AINetANTX AI Prize Challenge Q&A

Updated 16 July 2021

1. **Where can I find the White Paper and Quad Chart Template?**
   a. Please see **White Paper Submission Guidelines** and **Quad Chart Submission Guidelines**. They are hyperlinked at "Appendix A" and "Appendix B."
2. **Do we need to have a SECRET clearance to attend or participate in the AI Challenge?**
   a. No, you do not need a clearance to participate! The competition is UNCLASSIFIED. However, a Common Access Card (CAC) is required for the AI challenge.
3. **Are you able to discuss the specifics of the datasets for the AI Challenge?**
   a. Please see the example data description at the end of this document. We're working on a releasable list that includes the type and format of information in each.
4. **We don't identify the specific problem our tech is intending to address as it relates to the overarching problem the challenge speaks to, correct? If that is the case, it seems the reader will have to figure that out if we don't call it out for you.**
   a. Please do identify what types of problems your solution addresses and the expected operational impact, including possible future areas of application. Propose quantitative metrics for evaluating the performance of your submission.
5. **We are an international company. Is this challenge open to non-U.S. companies?**
   a. Thank you for your interest, but per the Prize Challenge eligibility requirements, participant entities must be incorporated in and maintain its primary place of business in the United States.
6. **What Technical Readiness Levels are required for this Challenge?**
   a. It should be at least TRL 5. We are looking for effective solutions that needs to be at the point where it can be evaluated.

      NAVWAR may award, pursuant to Title 10 U.S.C. § 2371b, a follow-on prototype agreement or transaction, or Limited Procurement for Experimentation Title 10 U.S.C. § 2373 to one or more participants who successfully demonstrate an operationally relevant networking technology during the Challenge. If the selected technologies are not yet mature enough for prototype awards, other agreements such as Cooperative Research and Development Agreement (CRADA) may be utilized. This Challenge, however, does not in any way obligate NAVWAR to procure any of the items within the scope of this challenge from the winners. For full language, please see **Agreements**.
7. **Does our product have to be FedRamp Certified prior to the submission?**
   a. No, it does not have to be FedRamp Certified.
8. **Can we reuse and update open source code and technology to use as part of our submission?**

a. Yes, as long as you have the rights to do so according to the licensing of the software.

9. **Can you provide a generic example of an operationally relevant dataset?**
   a. Please see the example data description at the end of this doc.

10. **Will the data sets be raw or curated to some degree?**
    a. Both. Data sets will have the metadata in the sets. You will have access to both raw and curated data sets. Please see the example data description at the end of this doc. We're working on a releasable list that includes the type and format of information in each dataset.

11. **What are the requirements for getting access to the data repository so that we can understand what data resources are available? What is the timeline for getting permission to access the data?**
    a. A Common Access Card (CAC) will be required to access the data and file sets for the AI challenge. Access instructions and timelines will be sent to participants once white papers are reviewed. Please see the example data description at the end of this doc. We're working on a releasable list that includes the type and format of information in each dataset.

12. **Could you provide an example of the types of analysis / decision support that you are looking for with beamformed data and optask messages?**
    a. We are seeking technologies that can help to automate the visual identification of patterns from platforms while transiting at sea or in the air.

13. **As an academic institute, we have several departments and schools. Would a submission from us be considered as a single academic entity or does the department or school constitute an academic entity?**
    a. Individual academic departments may submit one entry.

14. **What are the OSA access prerequisites? Network, equipment, CAC, 3rd Party DoD token, clearance, background check, LDAP account at NIWC PAC, etc.?**
    a. Access to the Overmatch Software Armory (OSA) will be via a Skydesk account which does require a CAC. Access instructions and timelines will be sent to participants once the white paper down-select is complete.

15. **Are solutions expected to stay within the challenge furnished data sets? E.g. Does the data set include or can we provide / use red or commercial track information?**
    a. You may train on or include supplemental data. For the purposes of the prize, submissions will be evaluated using our datasets. Also see challenge statement: "Required out-of-band information needed for the AI technology to operate" or "Limitations to use due to Intellectual Property ownership" can be considered in the scoring.

