services will be assessed based on risk profile and sensitivity of data.
18) The topic names several data classes — digital files, images, PDFs, written documents, bills of material, cable sheets, and network configuration data. Are sheets and network configuration data are where automated ingestion is both hardest and least addressed by existing tools ? No response
19) Which CATIA Magic / Cameo Systems Modeler version does the target environment run, and is Teamwork Cloud in use? Capabilities should work with multiple versions of Cameo / MagicDraw
20) We'd also value knowing whether the program intends to remain on SysML v1 through the Phase II horizon or is planning a v2 transition, given the February 2026 OUSW(R&E) SE&A transition guidance No determination has been made relative to transition to SysMLv2 for phase 2.
21) For the automated generation of SysML models with minimal human validation, does NAVSEA have a reference legacy documentation set and a corresponding hand-built Cameo model that Phase I performers can use as ground truth for measuring model accuracy, or should offerors plan to construct their own surrogate baseline from unclassified analogue documentation during Phase I? No
22) The topic identifies the continuous development of the AEGIS Combat System for nearly 30 years. Are there particular baseline(s) or configurations the Government considers most important in determining whether the generated model is operationally useful during Phase I? Similarly, are there particular artifact types that are especially representative of the initial AEGIS transition problem and that the Government expects to emphasize during Phase I concept feasibility or Phase II prototype development? There is no specific transition problem emphasized during phase 1.
23) The topic states that the solution must automate development of models “with no human input” but later indicates that finished models “may require some human modification”. Can the Government please clarify the acceptable level of human-in-the-loop involvement, such as autonomous model development followed by SME refinement, that would be considered acceptable for Phase I? Vendor approaches will be assessed as they are received. Vendors are encouraged to maximize the capability and maturity of their offering.
24) The topic states that the solution most ingest documentation and automate development with a "high degree of accuracy". Can the government please share their accuracy threshold and/or or anticipated method for measuring the accuracy of automatically generated models? We have not published a set of assessment metrics for the tool / product quality assessment, but will likely assess based on accuracy, alignment to desired model structures, compatibility / interoperability, flexibility, ease of use / LOE required to produce functional products, etc

25) Can the government provide additional information on how it will evaluate the feasibility of the concept? For example, what level of technical demonstration or representative data processing, if any, would the Government find most useful in establishing Phase I feasibility? We have not published a set of assessment metrics for the tool / product quality assessment, but will likely assess based on accuracy, alignment to desired model structures, compatibility / interoperability, flexibility, ease of use / LOE required to produce functional products, etc
26) Can commercial AI be used to generate an initial SysML model, which is then refined to accurately capture the system architecture, interfaces, requirements, and functionality? In phase 1 with non-AEGIS data commercial AI may be used, but may not be available in phase 2.
27) Is the intent of this project to develop an AI/ML capability that ingests legacy technical documentation and data as inputs and automatically generates SysML models as the output? The intent of this topic is accomplish the topic in whatever way offers deem appropriate given provided info.
28) Must Phase I modeling artifacts be delivered as native Cameo Magic Draw files, or is standards-compliant SysML XMI that imports cleanly into Cameo acceptable? So long as the output is ingestible it is ok, but should be demonstrated to be ingestable.
29) The topic states the selected contractor must be able to acquire and maintain secret level facility and personnel security clearances. At what point in the effort must the facility clearance be in place: at time of Phase I award, or prior to Phase II start? Is an FCL that is in process, sponsored under this effort, acceptable at Phase I award? The facility and information systems must be approved for SECRET NOFORN data in order to support phase 2 which will use actual AEGIS design information.
30) What representative data, artifacts, models, or mission traces will be available to performers for the legacy technical-data ingestion into model-based systems engineering prototype? The government will NOT provide AEGIS, or other GFI data for the phase 1 portion of this SBIR, nor licenses for CAMEO Magic Draw. Synthetic data, or data from another system of similar complexity may be used for phase 1 demonstrations. The Government reserves the option to provide or request offerors utilize synthetic data again in phase 2; however, no decision has been made yet.
31) What government-furnished interfaces, environments, or software versions must the prototype support in Phase I and Phase II? The government will NOT provide AEGIS, or other GFI data for the phase 1 portion of this SBIR, nor licenses for CAMEO Magic Draw. Synthetic data, or data from another system of similar complexity may be used for phase 1 demonstrations. The Government reserves the option to provide or request offerors utilize synthetic data again in phase 2; however, no decision has been made yet on phase 2.
32) What evaluation metrics will matter most at selection: accuracy, latency, operator workload, explainability, integration effort, or transition readiness? We have not published a set of assessment metrics for the tool / product quality assessment, but will likely assess based on accuracy, alignment to desired model structures, compatibility / interoperability, flexibility, ease of use / LOE required to produce functional products, etc
33) Will performers be able to use unclassified or synthetic data for the Phase I feasibility demonstration? The government will NOT provide AEGIS, or other GFI data for the phase 1 portion of this SBIR, nor licenses for CAMEO Magic Draw. Synthetic data, or data from another system of similar complexity may be used for phase 1 demonstrations. The Government reserves the option to provide or request offerors utilize synthetic data again in phase 2; however, no decision has been made yet on phase 2.

