Q68: In addition to image/video analysis, accurate DAC of falsified multi-modal media, in particular for contextual semantic understanding, would require a focus on the dimensions of online social media diffusion, including characterizations of involved actors and information sharing dynamics. Is this in scope for SemaFor?
A68: The primary focus of the SemaFor program is on DAC using the content of the media assets provided.

Q67: Is it required for all TA1 subcontractors to attend each hackathon? For a large team with 5-6 subcontractors, the travel cost will be high. If the prime is able to mitigate any risk due to a subcontractor missing the hackathon, would DARPA allow the subcontractor to miss some of the hackathons?
A67: TA1 performers are expected to participate in the hackathons. We expect that TA1 teams are adequately represented at every hackathon.

Q66: If a proposer wants to leverage in their proposal the use of their own IP that has been built on IP licensed from another organization or university, could they effectively develop new capabilities leveraging their own and/or the licensed IP and be a performer on SemaFor? Or would these IP requirements effectively disqualify them?
A66: DARPA is interested in capabilities that do not provide additional barriers to transition, particularly to the US government. Leveraging capabilities that are already developed should be to the benefit of progress on the program goals. However, any inability to share source code with program performers because of IP restrictions could hamper progress and transition activities. Any proposer claiming the use of proprietary technology or choosing to explicitly exclude their technology from the open source regime should provide justification in their proposal.

Q65: Can one performer be a sub-contractor on both technical area 1 and technical area 4?
A65: Yes. See BAA p. 9: “TA4 performers may be awarded contracts on other TAs of the program, but conflicts of interest plans will be required in the case of TA1 or TA2 due to potential conflicts of interest with the evaluation process.”

Q64: If a TA4 performer developed any technology as part of SemaFor, would the DoD or DARPA limit that technology’s resale/reuse in the private sector?
A64: Export Control laws and/or ITAR regulations may determine whether or not restrictions are imposed on technologies developed under this program.

Q63: If a team is submitting proposals to two technical areas and they believe their approaches are more effective when tightly coordinated, could you comment on the overlap of content on technical approaches allowed in the proposals? Is it possible to share required common
elements in both proposals, e.g. teaming and background description and prior DARPA work?
A63: Each proposal must conform to the requirements in the BAA. Proposers should not assume that multiple proposals from their organization will be reviewed by the same individuals. Repeating text, such as teaming and background description, across multiple submissions from the same proposer is allowed.

Q62: If a team is planning a joint proposal and the team is considering technical areas TA3 and TA4, per the solicitation it seems these two technical areas would be 2 separate proposals. Is that correct?
A62: Yes.

Q61: When/where will the PowerPoint slide deck that was used at the SemaFor Industry Day be available?
A61: The Industry Day slide deck can be found here: https://www.darpa.mil/program/semantic-forensics.

Q60: Where can we also find other team members if we are lacking some experience in some area?

Q59: Is the Industry Day event video stream available to watch? And the slides? Where can we find it?
A59: The Industry Day slide deck can be found here: https://www.darpa.mil/program/semantic-forensics. We are in the process of making the video available on the same web page.

Q58: Under TA4 when reviewing State-of-the-Art (SOTA) generation and manipulation techniques, we discover a better technique, may the better techniques we discover be used in challenges?
A58: Yes.

Q57: In the BAA, TA4 is to develop challenges and treat models based on State-of-the-Art (SOTA) capabilities, but not to enhance the SOTA in media manipulation and generation. As a TA4 performer upon evaluating SOTA techniques for challenges, limitations are discovered and alternate methods are considered for manipulating or creating media, are these new discoveries within scope of TA4 challenges to exercise the TA1 analytics?
A57: Yes.

Q56: Will DARPA provide computational resources similar to MediFor Compute?
A56: TA2 is responsible for providing compute resources for evaluations, demos, hackathons, and system integration. TA2 is not responsible for providing compute for development and training of TA1 or TA2 algorithms.

Q55: Is it possible for a government agency to be a subcontractor for TA1?
A55: Yes, a government agency may be a subcontractor. See the BAA, Section xi. Appendix A for the required eligibility documentation. Proposers must follow the BAA instructions.

Q54: If interested in a TA in some specific sense but not necessarily teamed, can we submit proposal for a part of TA in hopes of being teamed by PM upon award?
A54: Strong proposals will address all aspects of a technical area, as outlined in the BAA.

Q53: For TA4, how will hackathons and SOTA challenge development be managed? Will there be a lead from TA4?
A53: Multiple TA4 teams will need to collaborate closely with TA3 and DARPA to execute hackathons.

Q52: In terms of generating falsified media, what is the core difference between TA3 and TA4? Both are generating and manipulating. Is TA3 precluded from anything in particular?
A52: TA3 has significantly more responsibilities than TA4. DARPA expects that TA3 will pull state of the art capabilities from TA4 to inform formal evaluations throughout the program. TA4 is explicitly responsible for staying current on the generation and manipulation state of the art.

Q51: From the BAA, page 15 top, all falsified media assets to contain two or more modalities. Shall we consider 1 modality must be text?
A51: Text is not required to be one of the modalities.

Q50: Can we have non-US citizens as team members?
A50: Yes, if there are no eligibility restrictions prohibiting them from receiving federal funding. Proposers should create the strongest possible teams to address the challenges in the BAA.

