

HR001121S0008
Influence Campaign Awareness and Sensemaking (INCAS)
Frequently Asked Questions

November 4, 2020

Q96: Regarding the abstract due in 13 days, it is unclear where to put referenced work. My team has a selection of relevant academic publications and software that would be useful to include for reference, but the BAA does not outline such a section in either the proposal or abstract. Is the capabilities section of the abstract the preferred place for us to include these references, or is there an allowance for references outside the main 5 pages?

A96: References (e.g. to publications, etc.) should be included within the 5 pages. Referenced material (e.g. actual text papers) should not be included in the abstract.

Q95: If the solutions requested are to be general to the sorts of populations, then what elements of the program will validate that the tools developed perform with cross-cultural relevance?

A95: The evaluation scenarios will be chosen by TA5/DARPA in part to validate cross-cultural relevance.

Q94: If the solutions requested are to be general to the sorts of populations, then can you provide any guidance on how large of populations one would study? Country level? Smaller or bigger regions?

A94: Solutions should be general and expected to apply across varying sized populations. However, the program is tracking population at the population segment level and the sizes of those segments could vary based on the distribution of responses and size of the overall target population.

Q93: If I'm interested in understanding extracting agenda, concerns and sentiment (TA1) with help of images AND text (how they work together), but not interested in computer vision techniques for computer vision's sake, is this in scope, beyond TA4? Similarly, if interested in extracting psychographic markets from images people share online (TA2), is this in scope?

A93: Yes.

Q92: During anticipated future deployment, would TA1/TA2/TA3 tools have any ability to actively elicit and collect reactions from the hypothesized target population or hypothesized adversary? For example, distributing an advertisement over social media that is designed to catalyze some diagnostic reaction.

A92: Active engagement with the target population is out of scope. Performers may survey relevant sampled populations subject to HSR restrictions and DARPA approval.

Q91: Can you provide any details on the types of operational analysts that TA3 and TA4 performers will coordinate and transition tools to?

A91: Not at this time. INCAS tools are expected to have broad applicability across DoD, IC, and other USG organizations.

Q90: Can TA1 look at multiple messages?

A90: Yes

Q89: A key element in detecting a foreign influence campaign could be looking at changes/evolutions in influence indicators over time (which is different from simplifying being able to identify/characterize the indicators). In DARPA's view, would tracking/assessing changes in influence indicators fall under TA-1, TA-3 or somewhere else?

A89: Detecting and handling changes in indicators over time could fall under both TA1 and/or TA3.

Q88: Re: Teaming. Are you going to share others interested in this program to facilitate teaming?

A88: A list of attendees will be provided with the names of those who consented at registration to share their contact information.

Q87: How will the TA5 group be selected?

A87: The TA5 performer will be selected outside the BAA process.

Q86: How will the SMEs for TA5 be selected?

A86: SMEs will be selected by the TA5 team in consultation with DARPA.

Q85: The information flow diagrams makes it seem like a completed TA4 is a dependency for all the other TAs. How are we going to work with this?

A85: TA1-TA3 may need temporary mechanisms for ingest of TA4 collected data until the TA4 APIs and underlying capabilities are in place. This will be refined at program kickoff.

Q84: Will the SMEs from TA5 be available to the other TAs for consultation, or should TA3 proposals include their own SMEs? Staffing will be challenging since DARPA has not identified the scenarios.

A84: TA5 will provide some access to subject matter expertise through the Program's SMEs group, but performers may wish to supplement as appropriate.

Q83: The BAA states that for TA2, "Technical approaches need to be data driven". Could you provide an example of what will be considered NOT data driven?

A83: Data-driven refers to bottom-up segmentation of the responding population to influence messaging by response type, rather than just segment populations into a fixed set of segments (e.g., age 20-29) and attempting to gauge reaction by fixed categories.

Q82: Are TA2 performers supposed to infer demographic attributes from social media messages or are they supposed to obtain such information (whenever available) from public user profiles?

A82: Demographic and psychographic attributes should not be assumed to be available from the platforms for target (sub)populations of interest.

Q81: Do you anticipate that TA2 will need to utilize Natural Language Processing (NLP) techniques, or can TA2 assume that TA1 will have brought all data into a machine-understandable data model?