34) Are there required hardware, sensor, platform, or edge-compute constraints that should be assumed before proposal submission? For Phase 1 concept development there are no specific hardware/sensor/platform or edge compute constraints
35) Who is the intended transition owner or operational end user, and can proposers cite that office as the target customer? Not relevant to phase 1 demonstration
36) Does Phase I require demonstrated output into the Cameo Magic Draw SysML environment, or is Phase I limited to concept development and feasibility established through analysis and modeling, with Cameo integration addressed in the Phase I Option and Phase II? Yes
37) The description calls for automated model development with no human input and a high degree of accuracy, and also notes that finished models may require some human modification. Which of these governs the Phase I feasibility criteria We have not published a set of assessment metrics for the tool / product quality assessment, but will likely assess based on accuracy, alignment to desired model structures, compatibility / interoperability, flexibility, ease of use / LOE required to produce functional products, etc
38) Is there a Phase II path that remains unclassified, for example working from surrogate or unclassified system documentation, with classified AEGIS data reserved for Phase III? Notionally, phase 2 would be classified; however, we could consider an unclassified phase 2 for promising concepts.
39) Will unclassified or surrogate technical documentation be furnished for Phase I feasibility work, or should the performer plan to demonstrate feasibility on publicly available analogs? The government will NOT provide AEGIS, or other GFI data for the phase 1 portion of this SBIR, nor licenses for CAMEO Magic Draw. Synthetic data, or data from another system of similar complexity may be used for phase 1 demonstrations. The Government reserves the option to provide or request offerors utilize synthetic data again in phase 2; however, no decision has been made yet on phase 2.
40) Will Government-furnished representative technical documentation be available to Phase I performers for accuracy measurement — and if so, at what classification or handling level (unclassified, CUI), and through what mechanism and timeline? If GFI is not available in Phase I, does the Government have a preference for the surrogate artifact classes offerors use in its place? The government will NOT provide AEGIS, or other GFI data for the phase 1 portion of this SBIR, nor licenses for CAMEO Magic Draw. Synthetic data, or data from another system of similar complexity may be used for phase 1 demonstrations. The Government reserves the option to provide or request offerors utilize synthetic data again in phase 2; however, no decision has been made yet on phase 2.
41) Does the Government intend to furnish Cameo Systems Modeler / CATIA Magic licenses, or access to a Government-hosted Cameo and Teamwork Cloud environment, for Phase I performance? If not, should offerors price commercial licenses within the Phase I Base, and is a Government modeling environment expected to be available for the Phase I demonstration? No
42) What SysML version and Cameo Systems Modeler / CATIA Magic version should offerors target as the Phase II baseline? Specifically, is the target SysML v1.x on the UML profile, or is a transition to SysML v2 anticipated within the period of performance? Should work with latest versions of CAMEO but be compatible with multiple. SySML V1 is find but offerors should be prepared / have a plan to migrate to v2; however, it is not specifically required for phase 1.
43) Does the AEGIS program or PEO IWS require generated models to conform to a specified modeling method, SysML profile, or model architecture — for example MagicGrid, UAF, a

DoDAF-to-SysML mapping, or a program-specific architecture — and is a model naming or identifier convention specified that generated elements must follow? Should be able to work with multiple schema / style guides.

