A3ML FAQ

Q1: We believe we have already developed exactly what you are looking for in our commercial software product. Why are you running this program / why doesn't the U.S. Government just buy our product? (Particularly common question pertaining to TA-1.)

A1: Thank you for your question. Please submit an abstract to either TA-1 or TA-2 of the program, providing coherent, accurate, and precise technical information to demonstrate that your commercial software meets the technical objectives of the program (i.e., no marketing material or jargon).

To that end, A3ML has a variety of structural requirements for a TA-1 solution that have not been demonstrated in any known commercial software product. The following are distinct discriminators from currently market capabilities:

- A3ML is not interested in fraud; we are interested in illicit finance. While the overlap of the two are substantial, so is the divergence.

- A3ML is not only interested in banking or cryptocurrency and digital asset transactions but transactions across vastly different means that value can be transferred.

- A3ML is interested in generic pattern extraction. ("generic" in the programming languages sense.) A3ML is uninterested in the ability to classify with high accuracy and precision ***alone***. Additionally, A3ML is uninterested in a centralized "champion model"; an ideal prototype system created by A3ML should be entirely asynchronous and not rely on the existence of a central third-party.

Q2: In reviewing the abstract requirements for the Research Track, we see: Section 1: Technical Understanding – shall not exceed 1 page; Section 2: Technical Challenges – shall not exceed 1 page; Section 3: Technical Abilities – shall not exceed 2 pages. It looks like the effective limit is only 4 pages, and no section has innovations or technical plan in scope. Could you please clarify how respondents to TA1 or TA2 Research Track might use all 5 pages?

A2: A compliant abstract shall contain at least the enumerated sections within the allotted page limits.

Q3: The PS states that DARPA intends to issue an OTA in accordance with 10 USC 4022. Per 10 USC 4022(d)(1)(A), both nontraditional defense contractors and nonprofit research institutions who are involved with the program to a significant extent are exempt from cost

share requirements. Based on this, can you please confirm that nonprofit research institutions would be subject to the Model OT for Nontraditionals?

A3: Nonprofit research institutions may use the Model OT for Nontraditionals as a baseline for their proposed OT.

Q4: Section 6 of the Program Solicitation (PS) states: "Proposers may submit separate abstracts (and, if invited, oral presentations) to TA1 and TA2, but such abstracts/proposals must contain specific, detailed information regarding a firewalling of the TA1 and TA2 efforts from each other." Where in the abstract should this information be included?

A4: If a proposer wishes to submit an abstract to each of TA1 and TA2, they could create a new section in each abstract related to firewalling of efforts. Because the abstract is limited to five pages, we understand that the information conveyed in the abstract is necessarily more limited than the plan the proposer would detail in a full proposal, but the (potential) conflict of interest between the TA1 and TA2 effort must still be addressed in the abstract.

Q5: My specific SCIF accreditation means that the majority of the information required to request a copy of the classified addendum for A3ML is classified. Is there an email address on JWICS that we can send this information to?

A5: Please email A3ML@darpa.ic.gov (JWICS), A3ML@ascendc.local (SAV), or coordinate with DARPA regarding sending the information on SIC.

Q6: Can a team submit a fully unclassified response and work in a fully unclassified environment if we do not yet have access to a classified facility?

A6: A prime performer must meet the security and information technology requirements stated within Section 5 of the PS. Thus, a team cannot work in a fully unclassified environment, though it is eminently possible that a subcontractor to a prime contractor may do fully unclassified work. To that end, A3ML will accept unclassified proposals for both TA1 and TA2. A3ML does anticipate many excellent proposals will be written at the fully unclassified level.

Q7: We are a contractor or subcontractor to A3ML's test and evaluation team or a transition partner. Does this disqualify us from performance on A3ML?

A7: You should submit a justification / description of how your effort will be firewalled from your existing performance; see A4.

Q8: How does A3ML align with the current U.S. administration's priorities in the areas of (1) countering Russian threats, (2) rule of law, (3) global efforts in anti-corruption and towards financial transparency?

A8: DARPA does not make policy or take policy positions.

Q9: What specific initiatives of the current U.S. administration can you point to that A3ML will support?

A9: Examples of specific threat actors are included in the classified addendum to the PS.

Q10: Has the deadline for the Competition Track Abstract has been determined?

A10: The A3ML program solicitation is not soliciting for the competition; information regarding the competition is included in the program solicitation only for situational awareness. The A3ML competition will be advertised widely at a later date.

Q11: We do not see what DARPA BAA website we should be utilizing – could you provide the applicable link?

The program solicitation for A3ML references:

Unclassified or CUI: Unclassified or CUI abstracts can be submitted through the DARPA BAA Website, per the instructions in "Unclassified Submission Instructions (Proposers Not Requesting Grants or Cooperative Agreements)."

A11: https://www.darpa.mil/about/offices/contracts-management/proposer-general-terms

Q12: We intend to submit an abstract and part of that abstract is classified. Must we submit the entire abstract via classified medium or is it acceptable to submit the classified part of the abstract via classified medium and the unclassified part of the abstract via unclassified medium?

A12: It is acceptable to submit the classified and unclassified components of the abstract separately. Submit classified components of abstracts to A3ML@darpa.ic.gov (JWICS) or A3ML@ascendc.local (SAV), or coordinate with DARPA to submit via SIC, and reference the fact that you are also submitting an unclassified component and that the components should be considered part of the same abstract. Mixed classification abstracts are subject to the same page limits as other abstracts; the total number of pages submitted across both classified and unclassified components cannot exceed five (5) pages or the abstract will be deemed nonconforming.

Q13: Will DARPA send the classified abstract to [specific instance of a TS/SCI network that is compatible with JWICS, i.e., can be emailed to by a JWICS account]? Does this require a co-use agreement?