A81: TA2 may require NLP techniques as it may be dealing with additional data suggested by TA1-identified influence messaging.

Q80: In TA1, what are the limits of meta-data? Is it restricted to content/temporal/geospatial or can it include platform behavior (e.g., likes, shares, deletion, friending/following)?

A80: Metadata can include platform-specific behaviors to the extent they can be accessed through the data available by official platform APIs.

Q79: In TA1, are you interested in behavioral patterns that may be observed across multiple social media platforms (e.g., patterns that manifest in terms of a manipulation that exploits interactions between Twitter and Facebook)? Are you interested in transits to real-world activity (e.g., demonstrations)?

A79: Yes to both.

Q78: In TA1, are behavioral patterns of interest as potential indicators of geopolitical influence? For example, patterns that may be indicative of specific types of campaigns, or specific types of objectives; behavioral patterns that may be indicative of what may come next in a campaign.

A78: Yes.

Q77: A question of clarification regarding what is in-scope/out-of-scope for TA1: Did I hear the PM say that TA1 focuses on identifying and extracting the indicators of influence only, and that organizing the indicators into a cognitive framework of sense-making is not desired for TA1, since that is the human analyst's role?

A77: Yes. Organizing indicators for sense-making will be done by analysts, enabled by TA3 capabilities.

Q76: [How would the TA1 information be] interacting with the other TAs? (Small example – the outrage emotion – would that be given to analyst as an English semantic indicator of “outrage” or should this concept be presented/stored/transmitted in a multilingual way across the system)

A76: The representation of surfaced indicators and characterized population response is part of the campaign model design by TA3 working with the other TAs.

Q75: What is the rationale for TA3 not making contact with TA1/TA2? It seems odd to build a model of the campaign independent from its data sources; how will TA3 performers ensure compatibility with data sources available/processed through TA1 and TA2?

A75: TA3 will interact with TA1 and TA2 via the TA4 APIs. TA3 is expected to exploit the output of those TAs to help analysts build a campaign model. The campaign model and additional analyst feedback (from TA3) will be exploitable by TA1/TA2 via the TA4 API.

Q74: The BAA talks little about TA1 and TA2 coordination. Will they share a common representation of target groups?

A74: TA1 and TA2 will primarily interact through data sharing via TA4 APIs. Both TAs are expected to contribute to the evolving campaign model.

Q73: How will novel TA2 features/metrics on population response be integrated with TA3 models? How will potential conflicts be resolved?

A73: All TA2 efforts will contribute to the campaign model. The analyst, supported by TA3 interfaces and analytics, will adjudicate conflicts as necessary.

Q72: What kinds of population responses is the Program considering? Media only? Protests? Voting? Acts of civil disobedience?

A72: Both online and offline responses are in scope.

Q71: What is the thought process behind having TAs 1 and 2 exposed to the multi-lingual and -cultural challenges, but not the multi-media aspects of inputs, which are handled in TA4?

A71: Developing the program scope.

Q70: Program staffing and proposal may depend on the specific languages needed. Can DARPA at least provide a set of, say, four non-English languages from which the two languages will be selected?

A70: INCAS will investigate English and two additional languages.

Q69: It's clear that the program should not formulate or execute response campaigns, but is the program intended to be entirely passive? Specifically, is TA2 entirely passive, or can it involve active surveying to segment the population?

A69: Active engagement with the target population is out of scope. Performers may survey relevant sampled populations subject to HSR restrictions and DARPA approval.

Q68: Do you have any preferences or expectations regarding the areas of the world or population types that are of interest? Is there an explicit goal to develop methods that can adapt to a need to characterize influence campaigns anywhere in the world? Alternatively are there key areas of the world that should be kept in mind?

A68: The tools should be general enough to address a variety of topics and target populations for influence operations. TA1 and TA2 are specifically asked to describe the generalizability of their algorithms. TA3's campaign modeling capabilities should also be general. It is expected that some customization may be required for techniques to different topics or target populations.

Q67: Are simulations and "what if" questions considered out of scope for this effort?

A67: Yes

Q66: Can you elaborate on the definition of non-US population? Is INCAS tool only intended to be used oversea?

A66: INCAS is only developing tools for test and evaluation by operational users. For the INCAS program, those tools will focus on sensemaking of influence operations impacting non-US populations.