44) Will Government-furnished representative technical documentation be available to Phase I performers for accuracy measurement — and if so, at what classification or handling level (unclassified, CUI), and through what mechanism and timeline? If GFI is not available in Phase I, does the Government have a preference for the surrogate artifact classes offerors use in its place? Any silmilarly complex system information will suffice provided data ingestion demonstrates ability to accurately convert with sufficient quality. The government will NOT provide AEGIS, or other GFI data for the phase 1 portion of this SBIR, nor licenses for CAMEO Magic Draw. Synthetic data, or data from another system of similar complexity may be used for phase 1 demonstrations. The Government reserves the option to provide or request offerors utilize synthetic data again in phase 2; however, no decision has been made yet on phase 2.
45) The Description states documentation is 'only available on paper or in disparate legacy digital formats' and requires the solution to 'digitize, integrate, and ingest' it. For the installation drawings, installation and interface configurations, and cable run sheets that the topic's Description and Phase III text name, what mix of physical forms should offerors assume the Phase II prototype must ingest: raster scans of paper or microfilm originals (including handwritten field redlines), born-digital PDFs, or native CAD and drawing files (for example DWG, DGN, or Visio)? Will paper originals be furnished as scanned images, or does the required capability include scanning of paper originals? Offerors should maximize the number of supported formats and ability to adapt to new formats in their design.
46) **Phase I source-data priority:** Which source-data types should offerors prioritize when demonstrating AI/ML-assisted ingestion during Phase I? Does the Government prefer a representative combination of structured and unstructured engineering artifacts, or deeper processing of a smaller subset such as technical manuals, installation drawings, cable run sheets, bills of material, and IP/port configuration data? The government expects that the proposed capability work seamlessly to integrate data from multiple data sources and types. The preference is for accurate models based on available inputs. The ability to ingest a broad array of complex and simple data types may be considered beneficial if it is deemed a benefit the government.
47) **Existing versus greenfield SysML:** Should offerors assume that AI/ML-generated SysML must augment or update existing AEGIS SysML models, including preservation of existing identifiers, relationships, and configuration baselines, or is generation of a new standalone SysML model sufficient to demonstrate Phase I feasibility? The government is not providing AEGIS data; therefore, a stand alone model will suffice for phase 1. It may be deemed beneficial for the tool to update an existing model based on data ingestion and knowledge of the model schema.
48) **Cameo/MagicDraw extensibility:** Are third-party plugins, extensions, or externally hosted processing services permitted within the target Cameo/MagicDraw Government environment, or should offerors assume that the final capability must operate exclusively through Government-approved native interfaces and APIs? Should follow native interfaces and APIs.

49) **DOORS integration:** Does the envisioned operational capability require live, bidirectional integration with DOORS, or should offerors treat DOORS databases and exports as authoritative source data to be ingested and transformed into SysML? No bi directional DOORS integration is required. DOORS may one element that is used to develop models, but likely insufficient to produce the required level of detail on its own.
50) **Validation standard:** When the topic states that Phase II implementation in the Government modeling and simulation environment will "validate performance," does validation include successful loading, execution, and engineering use of the generated SysML model within the target environment, in addition to measuring completeness, correctness, traceability, and relationship accuracy against source documentation or a reference model? Not pertinent to the SBIR
51) **Government data and AI artifacts:** If Government-furnished AEGIS data is provided, what restrictions apply to its use for model training, fine-tuning, embedding generation, retrieval indexing, or other AI/ML processing? Should offerors assume that resulting model weights and derived artifacts must remain within a Government-controlled environment? The government will NOT provide AEGIS, or other GFI data for the phase 1 portion of this SBIR, nor licenses for CAMEO Magic Draw. Synthetic data, or data from another system of similar complexity may be used for phase 1 demonstrations. The Government reserves the option to provide or request offerors utilize synthetic data again in phase 2; however, no decision has been made yet on phase 2.
52) **Phase I execution environment:** Should offerors assume that Phase I Base and Option activities may be executed entirely at an unclassified contractor facility using Government-approved surrogate data, without access to a classified Navy network or Government-hosted modeling environment? The government will NOT provide AEGIS, or other GFI data for the phase 1 portion of this SBIR, nor licenses for CAMEO Magic Draw. Synthetic data, or data from another system of similar complexity may be used for phase 1 demonstrations. The Government reserves the option to provide or request offerors utilize synthetic data again in phase 2; however, no decision has been made yet on phase 2. It may be deemed beneficial if the offer's proposed approach shows an understanding of complexities of working in a classified environment and how their offering's functionality has considered that environment.
53) **NIME integration information:** Does the Government anticipate providing Navy Integrated Modeling Environment (NIME) interface documentation, APIs, schemas, sample data structures, or other integration information during the Phase I Option to support development of a realistic Phase II implementation architecture? Not as a part of the phase 1. We will consider as part of phase 2 however it may not be a direct connection; however produced SysML files will need to work in ime so compliant with recent CAMEO and SySML versions.
54) **Disconnected operation:** Should the Phase II architecture support disconnected, degraded, intermittent, or limited-bandwidth operation, including the ability to perform AI/ML-assisted ingestion and SysML generation without continuous access to externally hosted models or cloud services? Phase II will likely be in a classified environment so only services available on SIPR will be available.

