



SBIR 22.4 Q&A Telecons Transcript  
18 July 2022

**SBIR Programmatic Introduction & Q&A**

**1. Can you please confirm USSOCOM will not accept Phase I feasibility that has any prior SBIR/STTR? Or is it acceptable if such work is not the sole justification?**

If you have done this before in another SBIR, by law, that is by SBIR/STTR policy directive we cannot do it because this has been done before. So, if you tell me “yeah we have done this before” here’s the SBIR/STTR then we can say this one cannot be done for that reason. I find it hard to say that this is the exact same thing that is done by another entity because of the details provided in the statement of objectives. So, we want to see the technical piece. Our technical evaluators will want to read the technicalities of how you have done this. If you say we have done 1,2 and 3, okay great, but what did you do and how is it technically feasible and how are you going to reach what we’re looking to prototype.

**2. Are there specific cybersecurity requirements to which SBIR prototypes need to adhere?**

Eventually. We focus our SBIR/STTRs on transition and this one is just like any other one. We are looking to see how the prototype will be successful. With that said, there is some cyber security mechanism that is within a statement of objectives that may be required. From your company’s perspective, I would say, make sure you are adhering to the DFARs (Defense Federal Acquisition Regulation) for cyber security to make sure you are securing your network and your systems to protect yourself and protect your technology from others jumping in to take your stuff. So, that is also required. I don’t have the DFARs handy right now, but I can try to look it up in the meantime.

**3. Can you elaborate fully on the point "demonstrate that technical merit and feasibility has been established"...**

The feasibility is done. We are going direct to Phase II because we see through market research and the work done that this is feasible. Now the technical perspective of saying, yes I can create this is based on here’s the research, here’s the work done, here’s the technology we are working towards and providing that Phase I like feasibility study, the work that you’ve done already, is really the focus around the technical merit of feasibility.

**4. Can we apply for / be granted SBIR for more than 1 of the proposed 4 topics?**

Yes, you can apply to all four if you would like. All four are really under the same program office. There are relationships between them, and you can apply to all four if you see that you can get to it. Keep in mind that when you apply to all four of them, one of the things we see sometimes is same solutions, same proposal, same funding documents, same everything and we look at it and say it’s like a canned approach to submitting to all four. But all four have different approaches and different description requirements, so that’s something to keep in mind.

**5. Can you confirm this is Phase II?**

Yes, this is a Direct to Phase II.

**6. For the companies that have not gone through a Phase 1 before, is there a template for the feasibility study that we can model after?**

That’s a good question. Again, I can show you a Phase I report CDRL that we have done in the past, it’s more of a deliverable, and here is what we are looking for on technicality. But at the same time, it’s really just the study itself. Now you really shouldn’t be doing a Phase I saying it’s a Phase I, you should be doing a feasibility study. With that being said, let me see how we can put out a deliverable example of what the Phase I feasibility study final report looks like. Again, we do leave it up to the creativity of the small businesses to say what the technology looks like





from their research and study. There really is no page limit on the feasibility study, so make sure you use that wisely. Something I failed to say is we are aiming and working extremely hard to get all the awards for D1-D4 within this fiscal year. Before the end of September this year we are looking to reward those topics.

**7. For the Tech Vol 2, P1 proposals have 25 pages but this is limited to 10 max with the same template. What is the difference that drops half the pages?**

Our Phase I proposals have 5 pages max. It really depends on agency and DOD component, we all do it differently. If you look at our prior Phase I proposals we are asking for only 5 pages, but this one we are giving it 10 pages.

**8. What are the restrictions on non-US nationals working on these SOCCOM DP2 projects?**

With SBIR, that is only US nationals, US citizens and US residents as well, so green card holders working on the project. I believe there is an international traffic in arms restriction in there as well and of course as a SBIR, SBIR has to be done all in the United States. Even if we are talking about subcontractors, etc., all SBIRS must be done inside the US, all 50 states.

**9. Will a video of operation of a prototype capability be permitted as part of the feasibility study?**

I don't think so. I think everything has to be within the written document you provide to keep the fairness of our work.

**10. We would have a non-SBIR feasibility to use but also a previous SBIR that should be relevant to SOO. Want to make sure that would not be deemed "non-responsive"**

To answer your question, this is why we are doing a Direct to Phase II because we have done our market research, and we said "hey there are small businesses out there and the technology is out there, it can be done, it's feasible." I doubt you have the same exact solution for our use case or for our requirements. I echo what General Clark, our SOCOM commander, said at SOFIC that you have no idea how we can use your technology, so just let us know what it is. So with that said, if you have done something before and SBIR you have done before, its relevant but it's not the exact same thing. Keep that in mind when you are writing the feasibility study, the proposal, is it really the same exact solution or is it something that will need additional innovative work to get to the final solution?

**11. Does the presented solution need to already have an ATO for your environment?**

No. We are doing innovative technology here. We are doing anything from a TRL 3 to TRL 6-7. With that being said, that ATO would be something attained by the program office after the Phase II. It is not something that will be funded by SBIR dollars but the ATO itself will be something that the program office in a Phase III that is asking for the topic authors. So no, you do not have to have an ATO and keep working with cyber security in mind because eventually the final technology will probably be needing an Authority to Operate.

**12. Does the prototype have to be deployed into a SOCOM environment or will a contractor environment suffice?**

I don't want to answer for our technical POC. Right now, you will probably be working on a development solution. This may be something that you would be working on environment for now, but as this transitions to a Phase III and if and when it goes into a Phase III, it will probably be a government environment. But I don't know the answer to that one. I'd say take that to our technical POCs.

**13. Will use of prior work that is now part of a government owned baseline (i.e. services and containers) be permitted as part of this effort.**

You can develop something based on something you have already been excelling in to fit our needs, but it is a new development.





**14. Elaborating on an earlier question, if we cite prior work that is from a past SBIR, will we be deemed non-compliant?**

Yes, if you are citing something from past SBIRs saying “hey we have done this, but it is not the same work, but it’s the same concept”, I think that it should be fine. Again, focus on the technology. The technical evaluators will probably not focus on “oh that was a SBIR”, so I would focus on the technology itself.

**15. Are there commercialization restrictions on the work done here?**

It’s going to be an International Trafficking in Arms restriction. So anything that is going outside the US will have to be vetted and done through the Department of State as the law. Again, you have to think of the SBIR intent and that is for you to grow and develop within the US market.

**16. How do we engage your office and identify use cases for Phase I opps? We have a commercial product in beta mode that potentially can be used in cyber and comms.**

You can ask questions like these on SITIS which is on DODSBIR/STTR.mil but honestly, I don’t think we can put in use cases for Phase I for many reasons. Also, you cannot do a direct engagement with our program office. This is probably the best we can do in getting with the technical POCs. After the technical Q&A session, all other will come through the SITIS. We try to keep it fair with all the proposers and offerors.

**17. Is there a POC 004? If so, can you please provide.**

Again, this is the best interaction you will get with our Technical POCs. Aside from that, everything will go through SITIS for questions to keep things ethical and fair.





## SOCOM224-D001: Track Correlation/Data Deduplication for SOF Mission Command

1. **Since there is a correlation-tracking and data de-duplication assertion, does the solution also have to solve "source-of-truth detection" problem?**

I'm assuming "Source-of-truth detection" is referring to the fact that different sensors are detecting this thing, whatever this real-world object is and an associated confidence level, based on the sensor technologies. So yes – in the association, or correlation of tracks, the sensor confidence level should be taken into account to balance those type one (false positive), type two (false negative; aka a "miss") errors.

2. **Re: Track Correlator for MC, what kind of data is input, what is the format and content of that data, how is it ingested**

I think the easiest way to answer this question is – various. You're going to have data from commercial and government sources, organic sensors, in various formats, coming from various data sources. So really, the solution that we're looking for is something that's flexible enough to really not have a limitation in regards to the data input.