Q65: As a specific application domain, are you considering influence campaigns on cyber security facts which might have impact on the business and operations of industry actors?

A65: Addressing campaigns with elements aimed at business or industry actors is potentially of interest depending on the evaluation scenarios chosen.

Q64: The BAA mentioned multilingual. Is the proposed tool required to support multi natural languages?

A64: Yes, the BAA describes how tools must be demonstrated in at least 2 non-English languages besides English.

Q63: Is there an expected/allowed, real-time interaction between the human operator and the TA1 or 2 elements-- e.g., to tune assumptions, prioritize focus of attention, etc.?

A63: Real-time with the human user and the TA1 or TA2 elements is potentially in scope. Such interaction may also potentially be asynchronous.

Q62: We understand that proposers can submit proposals for more than one TA. Can proposers also submit different teaming arrangements for those proposals? For example, could a proposer

submit as a team prime with subcontractors for one TA and also submit as a subcontractor to a different team for another TA? Is a proposer prohibited from participating as a subcontractor for multiple proposals under the same TA? For example, could the same company be a subcontractor for "Team A" and "Team B", submitting proposals for the same TA?

A62: A proposer can be on multiple submissions, including ones for the same TA. Be sure to address any potential OCI issues if appropriate/applicable.

Q61: Given that TA3 tools will require analyst feedback, do you recommend that TA3 performers include intelligence analysts and operators on their team or will the government provide analysts for TA3 performers to interface with?

A61: TA5 will provide some access to subject matter expertise through the Programs SMEs group, but performers may wish to supplement as they find it necessary.

Q60: Are multiple org responses encouraged? If so, will there be a list of attendees provided so we can reach out to others interested in responding to INCAS?

A60: Teaming is encouraged, but is not required. A list of attendees will be provided with the names of those who consented at registration to share their contact information.

Q59: With regards to costing travel (for PI meetings), do you have any additional guidance as to how that costing should be performed with the current pandemic? Specifically, quotes for flights today are likely to significantly differ than flights in future years. Should we base quotes on the current cost of flights or project a future cost and, if so, how do we represent quotes for that in a cost volume.

A59: Proposers should propose as required by the BAA.

Q58: Would it be possible for us to recommend the additional languages in our proposal?

A58: DARPA will select the non-English languages based on the scenarios chosen. Proposals may suggest languages to focus on.

Q57: With regards to the place of performance, do you have guidance for how best to declare that given remote work during the pandemic? With more people working remotely, should we be listing all possible locations that people may be working remotely from or just from the physical site (e.g., company headquarters)?

A57: Place of performance should reflect the proposers SAM registration. The proposer may provide additional information as appropriate.

Q56: The BAA seems to be focused on very forward-thinking proposals. Research on already-known techniques like NLP and bot detection are considered out of scope. I'm wondering, to what extent will the selection process be tolerant of proposals whose forward-thinking nature means that they lie outside of some of the technical specifications? To be more specific, the BAA as it stands now, by having all TAs dependent on the TA4 data aggregator, does not allow for the possibility of emerging new channels of disinformation. For example, disinformation researchers used to primarily focus on open-forum social networks like Twitter and 4chan, but just this year there has been a great deal of talk about violence in India coordinated via WhatsApp chat groups.

A56: TA4 would work with other TAs to identify data sources potentially relevant to the evaluation scenarios. Proposers may suggest alternative sources.

Q55: Should we propose languages for Phase 1 and Phase 2 in TA1/TA2 or address how we will handle any language DARPA chooses?

A55: DARPA will select the non-English languages based on the scenarios chosen. Proposals may suggest languages to focus on.

Q54: Is it required for a proposal to TA1 or TA2 to cover all aspects of the TA? Is it acceptable to propose a solution to only part of the TA1 or TA2 problem set?

A54: A strong proposal should address the entirety of a TA.

Q53: Is it an intentional part of the structure of this BAA that activities should be limited to the data sources already specified via TA4? Would a proposal involving other data channels, or even the potential for detecting as-yet-unknown new channels, be considered responsive?

A53: Other data channels may be suggested.

Q52: If the same team or PI submits multiple proposals to TA1 and TA2, would the expected synergies have to be described in both proposals or is it sufficient to describe the expected synergies in just one?