16. **What level of curated ground truth will be provided for the challenge? Can you comment on the depth, breadth, and diversity of examples in the training set?**
    a. You should assume that ground truth data will not be provided within the data sets. However, several of the data sets may be used in conjunction with one another to infer ground truth. Please see the example data description at the

end of this doc. We're working on a releasable list that includes the type and format of information in each dataset.

17. **Is there a formal Performance Work Statement (PWS) that we can get in order to submit to the challenge?**

    a. All information pertaining to the challenge is located at the challenge.gov postings for the AI and Networks AINet ANTX Challenges. There is no Performance Work Statement (PWS) connected to this challenge as this is not a procurement; however, NAVWAR may award, pursuant to Title 10 U.S.C. § 2371b, a follow-on prototype contract or transaction to one or more participants who successfully demonstrate an operationally relevant networking technology during the Challenge.

18. **In the Challenge announcement, it states "Team entries or commercial entity entries must have an individual identified as the primary point of contact and prize recipient. By submitting an entry, a participant authorizes his or her name and organization to be released to the media if the participant wins the prize." Is it possible for multiple such entities to comprise a single team?**

    a. Yes, a team can consist of members from different entities.

19. **Charles mentioned that the 1st bullet on the Assessment Plan slide was most important. When I look at the white paper requirements, there is no specific section for this. Which section should we include information regarding how our tech positively affects operations and mission?**

    a. Operational Impact should be included in the Executive Summary and Critical Technical Parameters sections. Also, the evaluation criteria is included in the Judging criteria section of each individual prize challenge posting.

20. **Looking at the example REST calls, it appears projects will be able to perform a variety GET calls to query and interrogate data provided as part of the challenge. Are there POST/PUT endpoints for projects to submit data, either ephemerally or permanently, during processing (ie as part of a data pipeline or as a job) for persistence? If not, is there a persistent data storage service (ie database(s) or a data lake) for these types of operations outside of the REST API? As a follow-on, is there a data pipeline management service (ie airflow/kubeflow) that can be leveraged?**

    a. No Put/Post will be available. Assume a standard Anaconda distro is available. We have data pipelines on the backend to transform and prep data for the API Gateway… At this point we don't have Airflow or Kubeflow as a service or offering. If you have follow-up questions please reach out to us at NetAntx@niwc.navy.mil.

21. **Could you please discuss the challenges (e.g., related to standards) posed by the different types of data included in the datasets and the filesets (e.g., different formats across datasets, differences between structured and unstructured data)?**

    a. The main difference between the two is that the data may include raw binary, unstructured data, or dirty data that may appear unstructured. We're working

on a releasable list that includes the type and format of information in each dataset.

22. **Is the Navy open to papers that offer an environment for AI/ML development, training, testing? This would be an alternative to parts of the environment currently in place at the Overmatch Software Armory. If so, would the Navy consider running said environment in parallel to the existing environment during the AI modeling phase of the prize challenge?**
    a. Future challenges may focus on infrastructure aspects, but this is beyond the scope of the current AI prize challenge. Please refer back to the first slide of the brief for our focus areas.

23. **Are you expecting fully developed algorithms developed from the available data sets or is it acceptable to submit an architecture that is focused on how the data will be accessed, shared and analyzed by algorithms across the fleet?**
    a. The reason this challenge is open-ended is so you can surprise us with great ideas - so as long as your submission can integrate into the provided iRIL and OSA environment, it's okay. You're encouraged to propose strong quantitative metrics for evaluating your submission.

24. **Are participants only allowed access to datasets/filesets after the July 27th technology proposal date? Or is the data meant to inform potential technologies prior to a formal proposal?**
    a. No, our intent is to provide the data sets to invited participants after white paper review. Please see the example data description at the end of this doc. We're working on a releaseable list that includes the type and format of information in each dataset.

25. **Another question, there was a script for file retrieval earlier, can you please remind me what the programming language and UI that it would be written in?**
    a. Instructions will be provided to invited participants. We can say that technologies may use any type of shell. At this point, there is no constraints on the language that you choose to use.