3. **...and what logic do we need to apply to correlate it?**

It's obviously space-time, but as we talk about that, referencing the confidence level associated with different sensors, some sensors can not only detect and object and classify that object, but some can uniquely classify that object and so those correlations should account for different combat identification types of things, whether it's a friend or foe, or something like that. The other thing is some of these sensors have staleness and other information, metadata that are associated with those tracks and obviously you'll need to apply extrapolation and interpolation between points in that track because the sensor may update at different rates and they have different confidence levels in the location, so there's uncertainty associated with that object's position in the 3D space. Also, uncertainty in each of the metadata elements associated with that. So there's not only 3D or 4D location, but all the metadata associated with the track as well.

4. **Will components of a fully integrated solution be considered for award?**

It depends. If we're talking about a component that can directly apply to this solution set, then yes. If it's 30% of the solution, something that's never going to be able to be tied without additional integration and development, I'd say we'd want to look at the effort from a holistic point of view.

5. **Are we operating in unclassified environment or do we need to have a cross domain solution?**

You do not need to have a cross-domain solution, one will be provided to you. We have data across all three – unclassified, secret, and top secret. Not saying that we need to do development on top secret, but those data sources are at various levels. ***This will be an unclassified work.***

6. **Can you give an elevator pitch of what you are looking for? (pretend we are 5th graders)**

What we're looking to do is take one of the big problems that faces the DoD today which is Big Data and apply it to a specific solution set. Within mission command, we have thousands of disparate data sets coming in at different moments. Those data sets often contain the same information, and really restrict our ability to move data through pipes. So when we look at a track correlation solution, and we're getting data from an unclassified ADSB, from a classified Air Force sensor, and from three other sensors out in the environment, that's four (or more) potential duplicated tracks coming in. What we're looking to do is take that data, de-duplicate it, and cut it down to the authoritative data. That reduces the amount of data that we have to pass to an operator, and also provides greater information and less clutter on the screen.





**7. What is the dollar amount for award?**

\$1.225 million for this Phase II effort. All of these dollar amounts are stated in the SOCOM instructions. Please review them so that you can make sure you're following everything that we're asking for.

**8. What is the mix of structured and unstructured data?**

It's going to depend on the use case and at this point, it's not something that we'll have fidelity on. But there will be a mix of structured and unstructured data.

**9. Is this required to be solved only at central C2 or also closer to the war-fighter. If latter, is there a spec for compute and comms limitations.**

I'm not going to put you guys in a box for compute and comms limitations. Primarily this will be solved at the central C2 level, but I can see applicability to closer to the Warfighter. With regards to taking a central track data, and combing in with their organic sensors. I wouldn't say there's a specific spec for compute on those.

**10. What's the use cases?**

At the high level, we're talking about taking various sets of SOF, or DoD, or commercial data and de-duplicating and correlating it. With regards to specific use cases, this effort, once awarded, has two travel items in it to get you guys up front and working with the specific SOF user for a specific SOF use case. With regards to mission command, what you really have to take home from this one is that mission command operates across all levels of special operations forces. Whether it be maritime, land, sea, air, there are a myriad of use cases. What we're going to do is focus on one SOF user, one SOF use case, and that will be determined after the award is given.

**11. What terminals will this pull data from?**

It will be various sensors, various data sources, various Application Programming Interfaces (APIs).

**12. Is the government interested in any of the GOTS code from previously executed programs with similar intent? i.e. DARPA's INSIGHT track correlation tools?**

I'd say yes. If you can take something that's already out there, that's GOTS code, approve it and integrate it into a better solution, absolutely.

**13. How long is the period of performance?**

Up to twelve months. Review the Proposal Submission Instructions for more information: [https://media.defense.gov/2022/Jul/07/2003031160/-1/-1/0/SOCOM\\_22.4\\_R1.PDF](https://media.defense.gov/2022/Jul/07/2003031160/-1/-1/0/SOCOM_22.4_R1.PDF)

**14. Does the prototype have to be deployed into a SOCOM environment or will a contractor environment suffice?**

The contractor environment will suffice, as long as it is a relevant environment to what we will be looking to deploy in that SOCOM environment. Meaning that the technical specifications are near somewhat to what we're doing within that SOF environment. Basically, you're not doing something that could only be used in your environment and it's able to be integrated with different SOF environments.

**15. Is this data real-time, near real-time, or some lower refresh frequency than that?**

It is going to be all of the above. We could be pulling from real-time sensor data, or we could be pulling from historical sensor data, or historical data base.

**16. I heard talk of feasibility study at the onset of this webinar. Documentation states it is a direct to phase II. Can you clarify whether Phase I or DP2 effort?**

It is a Direct to Phase II effort and a feasibility study is required.

Mina: So, that said, by policy directives, if we're going to bypass Phase 1 and go directly to a Phase 2, the proposals that come in must have a feasibility study. There is no page limit to the feasibility study, but keep in mind that we're looking at awarding by the end of September.

**17. Is there an existing system that this looks to replace and update?**





No, we've looked for something that's complimentary to existing systems, and this would be something focused specifically on the SOF use case.

**18. Can proposed phase 2 prototypes focus on a single data type? (example: aircraft tracks)**

Yes. That is a single use case and not particularly a single data set, as those tracks will be coming in in various data formats.

**19. Will the effort work solely off of existing trajectories and metadata from the target tracking estimators? Is any of the raw data available?**

Raw data has already been pre-processed by whatever sensor we're talking about. These tracks are an output of some kind of processor associated with that sensor. With regards to raw data being available, that will be available within the effort not ahead of award. Again, that'll be determined with the SOF user that we put you guys up with, to then work through that use case.

**20. What are the expected operational scales for the tracking/correlation algorithm? What ranges of areas, target densities, and time scales are expected?**

This is getting a little "in the weeds" from an operational perspective so I'll say again: We'll determine that tied into the use case. We have some "known" correlations, you might have two sensors that do maritime, and they cover the same area. Then we have, sort of the ad hoc correlation situations where a particular user, let's say it's a group of people that are looking at a technical mission and they pull up a computer vision feed that is then supplying tracks and as you overlay that computer vision feed, onto some of the other sensor feeds, then that would be like an ad hoc correlation and it's an intractable problem to do that for all possible correlation examples. So if you talk about this question, we're also looking at the ability to quickly spin off another instance of a correlator engine to handle those ad hoc situations as well as the pre-determined overlaps.

**21. Where is data being collected from? What types of media, location, reports, etc. Sensors? Capabilities?**

This will be use case specific. They will be from various sources.

**22. In less than 40 words, what's the problem?**

The problem is that we have massive data sets coming in from various sources, whether it's commercial, unclassified sensor data, or classified exquisite data. We need to take that data, correlate and deduplicate it in a manner that allows us to provide a single data source coming into the mission command system.

**23. If needed, will SOCOM pay for the cross domain solution in Phase II, or will that be government furnished equipment?**

We will handle any cross-domain requirements as the government. We're not buying you a cross domain, though.

**24. Will existing fielded or tested SOCOM C2 systems have to be interfaced as a part of this effort?**

No. It will be in a relevant environment, that can then be integrated with those C2 systems.

**25. Where and how is data displayed? Is there an existing interface to be improved on or is this to design a new interface from existing data?**

This will be to design a new interface from existing data. Again though, within SOF and probably the greater DoD as a whole, those use cases vary. So we're looking for something that will integrate with case-by-case data sets.

**26. Does SOCOM have a preference for how deduplication occurs?**

No preference, just as long as the complexities of the situation are dealt with as we've talked about – a lot of the metadata itself has uncertainty associated with it. Correlation should account for different confidence levels for different sensors. Radar, for example, you can't generally tell what kind of aircraft it is, but if you had some other sensors that can tell what type it is, then whatever classification information you have about it, should be in there. And the





result is correlated tracks should have the metadata merged. If you've got sensors with other metadata sets, the resulting correlated track should probably incorporate the metadata of multiple sets and in particular with regard to the data accuracy and confidence level as well.