A52: Each proposal must stand on its own; therefore, expected synergies must be described in each proposal submission.

Q51: For TA2, we are asked to identify at least two additional languages that our approach could accommodate. Given that DARPA intends to specify the language(s) at the beginning of each phase, and these may not be the ones we identify in our proposal, should we include in our proposal tasks and cost necessary to adapt our approach to the languages we identify? If so, how does DARPA plan to handle any difference in the cost of work required to adapt our approach to languages other than the ones we have used as a basis for our costing?

A51: You can assume that the languages chosen have the same general level of resources.

Q50: Can we pick our own non-English languages or will they be assigned?

A50: DARPA will select the non-English languages based on the scenarios chosen.

Q49: Are you expecting that we propose the four influence indicators now or just propose some examples, which may or may not be used during the effort? Would these be fixed across the program?

A49: Influence indicator types include Agenda, Concerns, Emotion plus additional types that proposers must propose. See the BAA for details.

Q48: Where do you foresee human subjects research impacting this program? Which TAs should be doing human subjects research?

A48: Human subject research is not a requirement, it may be part of the technical approach.

Q47: Is Human subject research expected of TA1/TA2 performers?

A47: Human subject research is not expected for TA1 or TA2. It may be part of the technical approach.

Q46: Is experimentation with human subjects for self-assessment purposes in scope / of interest for TA1 and TA2?

A46: If this is part of the technical approach for TA1 or TA2, then it is in scope.

Q45: Does TA2 require IRB and HRPO approval?

A45: Consult your Institutional Review Board (IRB) office as you develop your institution's research plan.

Q44: Are we correct in assuming that if our proposed scope of work solely involves aggregated publicly available information obtained through online sources, we would still need to go through the HSR approval process?

A44: Yes

Q43: Would TA3 have IRB considerations WRT the analysts who would be using the tool?

A43: Consult your Institutional Review Board (IRB) office as you develop your institution's research plan.

Q42: Many TA 1 teams will be analyzing text from various sources from TA 4 (social media, news articles, blog posts,...). How does DARPA recommend that TA 1 performers conduct their IRB reviews given that obtaining consent from individual human users is not possible?

A42: Protocol is drafted stating that data is being collected from existing social media, news articles, etc. Then the protocol and supplement documents are reviewed and approved by the local IRB (of the performer) and then submitted for an administrative DoD review. Once both reviews are complete, then the performer can use the data.

Q41: Will the slides, especially Dr. Kettler's, be made available?

A41: Yes, The slides are located at: <https://www.darpa.mil/program/influence-campaign-awareness-and-sensemaking>.

Q40: For the purpose of scope, does this program have an overall funding amount?

A40: Proposals should be cost-realistic and considerate of the scope of this INCAS effort.

Q39: Can you share the anticipated funding amount for TAs per award?

A39: Funding will vary depending on the scope and quantity of proposals received.

Q38: What is ground truth for influence indicators? SOMETIMES automated assessors can be subtler than human assessment...

A38: Ground truth will be determined by the TA5 team.

Q37: Source information is critical to sensemaking. If attribution is not sought, can we assume that attribution information is available a priori?

A37: The program is not attempting to do fully automated campaign attributed. Some metadata may be available that would aid analysts in attribution, but this should not be assumed.

Q36: Related to the evaluation scenarios, are TA1/TA2/TA3 performers expected to only use social media, news, and other data collected passively from the open/public domain and provided by TA4? Or are they expected to also have additional monitoring or collection themselves?

A36: TA4 will be responsible for collecting data. TA1 and TA2 may wish to supplement that with additional training data depending on their algorithms' needs. They will not be responsible for monitoring online data.

Q35: Is the program interested in messaging platforms (e.g., groups in (WhatsApp, WeChat, Telegram, etc.) as a data source?

A35: Yes. Those platforms may be relevant to the evaluation scenarios chosen.

Q34: Is it in scope for TA1/TA2/TA3 performers to actively elicit and collect information from people (e.g. surveys, polls, social media quizzes, user studies) during the program, as long as HSR procedures are followed?

A34: Active engagement with the target population is out of scope. Performers may survey relevant sampled populations subject to HSR restrictions and DARPA approval.

