Q&A #1 for Navy SBIR N241-011 - 01/05/24

1. Can you clarify what is meant by "rapid generation of real-time, adaptive, high-fidelity scenarios"?

   Scenarios will include but are not limited to: entity starting positions, behavior/AI models for individual entities, terrain data, etc. Currently, scenarios are generated manually by a subject matter expert. The term rapid implies the generation of many scenarios with varying conditions or parameters that can be produces faster than the current time if would take for a subject matter expert and may include AI generated variations of the variables mentioned above (e.g., starting positions, behaviors, etc).

2. By "rapid generation of real time, adaptive, high-fidelity scenarios" is the generation process to be completed in seconds, minutes, longer periods? Will the scenarios be required to pass review and certification by trainers before use? What standards will apply for approval?

   Currently, scenarios are generated manually by a subject matter expert. The term rapid implies the generation of many scenarios with varying conditions or parameters that can be produces faster than the current time if would take for a subject matter expert.

   Scenarios will need to be verified and validated by a subject matter expert for approval. Final verification and validation for use by the Fleet will be done by the Fleet and is outside scope (TBD by government lead for transition). Goal of this effort is to give realistic, SME validated variations that can be selected and validated by the Fleet stakeholders for consumption.

3. How do you envision the adaptive behavior of red (adversary) threats in real-time? What are the key parameter or factors that should drive their adaptivity? Would this be handled by our scenario generation tool or would we select red threat behaviors via an NGTS configuration?

   Both the scenario as well as behaviors would need generated/edited to produce the desired outcome. It is expected that you would leverage the native NGTS capabilities including entity behaviors where it makes sense. Scenario adaptations could be driven by a number of factors related to complexity and training objectives, to include but not limited to threat density, entity starting positions, behavior/AI models for individual entities, terrain data, etc.

4. By "adapt in real time" do you mean that the scenario generation capability should be reconfigurable to suit changes in OpFor, TTPs, Threat Environment, Rules of Engagement?

   Yes

5. Should scenario adaptation be at the macro level (trial to trial), or the micro level (scenario is modified dynamically during execution in response to the flow of events)?

   Micro higher priority but for measurement and remediation over time both are important

6. What modalities are desired for the generative AI capabilities (e.g., text, audio, image, RF or radar data)?

   Audio and DIS/HLA/CH10; Audio data is currently being captured from the DIS/HLA network traffic and written to a ch10 file format. Tools for reading the ch10 file format are provided with NGTS.

7. Can you describe what type of analysis is envisioned for the blue communication? Can you provide more details on the desired outcomes of the automated analysis of blue (friendly) communications? What specific aspects of communication are most critical to assess?

   Aircrew are assessed based on comm brevity, comm accuracy, and timeline adherence. These data points will need to be associated with contextual data and events from the related data sources (DIS/HLA/CH10), so additional context or understanding of what is being said and what is related to (e.g., context) is desired. Audio data is currently being captured from the DIS/HLA network traffic and written to a ch10 file format. Tools for reading the ch10 file format are provided with NGTS.

8. Are there communications standards that should be applied in the assessment of trainee communications performance? Does an established doctrine for communications supporting integrated tactics exist and will it be provided?
   Existing doctrine for comms will be provided as GFI. Where not doctrinated (or more fluid), input will be provided by GOV supplied SMEs

9. Will the performance assessment be in relation to objective criterion referenced standards? Are these standards established and documented?

   Yes, to an extent. Most standards are established and documented but there are occasions in which deviation is expected and required. This needs to be assessed as well. Additionally, there may be some measures that are less defined such due to cross platform and complexity of mission set but that is more for us to determine.

10. Can you describe the unclassified sample dataset that will be provided in Phase I? What will be the size of the training data with which we will be provided? Scope? Number of exemplars? Number of parameters?

An unclassified version of the Next Generation Threat System will be provided as GFI to awardees to support this requirement along with an unclassified sample dataset will include an NGTS scenario file. Scenario files are NGTS specific formatted JSON files with a .nscen file extension. Sample DIS/HLA traffic logs will also be provided in ch10 format. Tools for reading the ch10 file format are provided with NGTS.