**27. What government furnished equipment will be provided by SOCOM?**

Nothing. That is also stated in the statement of objectives.

**28. What will the communication cadence look like with the SOCOM customer?**

Similar to other SBIR efforts. The one difference is that we're going to pair the awardees of each of the topics up with a specific SOF user and a specific SOF use case. There are two trips that will be technical exchange meetings with SOF users with that specific unit. So there will be more communications than traditional SBIR effort, but we'll also keep those regular programmatic communication cadences with monthly MSR, reviews, and technical exchange meetings.

We are expecting users to adhere to agile practices. And we are really attempting to get the SBIR performers into our software factory within our existing pipelines and so that will certainly be a goal of this effort. There's a feedback loop of the issue tracking and other things that facilitate collaboration amongst the program office, the end user, and the development teams.

**29. Does the presented solution need to already have an ATO for your environment?**

No. We look at Phase 2 to get it to a prototype, and then Phase 3 would be when we bring you in, and under that effort, get an ATO. With our mission command systems combined.

**30. Can you elaborate on "real-time"? Is this tactical edge, "guy on the ground", JOC? Trying to better understand if we are talking PALANTIR, ATAK, etc**

Primarily looking at the strategic/operational environments where you do have more compute, bigger data, and more data coming in. I'm not going to say no to a solution that does track correlation and deduplication at the tactical level as well. Ideally it would be a scalable solution to handle the most complex data, all the way down to the most tactical data, with regard to deployments.

**31. Can the Q&A referenced "central C2" and "SOF environment" be interpreted as JSTOF or JSOTF-like?**

Yes, that would be, what I call an operational environment. In terms of the deployment vs. a strategic environment, which would be held at a headquarters, type of location. So HQ SOCOM vs. forward unit doing something somewhere in the world.

**32. Instructions state, "Feasibility documentation cannot be based upon any prior or ongoing federally funded SBIR". So can't leverage work from prior Phase I work?**

It certainly can include work from a prior SBIR Phase 1, Phase 2, or even Phase 3. What you're doing is creating a new solution, and that prior work can be a part of that solution but shouldn't be the basis for the entire solution.

Mina: Yes, I agree. So don't just say, I've already done this, here's the Phase II solution, here's what we did and dictate that as your final solution. Prior Phase 1 work – I don't see a problem with that, the feasibility study is already done. I would provide the technical merit of the Phase 1. The technical feasibility study from the Phase I not the Phase I itself, focus on the feasibility study.

**33. Is Mission Command considering government owned solutions to interface existing track correlation capabilities?**

We're considering any solution to create track correlation and deduplication capabilities. Again, if you want to use something that's already out there, and improve upon that in a significant way, or use that as part of your track correlation capability, no problem there from us.

Some of the aggregated streams that we're talking about are already correlated. The problem is, other tracked streams then include overlapping tracks. So then you still have the problem of – you have two or more tracks of a single real world object.





**34. Is this solution intended for centralized location or edge processing?**

Again, scalable would be ideal, but for some of these solutions, that's not possible. With regards to the amount of exquisite data and compute required. So initially we'd be looking at a centralized location for deployment and then eventually moving to the edge location where we can have a lighter weight processing solution where there are fewer data streams coming in.

**35. Is JSON in and outbound appropriate?**

Maybe in certain cases. The problem is that we're going to have track streams that come in whatever the generating system has, so those would be a combination of JSON, XML, protobuf, etc.

**36. Source code is one of the deliverables -- Can we assume this code carries SBIR Data Rights?**

Yes

**37. How important is dealing with time-late data?**

We've got different sensors and the technologies that they apply and their processing chains and the latency associated with some of those... You know, a radar may scan every second, for example. Another sensor may get a track hit every minute. Some systems may not update for an hour. Maybe each of those systems are using different technologies, different update rates, different times, then we have to correlate those. Trying to figure out, is that the same object or not? Kind of reference back to that 4D space and extrapolation and interpolation and being able to apply Kalman-like filter technology to figure out if that's the same thing projected in 4D space.

**38. How will this solution need to integrate with TAK/TRAX solutions?**

This is very specific to the use case. I'm going to say that both of those are probably going to be involved in the use case. TAK is the data source, TRAX is the data parsing service, or tool. They don't need to integrate with them as a requirement. You can assume that there will be TAK data and there will be data coming through a TRAX solution.

**39. Is cloud-provider agnosticism an organizational and technical objective?**

I'd say it's probably a requirement, whether it's to your solution it must be able to run on-prem (premises) or on different clouds, Amazon Web Services, etc. We don't want to be kind of put into a corner with the specific tool that then restricts our ability to run it on-prem or run it on each other's clouds.

**40. Can data collection sources be assumed as autonomous sensors, user created reports (SPOTS), unmanned human controlled? or is all data autonomously generated?**

These are various sensor types that are feeding this. Some of those sensors act autonomous in nature. Some of those sensors may be human interaction. I don't think you could consider all the data autonomously generated.

**41. If there is an existing GUI to be targeted, is there an ICD to consider?**

There is an existing technical interface that we would like to present this data in and we would want well-known, well-supported applications programming interface spitting out these correlated tracks, so they can be presented in this GUI. And from our user's perspective, we're going to want to be able to toggle between uncorrelated tracks and correlated tracks in case they want to view the original source or see the correlated result.

**42. Will all code be delivered to the government and owned by the government?**

The data rights – SBIR/STTR they protect the data rights, so we have government purpose data rights, I believe at the end. That said, there's no requirement to deliver the code to the government unless the company agrees to it and the company understands that this is government purpose rights, but we want to give the code to the government, want to give them unlimited rights. So, with that exception. Otherwise, it's government purpose rights, they don't





need to deliver anything, within the boundaries of the requirement, and that really depends on the effort that's being done in itself.

**43. What do you mean "don't have to deliver something within the boundaries of the requirements"? You mean outside the boundaries?**

Mina: Government purpose rights are probably going to be the outcome to protect your data rights, intellectual properties as a SBIR awardee. So that said, is there a specific set of code that is required to be delivered by MCS COP?

Chad: I'd say no.

Mina: So that said, there's no need to deliver any code. We'll just be looking at different ways of leveraging the final tool, which is buying licenses, whatever it is that we can agree on in a Phase III after the awards.

**44. Please characterize the envisioned users for this capability. MOS? Training?**

These will be common operating picture operators, so they'll be used to manipulating incoming data to use sources for commander there will also be intelligence use cases where they'll be intelligence analysts.

**45. [1/2] "Confidence in incoming data": Is it correct to assume the data is not clean nor always (correctly) labelled and that solution needs to deal with this.**

I would think that the data would be correctly labeled within the uncertainty associated with whatever subject technology that you're dealing with. Though you could be dealing with a radio frequency spectrum, radar, cooperative and non-cooperative signals. The radar – there's an accuracy associated with its ability to determine something's position in 4D space. Also, as I mentioned before, some sensors can positively ID a particular object, and others cannot. So that's where we get this mixture of confidence levels associated with different sensors.

**46. [2/2] "Confidence in incoming data": Is that confidence level available or does it need to be computed by the solution**

Some of the sensors report their confidence levels, other sensors do not. The solution will probably have to apply an educated guess to the confidence level from a particular sensor technology. That would likely be within the trade space, to consider when you're trying to correlate and figure out what is the best information associated with the track.