Q33: Is it acceptable for TA1 or TA2 to use internal repositories for data specific to their processing and analytics, or are they expected to use the TA3 knowledge graph for all processing?

A33: Yes for development/test purposes, but TA1 and TA2 should provide output to and access data through the TA4 API. For tool deployment, using internal repositories may not be feasible.

Q32: In addition to using real data, are there any prospects of collaborating with the performers on the DARPA SocialSim program <<https://www.darpa.mil/program/computational-simulation-of-online-social-behavior>> (same PM) to explore their synthetic data?

A32: Depending on the evaluation scenarios chosen, DARPA may provide SocialSim data as Government Furnished Information. But proposers should not assume this.

Q31: Do you expect TA4 to provide all the ML training data to meet TA1 and TA2's needs, or only the operational scenario data? The relative ratio of the sizes of these corpora could be substantial.

A31: TA4 will collect and curate data for the evaluation scenarios. This will include some training data, but TA1 performers may need to supplement this depending on the needs of their algorithms.

Q30: Can you say something about how big the firehose of data will be? What does being able to work "at scale" mean for this program?

A30: This will be dependent on the scenarios chosen, but all scenarios will likely involved high volume collection for a target population of interest.

Q29: Can TA1 teams complement the data provided by TA4 with their own data collection, in case their indicators require ad hoc data?

A29: Yes.

Q28: The BAA requires YouTube.... However, program is not focused on video... What is the program expecting to do with YouTube data?

A28: The BAA does not require the use of YouTube. YouTube comments may be exploited.

Q27: Automatically detect implicit and explicit indicators of geopolitical influence in multilingual online messaging. Should we assume the analyst will get indicators that are "language-independent" or will they be able to read these other languages?

A27: TA1 surfaces indicators from the native language, but labels for these indicators could be in English. The design of the HMI in TA3 may show both the native language and an English translation of the actual messages.

Q26: Is data curation a part of the effort?

A26: TA4 will collect/procure and curate data for the evaluation scenarios.

Q25: Will TA4 provide labeled data for training?

A25: TA4 will collect and curate data for the evaluation scenarios. This will include some training data, but TA1 performers may need to supplement this depending on the needs of their algorithms.

Q24: Is TA4 responsible for data collection, or just ingestion and management of data that DARPA will obtain?

A24: TA4 will collect/procure and curate data for the evaluation scenarios.

Q23: For the TA3 HMI users, will the HMI be required to support native language displays for the other foreign language? Or can all users safely be assumed to be English speaking.

A23: All users can be assumed to be English speaking.

Q22: For foreign languages, will machine translation services be provided by TA4 – to include services for idiomatic expressions in social media? Translation services for idiomatic expressions in social media?

A22: This will be up to the TA4 performer and depend on which languages are selected by DARPA for the evaluation scenarios.

Q21: Can you say something about the nature of the additional languages and the evaluation datasets you are envisioning? Particularly of interest is whether low resource languages and/or novel subcultures (with novel jargon, memes, etc.) will be involved?

A21: Languages besides English will be determined by the scenario chosen by DARPA, working with TA5.

Q20: The BAA states that TA2 needs to have the capability to ingest data from TA1. In general, how much does the proposal need to address compatibility and interoperability with other TA1 and associated systems/data types?

A20: TA1 and TA2 will work with TA4 to develop an API that handles this once the program is underway.

Q19: Will this program be procured via DARPA Contracts or via an Agent (e.g., AFRL, etc.)?

A19: The program will be procured via DARPA Contracts Management Office.

Q18: How stringent is the requirement that TA5 be only FFRDC/UARC?

A18: Per the BAA, TA5 will be awarded to a UARC or FFRDC.

Q17: Given the program is intends to collect and store social media content from the Internet (not synthetic content), has DARPA adopted a position on copyright implications and limitations regarding the use of the data? Is this considered fair use – or is the original content expected to not be stored on project computers (e.g., only a link to the original source?)

A17: No, DARPA has not adopted a position. The data made available to the program via TA4 will be lawfully distributed to TA1-3 performers.

Q16: Can you say what an Associate Contractor Agreement is?

A16: "Associate Contractor Agreements" (ACA) are agreements between contractors working on government contracts or projects that specify requirements for them to share information, data, technical knowledge, expertise, or resources.