Additionally, Awardees will be expected to use NGTS to create relevant training data sets as necessary. Awardees will be provided access to NGTS SME's and aircrew as deemed appropriate by TPOCs.

11. Can you describe the types of parameters which will appear in the training data: weather, enemy positions, distance, azimuth, etc.?

Scenarios will include but are not limited to: entity starting positions, behavior/AI models for individual entities, terrain data, etc.

Sample DIS/HLA traffic will include time series data for each entity and event occurrence (e.g., detonations, fires, maneuvers). Entity data will include geospatial positioning (lat, lon, alt, heading, etc.). Event data may include things such as weapon fires, radio communications, Link-16, radar modes, etc.

12. Will sample recorded pilot communications be matched to sample geospatial training data? How?

Radio communications as well as geospatial data are logged to a ch10 file using NGTS's native recorder. This data can be parsed for analysis.

13. Is there any recommended game that we can use to represent the high threat density scenario? This can help us create the initial training data for the generative model.

An unclassified version of the Next Generation Threat System will be provided as GFI to awardees to support this requirement

14. What are the expected outcomes of Phase I?
Research and develop an integration plan for development of a proof of concept, standalone, capability to rapidly generate high-threat density scenarios with tactically

representative red threats that adapt in real time. This plan shall include a plan for integration into NGTS during Phase II of the project

15. Are there any recommended performance metrics to evaluate success in Phase I? For analysis of pilot accuracy in reference to ground truth, what is acceptable time lag between pilot observation and communication?
Phase I will not be specifically measured against objective standards - because this is an S&T effort, we're looking to understand what the acceptable time lag is through this research. This has not yet been defined in the domain space. It is expected, however, that Awardees report out on accuracy of communication translation and SME subjective assessments will be used to identify potential capabilities. Additionally, the feasibility of integration and transition into NGTS will be determined by NGTS SMEs.

16. Will the program be running on NAWCTSD platforms (e.g., NGTS) or on proprietary platforms?
The program will be integrated into NGTS which is used to provide the simulation environment for a variety of propriety devices (e.g., F/A-18 Tactical Operational Flight Trainer)

17. Will we be producing entirely constructive simulated pilot communications (with embedded noise) or will this be a virtual environment with pilot trainees providing test communications?
Both. Sample data will consist of both constructive pilot tactical behaviors embedded with matching live communications. Additionally, sample virtual data from trainees acting in the virtual device (e.g., F/A-18 TOFT) with associated communications.

18. Is the intent to generate scenarios and communications assessment for LVC exercises aboard ship as well as for exercises ashore? What can we assume that the exercise delivery environment will be? Traditional simulator? Lightweight (VR) simulator? Airborne exercises on the range with supporting range instrumentation (e.g., at Fallon)? Is there a specific LVC environment that the scenarios are intended to support?
While all of the above are potential delivery environments, the intent is that the technology developed for ashore training can be applied in the future to shipboard exercises. However, a "crawl, walk, run" approach for this transition will occur starting with constructive shore based, followed by virtual shore based, and so on with the end goal concentrating on LVC \ training environments.

19. Is there a target level of warfare that the air defense scenarios should be tailored to? Operational, tactical, both?

The scenarios should focus on the tactical level of employment with a future eye toward operational

20. What are the roles/positions of the intended end user to the LVC training?

Focus on fighter to fighter integration (F35 an F/A-18) with inputs from the E-2

21. In what form should the generated scenarios be produced? Nine-line? ATO? Knee-board materials?

Scenario files should be compatible with the Next Generation Threat Simulation (NGTS). Scenario files are NGTS specific formatted JSON files with a .nscen file extension. Sample DIS/HLA traffic logs will also be provided in ch10 format. Tools for reading the ch10 file format are provided with NGTS. In addition to providing NGST as GFI, info related to how to create and develop these scenarios will be provided as needed.

22. What kind of debrief artifacts are desired?

Reports and visualizations for debrief. Formats can vary but ideal these reports could be integrated with the NGTS Analysis and Reporting tools and will also be provided with Unclass NGTS software.

23. Is performance assessment to be entirely developed by intelligent automation or can there be a role for qualitative input by trainers as well?