**47. Does the government look at Entity disambiguation as part of track deduplication?**

I'm going to assume that the entity disambiguation deals with the positive ID that I mentioned before and so if there are two tracks and one sensor says it's a vehicle, one sensor says it's something else, then you need to figure out whether that is the same object in space; in the context of earlier breadcrumbs or positions associated with that sensor system and trying to figure out whether that is the same object or not.

**48. Is there an existing solution SOFWERX is trying to improve upon or is this solicitation is trying to fill a gap where there is no solution?**

We are trying to fill a requirement for SOF. And SOFWERX is not the one looking for it. They are a neutral facilitator. It's the SOCOM Program Office that is looking for the solution. The awards will be done by SOFWERX as an intermediary, a Business-to-Business (B2B) contract. That's what gives great flexibility and speed in the awards for SBIR. There are cost limitations, but it's a great benefit.

**49. Are we assuming that some of the data is "evil"? In other words that there are actors that are trying to attack the capabilities of this system to throw it off.**

I think we come under the assumption that a sensor system or a sensor feed has not been corrupted by an adversary, however, a particular sensor or the technologies associated with that sensor could be operating within a contested environment that is operating within adversary countermeasures. So one of those may be a global navigation satellite system jamming, for





example. If the solution had a way to determine that from another sensor and account that into the confidence level associated with any sensor types that are trying to rely on different sensors that would be affected by those adversarial actions and incorporating that in to the solution space would be a positive.





## SOCOM224-D002: Natural Language Processing for Special Operations Forces

- 1. Page 1, 4th line from bottom of page, indicates an ‘SDA solution’. ‘SDA’ was not defined elsewhere in the solicitation description. What does ‘SDA’ stand for?**

Stands for SOF Digital Application – that is, the program executive office that this will be executing out of. And it is the Software program that is used at SOF AT&L.
- 2. Is there a foreign language of choice that SOCOM would like prototypes to work with?**

We do not have specific languages right now, can guess that it will not be English or one of the common languages that there are several commercial solutions for. When we look at this problem it is that there are commercial capabilities for well-known languages, but SOF operates around the globe and we need a solution that is applicable to the less common languages.
- 3. The ‘OBJECTIVE’ states, “Social media posts CEM, printed material and signs”. Are the signs captured in digital photograph images, full motion video, or both?**

Both, but full motion video may be a bridge too far, but it is something that we will eventually want to get to.
- 4. Are you interested in a capability that handles the OCR function, Language Translation and the automated data upload to the SDA? Or are you looking for an SDA?**

Again, SDA is a reference to the program office – not a particular application. So, optical character recognition function is something that would be involved as in the previous question, in taking a picture of a sign then optical character recognition would then translate that or perform other Natural Language Processing functions from that result.
- 5. What types of NLP use cases are you looking at? What are the data sources/sets?**

We have already said that we aren’t going to get into specific use cases. Those data sources could be documents, a picture of a sign (previous question). The real focus is trying to get these NLP functions done in the foreign language before translation. That is the difference we are focusing on here in the topic discussion. Because you have different colloquialisms, figures of speech that don’t translate well – and even if they do, it loses its meaning in translation. So, we want to be able to do the NLP functions in the foreign language prior to translation.
- 6. What foreign languages are of interest? What are some AFSOC analyst use cases of the desired technology?**

Again, we are not going to get into specific use cases and with regards to foreign languages that will be used those will be within the use case. What we are really looking at here is identifying a technology that is applicable across a wide range of use cases and languages. Again with regards with the user use cases or the integration with the SOF operator, when this contract gets awarded there are two specific events called out in the contract that allow for technical exchange meetings between company representatives and a specific SOF unit with a specific SOF use case. And that will be the target of testing the technology that is being developed.
- 7. Can we initially focus on one foreign language?**

Yes, that will be determined through the use case and it probably will not be a common language.
- 8. Any NLP performance benchmark that we should utilize?**

I don’t think there is any particular benchmark that we are dealing with, of course with any statistically related data analysis problem we would want to balance those Type 1, Type 2 errors, as well as have accuracy that focuses in on something that you can have some confidence in. We didn’t mention this before, but this should be focused in on users that are not





necessarily language experts and so think English only user being able to apply NLP technologies to foreign languages.

**9. Are there translation speed and/or data volume metrics for the prototype to achieve as threshold or objective measures?**

I would say that I am not going to restrict you to speed or data volume metrics. Obviously, if it takes three days to look at one document, we are looking at something that is infeasible for our application. But there is nothing specifically that I am going to call out. Because what we are going to do here with less standard languages is cutting edge. We are looking to find out what those processing and time requirements are going to be as its developed.

**10. Is there an operational need coming from the services? Who would be an end user?**

Again, I have identified the requirement as SOF. SOF as a force is global, meaning that we operate all over the world which is certainly not unlike the services. But it certainly spreads us real thin to language expertise within the ranks. And that is where we are looking to build the requirement there.

**11. With an expectation of TRL-6, is "[in the colloquialism of the native language] sarcasm, figures of speech, and jargon" realistic for TRL-6 in 12 months.**

I'd say partly it would be realistic for TRL-6. Again, we are doing innovative research here, so TRL-6 is a goal. If we show good progress toward this requirement with our objectives, then it is something we continue beyond that 12 months in a phase 3 where you would be continuing that work. But really it is identifying/developing the technology that can apply to this use case and then we can develop from then on out. So, not looking at you guys to eat the whole apple here, we're just looking to take a bite off and show promise and capability for investment to move forward.

**12. What gaps do recently release large language models (from Facebook and Hugging Face (e.g. Bloom)) have that you're seeking to fill?**

Mainly the collected, exploited material, the various languages that are not covered under the traditional commercial models, and then the SOF specific use cases that the user may be looking for that is not standardly covered in commercial capabilities.

**13. what type of data sets are available? How are they labeled; what features/entities are labeled?**

Again, this will get into use case specific. So, you can think various data sources. Right now, a big focus for you guys would be publicly available information. Social media posts, news articles, anything in that regard in terms of actual paper documents are certainly going to be covered under collected exploitable material. That is not something that will be determined until a time when the contract is awarded or effort and you're working with those SOF users.

**14. Are you looking for a solution hosted on SOCOM premises/infrastructure, or are will you be able to use the public web to access the solution?**

This will all be determined through classification of the material. If we have classified material that we are looking at that will be on premise solution. But in terms of publicly available information (news articles, social media) that would be absolutely be acceptable to access with the public web.

**15. What kind of proof are you looking for in the feasibility study? If we are bringing a COTS solution will a discussion of success w/Other DoD missions suffice?**

You really want to discuss the technical merit of it. You want to discuss, that you have this commercial off the shelf solution that is going to be the basis for innovation, and you want to continue to provide the information needed for how it meets our requirements. It depends on what commercial off the shelf solution that is being discussed. If it is Google translate or





something that is well across the enterprise, as a capability we want to see more with regards to how you are going to modify that COTS solution. If it's not one of those big COT solutions that is already out there, then it is innovating on top of that.

**16. Would you explain how a solution could potentially satisfy the need of the end user?**

Again, I'd say it's going to vary by use case but they will take various data sources, and will pull out the language from those various data sources for a user that is not inherently a native speaker of said language, or languages.

You have an English-speaking user that does not have expertise with regard to the particular foreign language they are dealing with or a particular context. So the information that is presented to that user needs to be in English, since - we are talking about doing these operations in the foreign language prior to translation. Of course, then we are going to need some translation so the English-speaking user can understand what is being conveyed. So then the user interface has really got to pick out the context while in the foreign language and obviously present that in the English language for the end user. That's an additional context with regard to how the user uses the interface.

**17. How many awards do you envision granting under this topic?**

I don't know. We do have limited funding available within the program, but if we do see more than one of these that will benefit the user it could always jump from there. It could be multiple, but I'm not going to put numbers.