26. **What openshift roles will users be given? What components of ACS and openshift will users have access to?**
    a. User should get the regular user role in OpenShift and be able to create, start/stop and monitor projects, deploy images from the image catalog or from external sources, configure and manage containers and pods in projects. I believe it is the standard OpenShift cluster configuration. We can provide the list of container images that are currently available in the OSA OpenShift instance.

27. **Will we have any access to any data before the White Paper is due? Or at least Metadata? (I have a CAC)**
    a. Yes, Please see the example data description at the end of this doc. We're working on a releasable list that includes the type and format of information in each dataset.

28. **Is the intent to field AI technologies which can operate on a meta-level, i.e. process a wide range of data types and adapt quickly to new types, or is it also of value to provide technologies which are more domain-constrained?**
    a. Either is ok. Technologies will be evaluated against the judging criteria which includes operational impact among other factors.
29. **How will we know that you have the data we need for our proposed problem if we don't know what is currently in OSA, we can't use external datasets, and we can't see it until after the white papers are due?**
    a. Please see the example data description at the end of this doc. We're working on a releasable list that includes the type and format of information in each dataset.
30. **Are there any latency requirements for decisions?**
    a. Real-time or near real-time decision speed is related to operational relevance.
31. **Do all platforms have access to the same data available?**
    a. No, we should assume that the platforms are heterogenous. The different platforms will have different sensor suites.
32. **Regarding structured vs unstructured and dirty data - are these considered significant challenges given the various data types that need to be integrated operationally?**
    a. All of the data has been put through a data engineering pipeline. You will have access to both, to include raw data. Please see the example data description at the end of this doc. We're working on a releaseable list that includes the type and format of information in each dataset.
33. **Since SageMaker natively uses ECS/EKS to host the containers for training and predict, is it requested that solutions deploy SageMaker containers OSCP 4.6 (for training and prediction) instead of ECS/EKS? Or is OSCP 4.6 requirement for non SageMake Containers?**
    a. OSCP 4.6 is preferred, but we do not expect this to be material to the evaluations.
34. **Does the platform data sets (e.g. UUV, USV, etc) include data collected from sensor payloads?**
    a. We're working on a releasable list that includes the type and format of information in each dataset. An example dataset is below.

Sample Air Tracks Data

"description": "A dataset containing aircraft flight location and status over time", "allowedFilters": ["geospatial", "temporal"], "columns": [{ "name": "icao24", "type": "varchar", "description": "Unique ICAO 24-bit address of the transponder" }, { "name": "callsign", "type": "varchar", "description": "Callsign of the vehicle" }, { "name": "origin_country", "type": "varchar", "description": "Country name inferred from the ICAO 24-bit address" }, { "name": "time_position", "type": "bigint", "description": "Unix timestamp (seconds) for the last position update" }, { "name": "last_contact", "type": "bigint", "description": "Unix timestamp (seconds) for the last update in general" }, { "name": "longitude", "type": "double", "description": "WGS-84 longitude in decimal degrees" }, { "name": "latitude", "type": "double", "description": "WGS-84 latitude in decimal degrees" }, { "name": "baro_altitude", "type": "double", "description": "Barometric altitude in meters" }, { "name": "on_ground", "type": "boolean", "description": "Indicates if the position was retrieved from a surface position report" }, { "name": "velocity", "type": "double", "description": "Velocity over ground in m/s" }, { "name": "true_track", "type": "double", "description": "True track in decimal degrees clockwise from north (north = 0 degrees)" }, { "name": "vertical_rate", "type": "double", "description": "Vertical rate in m/s" }, { "name": "sensors", "type": "integer", "description": "" }, { "name": "geo_altitude", "type": "double", "description": "Geometric altitude in meters" }, { "name": "squawk", "type": "varchar", "description": "4-digit transponder code" }, { "name": "spi", "type": "boolean", "description": "Whether flight status indicates special purpose indicator" }, { "name": "position_source", "type": "integer", "description": "Origin of this state position: 0=ADS-B, 1=ASTERIX, 2=MLAT" }, { "name": "date", "type": "date", "description": "Date of the event" }] }