Data should provide both objective assessment, as well as, support the qualitative assessment of performance

24. Will NAWCTSD facilitate access to SME aviators/IPs to assist with the definition of objectives and evaluation criteria

Yes

25. How will "tactically representative red threats" be defined and measured?

Red threats generated using the tools developed under this topic will be compared against instructors performing in virtual devices

26. What specific real-time adaptation capabilities are envisioned for red threats?

NGTS utilizes behavior files for reactive behavior modeling. Real-time adaptations would include generated behavior models as well as AI/ML models for controlling entities that would enhance NGTS Behaviors.

27. What specific limitations do you anticipate for AI-generated scenarios compared to human-created ones?

There are no specific limitations expected, although it is acknowledge that limitations may exist. Part of this effort will help GOVT leads determine limitations and inform assumptions about the AI generated scenarios

28. What specific format will the blue communication recordings be in (e.g., audio, text transcripts)?

Audio will be captured via the DIS/HLA network in ch10 format. It can be extracted to a standard audio format. Additionally, audio will be transcribed to text using built in speech-to-text tools native within NGTS.

29. What specific capabilities will the automated debriefing provide based on the communications analysis?

Debrief capabilities should automatically enhance current assessment and reduce manual assessment of communication (e.g., replay and manual capture)

30. What specific transition challenges do you anticipate for moving the AI capabilities from a standalone prototype to an operational LVC environment?

Transition will be via NGTS so there will be challenges associated with integration and IA/SA concerns. Based on classification level of eventual system, some avenues for web based efforts and cloud based system support will be limited for AI and code formats and standards must be in line with current Navy system requirements.

31. How will the SME evaluation be conducted and what specific criteria will be used to assess the capabilities?

SMEs will manual verify and validate the generated scenarios.

32. Is there a specific transition partner already in the works for this effort?

    Yes, transition will work through NAWCAD's Next Generation Threat System and PMA205

33. Do the two threads of this topic (generating adversary constructives and voice comms assessment) carry equal weight?
    Yes

34. Does the Government expect one integrated solution to these two requirements or can a proposer offer two stand-alone capabilities?
    While standalone capabilities are expected during phase I, phase II should focus on a single integrated solution within the Next Generation Threat System.

35. Does the commercialization plan need to speak to both capabilities or can a proposer emphasize commercialization for one over the other?
    It is preferred that the focus be on both capabilities.

36. The topic states "While communication is critical to cross platform coordination and overall tactical execution, it remains one of the most challenging training objectives to meet during Air Defense events". For what reasons, germane to this topic, does communication remain one of the most challenging training objectives to meet during Air Defense events?

    Accuracy of speech translation remains a challenge due to noise, specific communication translation (comm brevity), speed, cadence, etc. Additionally, the speed at which translation occurs can also impact utility – debriefs occur within 30min of an exercise.

37. Will the Phase I sample data be applicable to both constructive adversary generation and comms assessment?
    Yes. There will likely be a combination of both combined data (adversary and embedded comms) and separate data sets for comms and scenario generation.

38. Will the Phase I sample data include operational noise effects?
    Because of the UNCLASS environment it is not likely that noise will be generated until Phase II.

39. What was the genesis of this SBIR topic? Have Generative scenarios been built before or is this a new requirement given LVC/NGTS program plans?

This is a new concept to be explored as it is believed to benefit the speed of scenario development and accuracy of communications analysis given the burden of development on instructors/operators. This topic will help inform future requirements for such capabilities pending utility as determined by stakeholders and end users.

40. What are the data formats and standards for the training scenarios and communication data provided in Phases 1 and 2? Are ontologies or more information available on NGTS that would address using this data? Are there specific performance metrics for evaluating the AI-generated scenarios and communication analysis?

Data formats – DIS, HLA, DIS Voice, ASTI radios; An unclassified sample dataset will include an NGTS scenario file. Scenario files are NGTS specific formatted JSON files with a .nscen file extension. Sample DIS/HLA traffic logs will also be provided in ch10 format. Tools for reading the ch10 file format are provided with NGTS.

Phase I will not be objectively measured against a standard. However, awardees are expected to report out on accuracy of comm translation SME subjective assessments will be used to identify potential capabilities.

41. Is NAVAIR open to commercial LLM technology infrastructure, or looking for open source models (e.g. focus on MIT or Apache licensing) running in cloud and on premises? Colvin Run has capabilities to support both of these, thought it would help to understand transition planning in the proposal phase.