**18. Do you have a template for the D2P2 technical volume? The one that is linked appears to be for a much larger proposal than the 10-page limit defined here.**

The D2P2 we are asking for 10 pages and we mentioned that in the beginning. You might find something on the DOD website that says 25 pages but stay within the 10 pages and find as much information as you can put in there. Focus on the technical aspect of it and the merits of the expertise that says "here's how we can do it". Stay within the envelope, but we are capping it at 10 pages. And we have been doing that for several years.

**19. In prior related efforts (e.g. Maven), a major hurdle has been preprocessing that data and preparing it for a model pipelines. is this of interest to gov reps?**

The tendency has been to try to stay out of the model training lane and leave it to folks like Maven to attempt such things. So from that perspective we would like to have models that are in working order and be able to integrate them into our systems.

**20. If it's an or a set of under resourced languages, then do you provide data for them?**

I'd say that would be case by case basis. I am not going to say they are under-resourced to a point that they are public. But certainly something that the government would be able to step in and provide if required.

**21. Would you be open to considering partnerships by best-in-breed NLP companies with best-in-breed translation companies?**

If it follows the SBIR guidelines and isn't illegal – yes, absolutely.

**22. What is the ultimate goal of the project? The range of NLP applications is so broad that the goal doesn't reveal itself.**

The goal, much like SOF, is a broad category so we are going to have a lot of different goals. You guys are going to have a specific use case with specific language and that is going to tease out what the particular unit is looking for. So I would say the ultimate goal of the project is to show progress in NLP with a language that is not the most common with applicability to publicly available material and information to provide inference to non-native speakers.

We have seen lots of solutions that translate and then do the natural language processing, but what this project focuses in on is the NLP prior to translation.





**23. Can you provide an example of a "less common language"?**

Farsi

**24. Given that the users will be English speakers, what is your vision/expected solution for translating the results/output of the NLP into English?**

For example, if you did name entity identification or recognition in the foreign language and then the user interface for the English-speaking user you could show a context of the language and the named entity's that you picked out. Of course, the user interface is going to be one challenging aspect, and then how do you then show insight to the user on what you did and now you're presenting them in an English user interface.

**25. What do you envision the budget for this project being? \$750k? More? Less?**

The max budget is \$1.225 Million over the 12 months – that is for Phase 2, obviously for Phase 3 we will go much larger than that. In terms of mission command as a program, what we like to do is take innovative technologies, group them out, and then justify investment in the future with much longer and more robust budgets. Success with one effort could then lead to a follow up with additional languages and additional features. They are coming a part of the greater mission command architecture.

Please read the SOCOM Instructions – read it in its entirety, it has a lot of information, how long it would be and other information that are critical in following to submit a proper proposal so that you don't fall into non-responsive category.

**26. Can SOCOM please define how much data (megabytes? gigabytes?) and what quantity of articles (10K? 1M?) is desired to be part of the prototype?**

The prototypes proving out the processing of that data, so I don't think that the quantity of that data is as important as the capability to process that data. Sorry to give you a roundabout answer. I just wouldn't put a number of megabytes, terabytes, theoretically it can scale across all of this – it would just be a matter of time and amount of processing required.

**27. What are the hurdles that prevent an existing solution from being applicable?**

The language that is going to be applicable to SOF is not always going to be applicable for commercial solutions. It's also just the number of those languages and the flexibility of the solution, and of course the accuracy of the results that are brought out. And then of course, the application of the programming interface that is well defined and well supported in industry with the technology perspective so that we can present those services.

**28. Are there any limitations in the technologies being used? Can we use cloud computing, for instance?**

Yes, absolutely. Preferred.

**29. What is the expected format submission?**

I don't know if this talks to the SBIR program as a whole or a specific part of the system. If you want to reword that we will help. I am going to refer back and say to read the SOCOM instructions – it has all the information that we covered in the first half an hour. It will tell you what we are looking for – how many pages, which template to use, what we look for in a volume, it has the presentation and feasibility study that needs to be filled to go to Phase 2, etc. those are pieces of the big puzzle so make sure to read the instructions.

**30. Will native speakers of the foreign language of interest be provided?**

That is full time employees provided by USSOCOM for language services, not tied to this effort. So that will be part of SOCOM. Having native speakers of languages of interest would be great as long as it follows the restrictions of the topic.

**31. If I understand the goal, higher accuracy should be preferred to faster output, correct?**

Correct.





Well, I'd say with the strategic use case level yes, there is a limit to the stalemate of information with regards to its usability, but the accuracy is very important within this effort.

**32. Will SOCOM provide examples of Collected Exploitable Material (CEM), printed material and "other potentially valuable sources of data" for the prototype?**

Yes, once the effort has been awarded, we will have those two technical exchange meetings with users and that will determine what we are going to be evaluating – both from a language and a material perspective.

**33. Doc says "This topic is accepting Direct to Phase II proposals only" but it also lists a feasibility study during Phase I. Why is Phase I listed if this is DP2?**

Again, the policy directive is saying that for us to bypass Phase 1 and go directly to Phase 2 we have to have a feasibility study. So, the feasibility study is not done. Now – Phase 1 equals to feasibility study as a final report, but a feasibility study does not equal to Phase 1. The need for it is required. The phase 1 topic write up in itself is something that DOD requires for us to post the topic in itself, but it is a Direct Phase. And it has that in the topic write-up. It says note that this is the Direct to Phase 2. So with that said, this is a Direct to Phase 2 with these requirements and that write up.

**34. What is the genesis of the topic? Who wrote the topic?**

SOCOM Mission Command Systems is the writer of the topic. The genesis of the topic is that SOCOM and DOD at large as requirements for NLP in order to make valuable insights with data that is not accessible by non-native speakers.

**35. The document mentions "current NLP solutions" that you've tried. Could you mention a few NLP solutions, models, etc. that you've experimented with already?**

I don't think we should mention any particular solutions. As we are working on our Market Research, it seemed like most of the solutions focus in on NLP after translation. And the concern that this topic tries to address is levels of accuracy and context when you do these NLP processes after translation.

**36. Follow up question re preprocessing (e.g. PDF-->JSON) - this is a big lift in and of itself. Would you support a proposal focused on this part of the problem?**

We definitely support this part of the problem. Again, \$1.22 million is not the end of this effort by any means, so what we really look for is taking that part of the solution with maybe one entity and partnering with another entity that can handle the other portion – basically showing the progress required to perform a Phase 3 effort to say – hey this is something that we will bring on as part of the program to start stacking bigger dollars against.

**37. The SBIR program needs to be periodically reauthorized. The legislative authority expires 30 Sep 2022. When do you expect to sign contracts for this effort?**

Ahead of 30 September 2022. We are working really hard to award this before the end of this fiscal year and within our current authorization. We have success and have been able to do that in the past. We are using a FY2022 NDA Section 852 which is an authority for SOCOM and SOCOM only to award SBIRs via SOFWERX as an intermediary. And that business-to-business contract has all the goodness and protection with the speed and agility of business-to-business contracts. This legislation that I just mentioned is within the boundaries of SBIR, so we have to have the authorization. That said, we can get it done in that speed and we are working hard and are going to get it done before the end of this FY in order for us to proceed forward.

**38. Can I assume that since it is an NLP task and speech recognition has not been mentioned, the modality is only written text?**

At this point, I'd say yes.

**39. What do you consider the greatest technical challenge?**





The figures of speech, the colloquialisms, sarcasms, in a foreign language is probably the biggest challenge and bringing those accurately into the English language.

**40. [1/3] "On source correlation during NLP processing": Does the NLP solution need to provide multimedia ingest (e.g. speech recognition, or speech-to-text).**

I think that previously we mentioned that this topic is not going to focus in on speech, but some sort of written document. The original format may have been in some form of verbal communication, but this SBIR topic is focused in on some sort of textual document.