A13: Yes, we will send it to that address, and no, it does not require a co-use agreement.

Q14: Does the government anticipate sharing/releasing government platform data to performers to integrate with "unique private financial transaction data" and support algorithm development outside of the government platform environment? Put another way, will DARPA provide copies or cuts of government platform data to TA1 performers to integrate with TA1 "unique private financial transaction data" on the TA1 performers' environments?

A14: DARPA will not release any government platform data outside of that government platform unless explicitly approved by the government agency that controls the information hosted on the government platform. DARPA does not anticipate such approval. (This barrier to data sharing exemplifies the need for A3ML technology; an ability to create generic, sharable representations of threat finance behavior that can be asynchronously updated in a distributed manner solves the problem of being unable to provide copies of government platform data to performers for further integration.) Part of the technical challenge of A3ML is figuring out how to create such a generic, sharable representation of threat finance behavior that can be updated asynchronously by executing algorithms across multiple logically and physically separated data resources.

Q15: Will it be acceptable for an A3ML prime proposal team organization to lease SCIF space and compute resources that are operated by a 3rd party?

A15: In principle, yes, though it will be incumbent upon the prime proposal team to conclusively demonstrate that they will continue to have access to a TS/SCI development network for the duration of the program (or that lack of access does not pose unacceptable programmatic risk to their technical solution) and to work out any relevant co-use agreements.

Q16: How and where will performers access a particular government data platform? If something like a Common Access Card (CAC) is required, how does DARPA anticipate performers will access government facilities for a particular government platform?

A16: DARPA will sponsor performers for CACs and facilitate access to government facilities as necessary.

Q17: Are non-US citizens able to access a particular government data platform? Can approved foreign nationals (e.g., green card holders/resident aliens) access a particular government data platform?

A17: The ability to access individual government platforms rests with the program offices responsible for the platforms and with the owners of data resident in those platforms.

Q18: Does someone with a Department of Energy (DoE) Q clearance who is SCI eligible count as a TS/SCI eligible person per the PS?

A18: Yes, DoE Q clearances and DoD TS clearances are treated identically for the purposes of A3ML.

Q19: How will TA1 performers interact with evaluation and/or Government data? Will they submit software that T+E runs?

A19: Please read the program solicitation.

Q20: Can TA1 performers upload our data to a particular government data platform and use their services to organize it?

A20: See Q17 / A17.

Q21: Will TA1 performers get access to the surrogate data intended for use by the TA1-Competition competitors?

A21: Unlikely.

Q22: What does "partially labelled" mean in "develop algorithms for inferring new TTPs from partially labeled transaction data"?

A22: Threat finance data will be labeled heterogeneously as it will be sourced heterogeneously. One can imagine many different labeling modalities, all of which (and more!) are likely in scope. For example, some data resources may label entire subgraphs as suspicious or associated with illicit behavior, while other data resources may label specific transactions or entities involved in transactions. Some data resources may apply only binary labels while others may provide probability judgements. Some data resources may perform attribution of threat finance actors while others may not. Crucially, performers should expect a uniform, consistent interface to data resources in which all data labeling modalities may be expressed consistently but should also, in general, expect that each data resource uses a unique labeling modality.

Q23: Does the abstract format have requirements for spacing, i.e, single or double spaced?

A23: Please read the program solicitation.

Q24: Can a bibliography be included?

A24: Responsive abstracts will respect the constraints laid out in the program solicitation.

Q25: Is a deliverables table needed in the abstract?

A25: No.

Q26: Will TA2 be using real or synthetic data?

A26: A TA2 performer effort may use whatever data it believes is necessary to achieve the objectives of TA2.

Q27: Do figures, tables, and charts count for the 5-page limit for the Abstract?

A27: Yes.

Q28: Should performers include compute costs in their ROM estimates?

A28: Yes, to the extent that performers intend to leverage non-Government-provided computing resources.

Q29: The Program Solicitation notes that proposals should "Explain how the structure of the proposed algorithm(s) enables graceful degradation of performance in operational environments with diminished access to high-value data sources." Are there specific examples of high value data sources here? Is it consistent to view these AML problems as, at the very least, containing some transactional information?

A29: Examples of relevant data sources are contained in the classified addendum.

Q30: The TA-1 references "top-down" (extraction of instances of known TTPs from financial transaction data) and "bottom-up" (inference of new TTPs from partially labeled financial transaction data) problems. Is it fair to consider "top-down" as the inference/prediction phase of machine learning, where predictions such as "instance" or "not instance" are issued by a model or system? Or is there something else that is intended to be referenced by this statement?

A30: The notional top-down and bottom-up problems refer to two distinct search problems. The top-down problem starts from a library of known or hypothesized illicit financial TTPs and seeks to find all instances or partial instances of those TTPs in a distributed collection of financial transactions and other information. The representation and implementation of this search problem – as machine learning prediction problem or otherwise – is specific to the performer effort and is likely to vary widely across excellent proposals; DARPA does *not* seek to impose a uniform interpretation of how this search problem is defined and computed. The bottom-up problem searches an implicitly defined

space of TTPs for those that, when reduced to practice, could have generated the observed data. Again, DARPA does *not* seek to impose a uniform interpretation of how this search problem is defined and computed and is interested in a wide variety of unique approaches.

Q31: For budgeting travel to Tampa, how many days will each evaluation be?

A31: Unknown, but likely one full workday (~10hrs of work during one 24hr period).

Q32: How many evaluations will there be in Phase 2?

A32: Unknown. DARPA understands that this causes uncertainty regarding ROM cost estimates; this should not be a major point of contention in an abstract.