HR001119S0085
Semantic Forensics (SemaFor)
Frequently Asked Questions

As of August 29, 2019

Q35: The BAA mentions an interest in scalable methods. Do you also want to prioritize real-time demonstrations? (I.e. an alerting system for real-time manipulation in telephone or video conferencing?)
A35: Transition partners will likely be interested in near-real-time capabilities.

Q34: For TA1, do you anticipate a small number (3-4) of large teams or multiple (6-7) smaller teams? Roughly what proportion of the program funding will be allocated to TA1?

A34: It is the proposer's responsibility to determine the optimal team size for their approach. Multiple TA1 awards are anticipated; the size of the TA1 team has not been preconceived. No information will be made available on the program funding allocated to each technical area. Proposals should provide a budget appropriate for the work proposed.

Q33: Can the same company perform as subcontractors in TA1 and TA2?
A33: Yes.

Q32: For TA1, do you prefer individual proposals with specific sets of proposed detection algorithms or larger teams with more comprehensive suites?
A32: Strong TA1 proposals will address all media types and the detection, attribution, and characterization tasks. Proposals should include whatever algorithms are needed to perform the work proposed.

Q31: Should knowledge semantics be automatically learned? Using your example of the mismatched earrings: should the importance of this symmetry be learned automatically or is it OK to assume that an expert defines all semantics ahead of time?
A31: Learned or manually-curated semantics are both in scope. Proposers should address how semantic information will be maintained and updated as falsification technologies evolve.

Q30: Given the long lead time for HSR approval and that the first eval will be at the six month mark, what are the expectations for when data will be delivered to TA 1 and 2 for the six month dry run?
A30: Strong TA3 proposals will explain how they can make relevant data available to the program as soon as possible.

Q29: Is data collection and management for TA3 going to be hosted on TA2's platform or will that be managed separately by TA3?
A29: If TA3 needs compute resources to collect or generate data, that should be included in TA3's proposal.

Q28: Do TA3 performers submit their scoring code to TA2 in docker containers in the same manner as TA1 and TA2 submit their code?
A28: It would be advantageous if TA3 performer's code was run on the compute provided by TA2, in order to support rolling evaluations.

Q27: Will DARPA provide background data for the semantic models and can TA1 performers use their own data for these models or is that handled by TA3?

A27: TA1 performers will need to provide their own training and semantic model data.

Q26: Do media languages matter?

A26: The program is focused on English. Multi-language capabilities may be of interest to transition partners.

Q25: Can TA1 participants handle text, images and video but not audio? In general, multiple content types but not all.

A25: Strong TA1 solutions will address all media modalities.

Q24: How many days do you expect each hackathon to last?

A24: Hackathons are expected to be one week in duration. See the BAA, page 10.

Q23: What is a 'probe', and how does it differ from an 'evaluation' and a 'challenge problem' (p15 of the BAA)?

A23: A probe is a media asset (or collection) for the system to review. A probe may be provided to the system during normal operation, an evaluation, or challenge problem.

Q22: Is there a TA1 metric for processing time and CPU, memory, GPU maximums?

A22: There are not currently such metrics specified in the BAA but compute requirements may be of significant interest to transition partners.

Q21: Will there be guidance for TA1/TA2 about algorithms supporting GPU as opposed to requiring it?

A21: Some algorithms may require GPU access to operate in realistic timeframes. GPU is neither required nor prohibited for TA1 or TA2 algorithms.

Q20: What might an explanation or prioritization score look like (i.e. binary, qualitative, etc.)?

A20: Proposers should describe and justify an explanation and prioritization scheme that they think will be most effective for an operational customer.

Q19: What mechanics are foreseen for attribution for TA3 (e.g. how will TA1 be able to effectively attribute to media sources or authors, without sufficient sample data)?

A19: TA3 will need to coordinate with TA1 and TA2 around organizations and authors (real or synthetic) embedded in the evaluation data.

Q18: For TA3 what world data content is off limits (e.g. are there copyright concerns)?

A18: The program seeks data that can be used across the program, meets PII and HSR requirements, does not violate licensing terms, and can be released to the broader research community.