**41. [2/3] "On source correlation...": Does the NLP solution assume multiple interwoven sources (e.g. multiple speakers) and provide contextual separation.**

Again, not doing the spoken.

But, within the problem space is the recognition – you've got entity linking, relationship extraction, co-reference resolution, so that if you have the context or paragraph, or document that you're dealing with, you need to disambiguate the different entities that are involved there – and as those get used with pronouns or anything else that is parallel to the foreign language then you would need to be able to provide the right context for those different entities that were involved in that document.

**42. Are we able to / is it advantageous to include an MOU with submission if we have committed Customer / End User from AFSOC or other command to support transition?**

End User and commands don't support transition, the program office does. So, I would say to support transition no. Any End User requirements or any End User support that you have is valuable to the program just not with regards to a transition. The program will transition a capability based on its merit and it's answering of the requirements. But I have seen a lot of SBIR proposals come in with memorandums of support, and similar from actual users or other components and that is definitely a possibility that we have been evaluating – but not required, technical merit.

So – this is definitely not required, but if it's something added in there, that is fine. The Air Force requires it because the way they do business is different than how SOCOM does business when it comes to SBIR – they have the almost \$1 Billion SBIR program, which SOCOM has a fraction of a percentage of that. So that said – our technical POCs write our program offices, and they are the ones that go out there and say: hey this is what we need and here is how you're going to fulfill it and here's how we are going to transition it to Phase 3. Even internally in SOCOM, the PO signs a transition agreement with our entity director of SOCOM to say: we are taking this and we have a plan for this for a Phase 3/transition. That said, we don't need any support from the outside because the Air Force mechanism is to say who really needs this? Go ahead and find the users that would say yes, we want this kind of solution. It's not required as I understand that the Air Force has gotten a lot of our small businesses in the habit of getting MOUs, etc., but it's not something we require or look at in that much detail.

Memorandum support is not considered at all SOCOM evaluation.

**43. [3/3] "On source correlation...": Are we assuming multiple languages in a correlated sources?**

If we are targeting a single language for the effort, no. If we are expanding the effort to target two or more languages, then yes. We are going to look for multiple languages within a correlated document. The assumption here is that the document is going to contain a single language and then that is the language that will be targeted and known at that time.

**44. Does the solution need to work on a secure environment?**





I would not develop any capability that would not be able to translate over to a secure environment. So, if you have dependencies on external applications that require it to function all in the unclassified that still has value but would be a severe limiting factor to our use cases.

**45. Is there a TRL level that is expected when submitting a proposal and when completing Phase I?**

This is a Phase 2 – the TRL of what we’re shooting for is Level 6. That 6 can be with 70% of the features, or 100% of the features. Really, we are looking for a technical capability that has merit and can be built upon as needed based upon user requirements.

**46. Can you provide an example of a less common language?**

Already provided one, but Farsi – something that is regional vs. global.

**47. When discussing uncommon languages, are you also including various dialects?**

Yes. This can get convoluted really quickly. We are looking at Use cases – if something is picked up that has a million different dialects, we will probably target the popular ones and then look at adding on other required dialects.

**48. FYI The question about preprocessing wasn't related to training...it was related to processing. Poor inferences from models if not investing in preprocessing**

I’ll answer this by saying, within the overall goals of the SBIR topic, the trade space we are dealing with. If you as a technology expert within this realm you need to decide where to put your emphasis in presenting a proposed solution to the Government.

**49. Regarding feasibility study question earlier: Should the feasibility study be done in advance of the proposal? Or during the Phase II?**

It has to be done in advance of the proposal, and those instructions are on the SBIR page.

**50. Is there a language preferred for NLP?**

Again, this is going to be based on user input that will be completed after awarding the contract and the technical exchange meaning.

**51. Is Chinese one of the languages of interest for in-language processing?**

Again, we have a requirement across the globe and probably every language out there has a requirement to it in some degree. The application of the technology you’re proposing will be based on a user case after award of the contract. Initially looking at the technological capability being proposed versus what somebody can do with a particular language.

**52. To clarify a previous question, will a language expert be provided by the government to aid in development and validation of this solution? Yes or No?**

You will be provided access to users, but we will not be providing a language expert. We are looking to get language expertise from a solution that may not have said expertise in language.

**53. Has the PEO Digital Applications individual changed in the last two months?**

That’s irrelevant to the topic and that’s irrelevant to the technical nature of the questions here. With any changes that happen with the PEOs there is no foreseen changes. People change from time to time, that’s the nature of our military staff.

**54. If we partner with a university (as a sub) that has foreign graduate research students on an F-1 student visa, is it acceptable for contracting?**

Negative. We cannot take any F-1 student visas, only US citizens and US permanent residents can work on this effort.

**55. Who is the end-customer for this effort? Which command within SOCOM/AFSOC is sponsoring this?**

USSOCOM Mission Command Systems is the program for this effort. And once, awarded, we will determine an end user, and a use case will then be provided. Again, when you look at topics like





these they spread across the DoD, they also spread across USSOCOM and that is where we are going to focus on specifics to bring out the technology and identify capabilities that can be brought out to these cases.

**56. Five Eye citizens/ Canadians? Are these usable as local Americans are difficult to find.**

Again, those are great partners, but they are not US citizens and because this is SBIR and is restricted. SBIR focuses on United States citizens, permanent residents, and solely doing the work in the United States, the 50 states.

**57. Where can I get more details on the use case?**

You are going to get more details on a use case if you get picked for award and we provide you access for user and trainings. Ahead of that, I have provided everything on the use case that I am going to provide.





## SOCOM224-D003: Low/No Code Data Manipulation and Discovery for Special Operations Forces

**1. What questions are Special Operations Force Commanders typically trying to answer from the data analytics?**

It varies. It is mainly tied to inference of any amount of data or disparate data sets.

It certainly depends on the mission thread we are talking about. It may be a position location information, it could be other data sets, just trying to do both supervised, unsupervised, semi supervised kind of data analytics to different data sets to answer specific questions and some of those things is asking the system to provide associations that aren't necessarily intuitive to the user and so there is a lot of flexibility within this problem set.

**2. Can you please give a couple of examples of what data sets (like audio, video, texts etc.) are of interest and which other systems would load them?**

Like I just mentioned, I may have a position location information where the user is trying to identify patterns of life, co-travelers, and common space time intersections. If they are trying to identify individuals or entities that are associated in some way, maybe those real-world associations aren't readily apparent to the end user and so just looking at the data analytics engine to be able to provide that insight.

**3. Could you please describe one or two realistic Special Ops scenarios in which commanders would need data analytics?**

No, but I will take a chance to answer this question as best as I can with regards to data analytics. One of the things that special operators and the DOD aren't, are data analysts or data scientists. Big data to them is a problem. They aren't always going to have that data analyst or data scientist on site to provide analytics, so what we are looking for here is to provide a data analyst or data scientist like capability set to a user with less aptitude than someone who has a four-year degree in. So, this is making data more manageable, configurable, and useful for users that this isn't their day job. That is applicable across all scenarios.

**4. What are the cybersecurity requirements for any software developed? Will the DoD RMF have to be applied in Phase II?**

No, it does not have to be applied in Phase II. Anything you develop will eventually have to be used on a classified system, so I would not do anything out of the ordinary with regards to "hey it is vital that YouTube is utilized to cross the formation for this". But the RMF for any RMF or ATO requirements can be covered in the Phase III. This is the initial development of a prototype capability and then Phase III would pick up that ATO work.

**5. The requirements state that the Contractor shall "deliver the software to a USSOCOM designated Cloud Based Software Factory"; can you describe what this is?**

This is a typical Dev Sec Ops environment that SOCOM has created for itself. It uses fairly common tools you see out there GitLab, Jira, Confluence, and security scanning tools to be able to generate the kinds of security artifacts that are necessary to provide confidence in the cyber security experts to allow that code to deploy. So, the thinking is, we would like to get the SBIR efforts within our environments. What that allows us to do is align transition and support transition ability. So, if we can, depending on the SBIR that we are talking about, if we can align it with the rest of the system that MCS/COP has out there, then being in that same environment allows you to get maybe a copy of the latest code if we are trying to integrate that with a particular SBIR technology.