Q15: Are individuals working on TA5 also allowed to work on a proposal or is TA5 a conflict of interest?

A15: Personnel working as a subcontractor or as a consultant to TA5 will not be permitted to work on any TA1-4 tasks.

Q14: Can we perform the tasks (related to RFP) outside USA?

A14: See A5.

Q13: You said you expected that most of TA1 and TA2 work would be done in Phases 1 and 2. Will they be expected/allowed to participate in the capstone exercise? Will that be contingent on their having SECRET cleared individuals? Will they be allowed/expected to adapt their algorithms and approaches, or even during, the capstone exercise?

A13: TA1 and TA2 will develop the bulk of indicator and population attribute extraction during Phase 1 and 2. During Phase 3 they may be increasing accuracy, maturing the capabilities, or customizing to the capstone challenge problem scenario. TA1 and TA2 will be able to participate in the capstone exercises to the extent enabled by the security level required. Ongoing feedback will be provided to all performers from the capstone experimentation so that they can adapt tools accordingly.

Q12: What is the expected collaboration between TA1 and TA2? The BAA seems to talk about collaborations between other TA's but not this particular pair. It may be important for TA1 to have the knowledge of targeted populations to determine if an attempted influence exists, especially for indicators of CONCERN.

A12: TA2 is expected to use the output of TA1, which TA2 will obtain through the TA4 API. Both TA1 and TA2 will contribute to and exploit the campaign model via TA4's API.

Q11: The BAA (Figure 3) does not indicate that TA2 receives TA1 indicators as input. Can you confirm that it is expected that TA2 will use these as an input?

A11: TA2 is expected to use the output of TA1, which TA2 will obtain through the TA4 API.

Q10: Is TA3 or TA4 responsible for developing and maintaining the campaign model? (Page 15 of the BAA says, "TA4 is responsible for implementing the campaign model and thus will work closely with TA3 on its design.")

A10: TA3 will be responsible for the design and development of the campaign model. TA4 will be responsible for data persistence in this model. TA3 and TA4 should plan to cooperate closely on this.

Q9: Can TA1 look at multiple messages or should focus on only individual messages? That is, the notion of "fragments" in the BAA sounded like it referred to multiple messages but in the briefing it sounded brief tweets rather than fragmented in time.

A9: TA1 can look at multiple messages.

Q8: Can DARPA provide references for assumptions we should make regarding required content for Preliminary Design Review and Critical Design Review?

A8: These documents should be treated as informal and capture the revised plans for each effort. Details will be provided at program kickoff.

Q7: Will abstracts be evaluated in the order they are received, thus incentivizing an early submittal?

A7: The BAA does not prescribe the abstract review process. DARPA highly encourages abstract submissions and will review all abstract submission received prior to the abstract due date.

Q6: Can proposals be submitted via email?

A6: No, per the BAA, page 43, "Since proposers may encounter heavy traffic on the web server, it is highly recommended that proposers not wait until the day proposals are due to request an account and/or upload the submission. Full proposals should not be submitted via Email. Any full proposals submitted by Email will not be accepted or evaluated. Proposals must be submitted via the DARPA BAA portal at <https://baa.darpa.mil/>.

Q5: Can the program tasks be performed outside of the USA?

A5: Yes.

Q4: Will travel be required for program events for performers outside of the USA?

A4: Performers are expected to attend semi-annual PI meetings in person, subject to any travel restrictions due to extenuating circumstances.

Q3: Can non-US companies be awarded contracts for this program?

A3: Yes.

Q2: I noticed in the BAA that there is a heavy focus on natural language processing. However, there is also mention of multimedia. Modern social media are all multimodal, so will supplementing of NLP with analysis of images, videos etc. be part of the scope of the program? Only in the service of TA-1 and TA-2 of course not multimodal content analysis for its own sake.

A2: TA4 will be experimenting with visual imagery (memes, ads, etc.) and their conversion to text for TA1-TA3 processing.

Q1: HR001121S0008 Abstract submission portal not available on <https://baa.darpa.mil/>. When will this opportunity be available on <https://baa.darpa.mil/> <<https://baa.darpa.mil/>>?

A1: The site is available.