Q17: Can you provide guidance regarding the team size for TA1?

A17: Team size should match the level of effort required to execute the proposed solution.

Q16: In Phase II, when dealing with multiple articles, are TA1s tasked with making one prediction for a set of articles? Or should TA1 make individual predictions for each?

A16: Proposers should tell us how they plan to address media collections.

Q15: Do the detection, attribution, and characterization objectives change in any way when the later phases address collections?

A15: The objectives do not change in later phases.

Q14: Does DARPA have a start date in mind for cost estimation purposes?
A14: Best practice is to assume 150 to 180 days from the close date of the BAA.

Q13: Is participation at every hackathon required from every TA1 performer?
A13: Yes, TA1 performers are expected to participate in the hackathons.

Q12: Will data be made publicly available to support the community, e.g., workshops and collaboration?
A12: DARPA is interested in sharing data with the community.

Q11: What are your top envisioned transition partners for getting SemaFor tools to production?
A11: The Intelligence Community, elements of the Department of Defense, and Law Enforcement agencies are the primary transition partners but we may also look to transition to entities like social media platforms.

Q10: A large part of determining legitimacy of attribution are the externals of the media object. For instance, video standards used or delivery mechanism employed. As example: a purported CNN video was linked on Twitter. However, the link was to a non-CNN online property and the video had PAL to NTSC artifacts. Are forensics on externals of the object in scope?
A10: Forensics on externals of the object, if available, would be in scope.

Q9: Is an AMT (Amazon Mechanical Turk) task considered human subject research?
A9: AMT could be considered human subjects research depending on how the task is designed.

Q8: In detection and attribution, you are seeking to detect fake material. Fake material may be generated artificially or manually (particularly on the short social media posts). Is manually-generated fake material in-scope for the program? (e.g. trolls)
A8: Yes, manually generated falsified material is in-scope.

Q7: Is it allowable for a company to be a prime on TA1 and a sub on TA2?
A7: Yes (and vice versa), that is allowable for TA1 and TA2 performers only, as long as there is sufficient value to the Government in the work proposed.

Q6: Should TA1 algorithms be able to detect, attribute, or characterize all possible semantic inconsistencies?
A6: No, however TA1 algorithms should target detecting, attributing, or characterizing broad ranges of semantic inconsistencies or semantic information. Of particular interest is the detection of inconsistencies that require significant increases in algorithm, training data, or compute resources for an adversary to overcome.

Q5: Are temporal constraints and causal dependences considered “semantic”?
A5: Temporal and causal information tied to internal or external inconsistencies could be considered semantic and used to automatically detect, attribute, or characterize falsified media.

Q4: Is it in scope to consider, e.g., the content of a speech and compare it with the content to be expected from a specific person based on other data sources?

A4: Yes. This is addressed in the BAA, see TA1 Detection, Attribution, Characterization.

Q3: Could you give us some insights about what kind of data is to be expected under this program? Are we supposed to leverage open data sources? Is the Government going to provide background data in addition to falsified (toy) data that allows us to check for semantic inconsistencies?

A3: See the BAA section on TA3 Evaluation for a description of the data that can be expected from the TA3 performer. TA1 performers will be responsible for developing their own training data sets across media modalities. Leveraging open data sources is encouraged as long as sources follow the program's CUI guide; meet human subjects research and personally identifiable information (PII) restrictions; and do not violate licensing terms associated with the data.

Q2: In addition to multi-modal, is checking for cross-modal inconsistencies in scope?

A2: SemaFor seeks to identify falsified media that contains multiple modalities (e.g., image, audio, video, text). Checking for cross-modal inconsistencies is in scope and a key challenge of the program.

Q1: What is considered to be "semantic"?

A1: Semantic information, or higher level information, could include elements such as words, phrases, signs, visual elements, audio elements, etc.; relationships between these elements; and relationships between such elements and the real world. Primarily SemaFor seeks to move beyond low-level statistical fingerprints, such as sensor noise patterns, which are relatively easy for a skilled adversary to mask.