**6. How do the current commercial solutions to no code low code data management not meet the government needs?**





I'd say, pull it back from the government, we say Special Operations Force's needs. Current solutions that are out there right now have not met the requirement for SOF users. They are not utilized by SOF users, and the goal of this effort is to create something that is applicable and tailored to a SOF user and a SOF use case. That could vary from the source of the data, or to the way that the tool is deployed.

**7. What platforms/environments will a solution be required to run on (e.g. Windows vs. Linux, headless server with a client GUI vs. locally installed on a laptop)?**

Probably not locally installed on a laptop, if you can do it on a laptop, that would be great. If it is something that is not as processing intensive and it can go on a laptop, it will certainly be valuable, but I see this more as a micro-service that can be run in a cloud agnostic environment, so it can either be run on-prem or multiple cloud environments.

**8. My interest is coercing SQL queries to run in real-time by logic table technology (LTT). Where to fit into the SBIR pie? Vendors use query-templates from 1980s.**

I am not sure how to answer this question. The main focus on this is that we have users that are not data scientists and we want to be able to provide them the tools that a SOF Operator or a SOF Intelligence Analyst can take different statistical tools and processing stubs to in some cases to build a processing pipeline that takes sources of data and runs it through this pipeline and then produces some output that is relevant maybe to further processing pipeline, but then being able to save and share these algorithms or processing pipelines with others and in doing so in a way that they don't really need to understand the nuances of the technology, but the user interface allows them (that non-data scientist) to understand the basics of what is going on in the processing chain and being able to piece those things together to provide a result they are trying to answer.

**9. Are the data sets stored? Near real-time data flows? Real time sensor feeds?**

They will be all of the above. When we talk about doing data analytics a lot of the time when we are doing data discovery, it is taking one set of data and being able to correlate or discover a new set of supporting data. That is a data analyst or data scientist tool saying "hey here is two sets of data that are correlated through X" – we want that ability for SOF to be able to perform that function. Not necessarily on the ground with a rifle in their hands, but certainly unsupervised from a data analyst or data scientist that is sufficiently educated.

**10. Is the emphasis on this primarily on back-end data analytics, or visualizations to help users discover the data and work through the analysis process?**

If you mean by back-end data analytics, like elastic search or graph store, it would be how to utilize those capabilities at the nacent level for per user versus people who are adapted at utilizing those systems. Then yes, absolutely visualizations to help users discover the data and work through the analysis process.

**11. How can a DP2 proposer show that the unspecified-in-the-topic "minimum performance parameters" are met? Will they be provided (and soon)?**

Since this is an unclassified process, the performance parameters are intentionally vague. I don't think that when I'm evaluating this topic that I am looking at any key performance parameters. I am looking at technology that is being proposed and the applicability to my mission.

**12. Given previously mentioned use cases (e.g. space-time intersections), is the desired solution to automate geospatial analysis across disparate datasets?**

I would say that that can certainly be one part of it, but not just geospatial analysis.

Yeah, there's a myriad of real-world things that have a geospatial component to them, and not all things do. So, from that perspective, it would be both.





**13. Will sample/representative data sets be shared with awardees related to specific use cases of interest?**

Yes, absolutely

**14. From a technical perspective, in what form are the provided APIs and sensor control interfaces a solution has to integrate with (e.g. HTTP REST, C/C++, etc.)?**

There isn't any particular API technology that the program office is looking for. It just needs to be well supported by multiple programming languages, as well as well supported in industry. We are not looking for any particular technology solutions from that perspective.

**15. Can you give one or two examples of desired data pipelines? E.g. (Data: troop elevation)->(Transform: filtering by time)->(Display: histogram of elevations)**

The possible scope of analysis and mission threads that our users deal with can certainly be this example, but they are so diverse it is difficult to properly characterize all the potential uses of data analytics that our users have.

**16. Please elaborate on "present complex theories, processes and products in a way that is easily trained and implemented across the SOF formation."**

This is speaking to data analysts and data scientist type activities and being able to implement that at the SOF level. Thinking with regards to simplicity, it something can be trained via a YouTube video, as a DIY scenario to create a particular user thread. Something that's at that level, a tool is able to simplify the process enough that it provides a usable tool for an operator to do that of discovery and manipulation.

Let me throw out one thing - you have all kinds of things within the data science space, things like hyper parameter tuning and things like that where somehow we have to be able to present that in such a way that the low code/no code user can understand and then understand what it means to do the task that they have been given and being able to use those tools to produce an output that they understand really what the output is saying.

**17. Can you elaborate on the "scalability of services" requirement? In what ways do you envision a solution leveraging/integrating with GPU enabled cloud resources?**

Really this is going to be determined by the area that the operation is being conducted in the strategic level where you have access to the GPU enabled cloud resources or is it in the tactical level where you have a tactical data base tied to a sensor data base where you are doing discovery on. Now, in terms of scalability of services, I'd say if the tool requires heavy compute across a multiuser base, say we have 1,000 users in the system, that would be at strategic cloud usage, if it's at the tactical level it would be at the most 10s of people would be using the capability and it would be much smaller from both a data set perspective and number of users implementing the services.

I'll throw one other thing out there, so of course the focus is on the low code/no code user, but from an algorithmic perspective if some of the algorithms that are running in the background are run much more efficiently on GPUs than CPUs, but even tensor processing units or field programmable gate arrays may be relevant in regard to particular algorithms, so the real emphasis is choosing the right compute engine for the right algorithms and presenting that to users so they can use it in a way that they don't have to be a data scientist. So maybe the solution itself picks the right processing compute engine for a particular pipeline that was created.

**18. From a technical perspective, what is the MCS/COP which is provided by the government for this effort and in what ways will solutions have to integrate with it?**

It will not have to integrate with it. We are not going to declare a single MCS/COP solution. This is meant to be a tool almost like a microservice with its own GUI.





**19. The Government describes a Direct to Phase II SBIR, but the SBIR program document includes Phase I and Phase II scenarios. Please clarify.**

It is a Phase II. The Phase I feasibility study is required as a requirement on a Direct to Phase II. This is a Direct to Phase II. Congress institutionalized it about 6 years ago for those situations where the government has determined that private industry has already done their internal Phase I feasibility studies. Since private industry has done them, the government does not need to do a Phase I. In order to be eligible for the Phase I, the firms, the proposers are going to have to validate that they in fact did a feasibility study and that what was determined to be feasible in Phase I would be the solution that they are proposing for your statement of objectives.

**20. We are a US-based firm w/employees in other countries. Is that an issue? We understand non-Americans can't be paid w/SBIR, my q is if we're eligible to apply.**

The general guidelines for the SBIR program that it will be conducted in the United States by American citizens. I don't want to guess at that one. Generally speaking, the answer would be no, but I think we need to get back on this one and give them the right answer.

**21. Is there a template for Phase I feasibility analysis?**

If you look at the DOD instructions, they do have a prescribed format on what needs to be included on the feasibility study. I would make sure you go back into the Phase I section of the OSD agency announcement to look at what those mandatory requirements are. As it relates to what you will be submitting here, there is a volume II, which is your technical proposal which needs to capture many of what the DOD requires to be there, but you also then have a presentation of a maximum of 15 pages where you can also balance the information that is required. So, between the two of them, the volume II and your pitch you should have enough to meet the DOD requirements.

**22. In the context of this topic, how interested are you in collecting/processing/getting insights from unstructured text? Or is this mostly about sensor data?**

I would say that this is about all data. Certainly collecting, processing, and getting insights from unstructured text would be part of it.