Any capabilities developed under this effort will need to integrate with the government owned NGTS software to fully transition. How this is achieved is over to the contractor to support.

42. What are the primary challenges you foresee in integrating AI technologies like ChatGPT into scenario generation and communication analysis?

Primarily there are Information Assurance (IA) and classification concerns. Technologies must be IA assess and approved. Any technology that requires web access would not be able to be transitioned, etc.

43. Are there any specific data security and privacy guidelines we need to adhere to when handling communication datasets?
All sample data during Phase I will be UNCLASSIFIED likely at the CUI level. Phase II will likely migrate to CLASSIFIED SECRET data and all processes and procedures for handling such including the ability to obtain/maintain a security clearance.

44. What specific metrics or KPIs are you looking to improve with the implementation of this effort?
Will vary by training audience/mission set. Should eb scoped for Phase I for proof of concept

45. Could you provide more detailed expectations for the deliverables of Phase I, particularly regarding the proof of concept and integration plan?

Research and develop an integration plan for development of a proof of concept, standalone, capability to rapidly generate high-threat density scenarios with tactically representative red threats that adapt in real time. This plan shall include a plan for integration into NGTS during Phase II of the project

46. What are the key milestones or criteria for transitioning from Phase I to Phase II in this project?

Phase one will develop prototype/proof-of-concept. The key for Phase II is to integrate/transition solution into NGTS suite of software tools.

47. Have there been any similar initiatives or projects undertaken previously, and what lessons were learned from those? Have you tried to accomplish this goal in the past? What lessons did you learn from those approaches?
Speech recognition and rapid scenario generation are ongoing research and development efforts being explored using varying state of the art S&T capabilities due to the burden of these tasks on instructors/operators. However, accuracy of speech translation remains a challenge due to noise, specific communication translation (comm brevity), speed, cadence, etc. Additionally, the speed at which translation occurs can also impact utility – debriefs occur within 30min of an exercise. This topic will help inform future requirements for such capabilities pending utility as determined by stakeholders and end users.

48. Is the data set mentioned in the RFP available to proposers during the proposal-writing phase, or is it only accessible to Phase I awardees? Understanding the data set in advance could significantly enhance our proposal, particularly in defining effective scenarios and configurations for the NGTS. (e.g., pre-mission briefing, ROE, weapons loadout/fuel status for blue forces) (e.g., information that would populate the NGTS configuration file)

    The dataset is only available to awardees

49. Can you provide further guidance on the scope of the air defense scenarios we should focus on? For instance, should we start with simpler engagements like a 2-vs-2 Defensive Counter Air scenario and then scale up to larger force-on-force engagements? What metrics will be used to define "high threat density"?

    Final plan and deliverables should be based on "high threat density" scenarios equivalent to 60+ red platforms

50. In the context of the multiple deliverables mentioned, could you clarify the relative importance of each? Specifically, how do you weigh the significance of scenario generation, including pre-mission briefing and NGTS configuration files, against the needs for unobtrusive communications measurement and the development of a debriefing capability?

    Weighting is based on technical merit and feasibility of transition

51. How critical is the integration with the Next Generation Threat System (NGTS) in the initial phase of this project?

    Research and develop an integration plan for development of a proof of concept, standalone, capability to rapidly generate high-threat density scenarios with tactically representative red threats that adapt in real time. This plan shall include a plan for integration into NGTS during Phase II of the project

52. Considering the potential security and operational concerns, would it be acceptable to utilize an on-premise AI tool for Phase I of the project, instead of open-source AI tools? This approach could address security concerns while still leveraging advanced AI capabilities.
    Yes.

53. That "red (adversary) threats...adapt in real time." In the context of NGTS, does the "real time" nature of this adaptation imply offline modifications to an NGTS behavior graph or online (active) modifications to red threat behavior?
    Adversary threats should be reactive in a tactically feasible manner to blue force behavior during event run-time. How this is achieved can be determined by the offeror.

54. The rapid generation of scenarios in real time. What is the time window implied by "real time?"
    TBD based the findings of this research, however, a more accurate statement would be "near real-time". It is intended that this research provide a better understanding of acceptable latency.