Q49: What platforms and types of social media data are interesting or required for the program?
A49: DARPA will not provide additional guidance beyond the BAA.

Q48: Many proposed solutions may build off of existing capabilities (core text, video, audio analytics) that are unlimited government rights, but which have been funded by and transitioned to various US government organizations that may not want this to be made universally available to everyone, as required by open sourcing. Is it sufficient to be able to share this sort of code among other program performers, or is it required that everything be provided to the world as open source? Is it preferable to start over and develop 100% open source-able version of key components?
A48: DARPA is interested in capabilities that do not provide additional barriers to transition, particularly to the US government. Leveraging capabilities that are already available to the US government should make future transitions easier. An inability to share source code with program performers could hamper transition activities. Any proposer claiming the use of proprietary technology or choosing to explicitly exclude their technology from the open source regime should provide justification in their proposal.

Q47: Does the use of human subjects research results from TA3 require TA1 performers to go through the institutional review board (IRB) and DARPA human subjects processes?
A47: Yes, when using human subjects research (HSR) data from an ongoing (or ended) IRB approved study, an IRB has to look at the new use of the HSR data that is being used. A more in-depth discussion and determination will only be possible when it is known exactly what data performers want to use, how they want to use it, and what risk level is associated with that data.

Q46: There is an emphasis on building and training submitted algorithms in the context of the SemaFor system, as opposed to merely running them. Is there explicit guidance on using only unencumbered build tools (e.g., disqualifying MATLAB) and training data (e.g., disqualifying paywall content)?
A46: The program evaluation of TA1 algorithms will be done as integrated elements of the TA2 system. Build and training processes would happen outside the SemaFor system, but would need to be used by transition partners and are program deliverables (see BAA p.19). The BAA requires that performers share any training data with the program (see BAA p.19). Tools such as MATLAB may incur additional license fees, adding barriers to transition.
Q45: Please discuss in more detail the scaling goals of the program. Goals, schedule, and how it relates to transition activities.
A45: DARPA will not provide additional guidance beyond the BAA.

Q44: Is studying adversarial attacks against the SemaFor defense in scope? Is making the SemaFor defense robust against such attacks in scope for TA1?
A44: Being robust to adversarial machine learning inputs to detection, attribution, or characterization (TA1) algorithms is of interest, but not a main focus of the program. TA4 is responsible for developing threat models that characterize how a media falsifier (i.e. an adversary) might use state-of-the-art technologies to challenge SemaFor defenses.

Q43: Will any "original" media be available for reasoning? E.g., provided for training, included in media collections, or provided as a high integrity bank?
A43: The Government will provide a portion of the media (images and video) developed on the MediFor program to all performers. These assets include both unmanipulated and manipulated items. A portion of the media collected by TA3, both unmanipulated and manipulated, will be provided to performers but will likely not be of sufficient scale for training.

Q42: For a TA1 system, will Attribution and Characterization be conditionally scored based on the output of Detection? Or will they be scored assuming Detection is solved? Or both?
A42: The expectation for the start of the program is that Detection, Attribution, and Characterization will be scored independently. The program’s understanding of how to evaluate Detection, Attribution, and Characterization will evolve over time.

Q41: Should TA1 teams expect that TA2 will fuse Detection, Attribution and Characterization modules independently? What happens if there is a dependency between them?
A41: The expectation for the start of the program is that TA2 will separately fuse Detection, Attribution, and Characterization scores. During the course of the program, better approaches may be uncovered.

Q40: Does each TA1 team need to address every type of manipulation in each media type? This would require virtually all of MediFor plus equivalent for text, audio, etc. If not, will the evaluation be partitioned according to those manipulation types that teams do have?
A40: Strong TA1 proposals will address all media types.

Q39: What is the appropriate way to address any disagreements with assumptions made (technical or otherwise) in the BAA? Is there any way to propose alternatives to the request as it is laid out without being disqualified?
A39: Proposers are encouraged to take advantage of the abstract process to receive feedback on specific technical approaches. Beyond abstracts, proposers are encouraged to be responsive to the program goals and to describe and justify their best possible technical approach.

Q38: Will SemaFor be focused on detecting semantic inconsistencies only or also on human generated manipulations that may appear semantically consistent but convey false information?
A38: Both are of interest.

Q37: Semantic knowledge needed to judge a detection and its characterization can be strongly dependent on the domain. For example, that violent protesters don't carry peace signs. Will DARPA provide a set of domains to focus on?
A37: A set of domains will not be provided by DARPA. “TA1 proposals should highlight datasets
and scenarios that are aligned with the proposed approach and would provide strong
demonstrations of algorithm performance.” (BAA, p.9)

Q36: Could you clarify ‘semantic models’ versus ‘statistical’ detection methods? What’s an example
of how you might measure the degree to which a model is operating ‘semantically’? Is the key
here that the model needs to offer human-interpretable explanations such as “there are
children at this supposedly-violent protest”? (Arguably a black-box deep neural network could
be implicitly capturing such semantic information even if it appears to be operating purely
statistically.)
A36: Many current statistical detection methods are completely reliant on noise patterns that are
unrelated to the meaningful content or context in a piece of media. For instance, sensor noise and
compression artifacts are often used in current image forensics. Such low-level detection
techniques can be fragile when presented with falsified media from a skilled adversary.
Understanding content and context is important for creating much richer detection, attribution,
and characterization algorithms.

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.