55) Does the requirement that automatically generated models be usable within the Government's modeling environment include a requirement that each generated model element carry a traceable reference back to the specific source document and location it was derived from, and if so, is that traceability expected to be visible inside the model itself or delivered as a separate artifact? Up to offerer to propose.
56) What Cameo/MagicDraw release, SysML version, profiles, and model templates define the required output environment? Why it matters: determines the conversion interface and compatibility tests. Source: SRC-002 pp. 71-72, direct Cameo conversion and Government modeling environment. Unsure of reference; however, output should be compatible with multiple CAMEO versions, including the latest and be prepared to convert to SysML2.0 when needed.
57) What representative technical documents, sample datasets, or reference SysML models, if any, will be available to Phase I performers, and when? If none will be provided during Phase I, are there stated requirements for the representativeness of contractor-provided or synthetic evaluation data? Why it matters: determines feasibility evidence and data-access dependencies without assuming early AEGIS access. Source: SRC-002 p. 72, Phase I feasibility and later Government/AEGIS use. Not sure what SRC-002 p. 72 is. The government will NOT provide AEGIS, or other GFI data for the phase 1 portion of this SBIR, nor licenses for CAMEO Magic Draw. Synthetic data, or data from another system of similar complexity may be used for phase 1 demonstrations. The Government reserves the option to provide or request offerors utilize synthetic data again in phase 2; however, no decision has been made yet on phase 2.
58) What computing, connectivity, and software-dependency constraints are known for the Government-approved modeling and simulation environment? Is access to any Government-approved AI inference service, meaning a service that runs model requests, expected to be furnished, or must performers plan to supply an approved inference capability? Why it matters: clarifies the deployment requirement and the boundary of Government-furnished resources. Source: SRC-002 p. 72, Phase II environment. This does not request approval of a particular model or architecture. Phase 1 will not provide government AI or modeling services, offeror should be familiar with Navy Integrated Modeling Environment and connectivity; however, these will not be provided for phase 1.
59) Is the use of SysML a hard requirement or would you also consider UAF or DoDAF? Hard requirement.
60) What kind of feasibility and performance telemetrics are called for in Phase I ? We have not published a set of assessment metrics for the tool / product quality assessment, but will likely assess based on accuracy, alignment to desired model structures, compatibility / interoperability, flexibility, ease of use / LOE required to produce functional products, etc
61) For the Phase I Option (and Phase II planning), may performers host the ingestion capability in a commercial U.S. cloud environment handling CUI, or does the Government require deployment to a Government-controlled environment such as a Gov cloud enclave or departmental infrastructure? If Government infrastructure is contemplated, would the Government furnish access to that environment,

including Cameo, during the Phase I Option? For Phase 1, commercial cloud environments will be acceptable; however, phase 2 may require a classified environment (TBD).

62) The topic states models must support cyber assessment and authorization. Which cyber-assessment artifacts should generated models be able to produce or support — for example ports/protocols/services matrices, system boundary and data-flow views, RMF control inheritance, or eMASS-ready exports — and will Phase II evaluation include use of the model in a Navy cyber assessment process? Phase 2 models should be able to support multiple use cases; however, their end use will be determined by model quality and government needs.

63) Phase III names reliability calculation and configuration management. Should Phase I concepts address ingest of reliability data (failure rates, FMEA/FMECA, RBDs) and configuration reports into the model, or is Phase I scope limited to structural, interface, and requirements content? Offerors are encouraged to take on as much as they are able within the resources and period of performance.

64) The Description states that the AEGIS program requires the ability to "rapidly and affordably" ingest and develop system models, and Phase I requires offerors to "estimate initial performance." That's correct.

65) What representative scale should offerors assume when estimating and evaluating conversion performance-for example, the approximate quantity of source documentation, the approximate size of the resulting SysML model in terms of elements and relationships, and the expected processing turnaround If the Government has not established a benchmark, may offerors define and justify their own representative scale and throughput metrics in Phase I? ? The AEGIS system consists of dozens of server cabinets, hundreds of nodes, and hundreds of connections. Any proposed capability should be able to scale to accommodate this or larger systems of systems and not be limited by system size or complexity. It should build the system dictated by artifacts.