Yeah, this would be the whole realm of AI/ML, it is kind of computer vision, natural language processing, you know associating different linkages, pattern of life, time series plots, so all kinds of data and sources providing analytics from that.

**23. Could you please describe one or two realistic Special Opps scenarios in which commanders would need data analytics?**

Not going to go into that one on this call, as previously stated. I will say that once the technology has been selected and awarded, there will be two technical exchange meetings with users to tease out and determine that use case.

**24. Can we assume that certain software (specifically a modern browser like Google Chrome) will be installed on a target platform?**

Yes

**25. Through which mediums, technology, or platforms will end-users interact/manipulate the data? Such as tablets, wrist-worn device, command center computers, etc.?**

Probably not going to be on a tablet or wrist worn device. Data discovery will probably at a workstation, whether that be a laptop or desktop.

**26. Based on the AI/ML references, is one of the goals to have AI/ML learn and anticipate how end-users would want data to be manipulated for a specific use case?**

I don't think that would be excluded from the problem space. Hypothetically speaking, if you had a solution that could learn from other users that this is how they normally manipulate a data set to learn a specific answer to performance of that answer in mind. I can't think of a specific use case for that, but we aren't giving use cases right now. I don't think that it would be





excluded from a particular solution where the machine would suggest a particular way to process information to get the answer within a particular mission thread.

**27. Can one company submit a separate SBIR for each of the 4 topics?**

Yes

**28. Would it be sufficient if the prototype is able to crunch data from csv data files without implementing APIs for database / historian accesses?**

I wouldn't limit an input format to comma separated values, but the solutions should also take in streaming kind of data in real time or near time, potentially as well as pulling out data from a historical data set. So users might be trying to figure out patterns of life based on historical data and maybe comparing that to something that is happening in real time or near real time.

**29. Is there a requirement for use of the platform at the edge?**

The edge most processing that you should be doing this against would be at the laptop and/or desktop level. Again, I don't want to throw a specific hardware requirement against this, if it's something like Alien ware gaming laptop or IBM business laptop that is something that is certainly in the realm of possible for SOF forces. Yes, both platform at the edge and strategic cloud and/or on prim can be used in most cases.

**30. Are there any specific use cases already in mind?**

Nothing that we currently have in mind. That will be determined by the user after award at the two technical exchange meetings that will be scheduled.

**31. What rules of inference are desired? Abduction, Induction, Deduction? Only the last is provable in bi-valent logic. Also, is fuzzy logic desired?**

As we mentioned before, the uses cases are so varied that depending on the mission thread and the data set that's involved and the answer that the user is trying to get out of that data set, it's hard to limit this solution set to a particular inference. I would also say fuzzy logic is desired within the overall solution space

**32. What is the purpose of the government provided "dedicated vendor integration team"? What activities do they perform?**

They will provide any requirements for integration onto a SOCOM designated cloud, integration within a SOCOM network, or any other required integrations. That could be tied to data sources or to existing capabilities that can be leveraged in the effort.

**33. Is the Government contemplating multiple awards?**

We aren't not contemplating multiple awards. It will all be tied to funding, availability of funding and if technical solutions meet or exceeds what we are getting after, then yes multiple awards are possible.

**34. Is analysis of unstructured data of interest or if focus on structured data (e.g. in databases)?**

Both.

**35. how should we expect the special Ops to acquire data? field collection or data assigned to them by a central command?**

Both.

**36. should the performer assume that this capability may have to run in DDIL environments and accommodate low SWP?**

We aren't going to hedge it on either scenario. Scalable capabilities are preferred. But within a DDIL environment, if you are not connected to data sets, you would be restricted to data on hand whether that be sensor data, or something brought with you that would certainly be a possibility.

**37. can you tell us a bit about capabilities that you think missed the mark - in terms of usability?**





I don't think it's so much a question of terms of usability, it's the ability of a user to use said systems. So the training required or inherent background of a data scientist or data analyst really limits the use of some of tools for SOF users.

**38. are geospatial, temporal, and link/node the most likely visualization output for this data? Does this also have to end up on TAK?**

Geospatial, temporal and link/node are certainly the ones that would most likely come to the mind if you talked about visualizations output. I don't necessarily think that that is the limit for visualization. Some of these data sets will be looking to industry to provide new ways to visualize things, especially as you are trying to convey provenance in other things associated with the data set and quickly as you are trying to show these visualizations and making sure the non-data scientists can understand the meaning of this data and so multiple visualization techniques involved. Eventually this may end up on a TAK UI for a prototype, but it is not a requirement.

**39. can you share your vision for 'low-code'?**

Hyper parameter tuning and things like that as the analyst like that picks different algorithms and models to apply to the data set. They may have some ability to understand python and some other TensorFlow things like that. So if we have some users that know a little bit of code, they should be able to use that expertise and leverage that in the overall solution where they would be able to get more out of the system than if the system was designed for a no code user only.

**40. Is this effort specifically for MCS-COP or for other SDA systems (DCGS / FADE)?**

This effort is for any SOF user or SOF program that deals heavily in data. So this would be tied to PM/DCGS for the intel side and MCS/COP for the operations side and also digital ecosystem for SOMPE in terms of other digital SOF applications.

**41. What do you think will be the greatest technical challenge?**

Taking tools that were built for data scientists and data analysts and moving it to the level that a SOF or other DOD operator can understand, infer and utilize said tools.

I think getting back to the visualization there is how do you present this data whether its tabular, pictorial, or whatever else that it doesn't escape the user what the really the gist of the inference is.

**42. Expected future classification of transition (i.e., Phase III) effort?**

Most likely at the top secret or secret level.

**43. The GFP/GFE/GFI section of the Statement of Objectives says "the government will provide the contractor access to [...] MCS/COP"; is this not accurate?**

That is accurate.

**44. What is the definition of success for this work?**

A non-technical SOF user able to do data discovery and manipulation utilizing tools and techniques that a data scientist or data analyst would use.

**45. can you characterize your requirements for this capability to accommodate real-time or streaming data?**

Some of these data sets may be coming in real time or steaming kind of nature or near real time and if the analyst has used the tools to provide some sort of processing chain based on historical data and the model provides some inference or output that is relevant within that historical data, whether that be pattern of life or whatever else, and then there was a streaming data set or near real time data set that they would like to run the same model on. I think that's some of the context of what we are talking about.





**46. What is the deployment platform/target? How will the end user be consuming the solution (web, mobile, Windows desktop) or something else?**

Most likely be web or Windows desktop in regard to a workstation connected to a network and the data sources.

**47. If we already have a phase SBIR award with AFWERX is there a specific person to reach out to.**

Not sure the applicability to this topic. Please reach out to the SOCOM/SBIR office.





## SOCOM224-D004: Human Machine Teaming for Reduction of Operator Cognitive Load

- 1. Do you want the operator to be able to say "hey Jarvis pay attention to this object on my HUD" (closed request) "hey Jarvis find me a new route" (open request)**

Without getting into the specifics of the mission thread, the computer will be told, and it needs to understand, what the mission thread is and what the mission goals are. It then needs to understand how objects in the real world can affect that mission. As we get into querying the end user parts of the missions, they will offload it into the machine, that's where we will get into that. I don't think we can properly characterize whether it's a closed or open request at this time.

- 2. Does technology that converts visual light to auditory sound fall within the scope?**

I would say no. Generally speaking, the human's visual cues are a much broader bandwidth to the human brain than auditory or tactile or others. Hypothetically speaking the human machine teaming could provide a solution that included multiple human senses, but obviously visual is a key one.

- 3. Can we request a GFE NIPR laptop (or equivalent) for testing, or should we assume that all test events will be conducted on contractor equipment?**

Unless we are loading the capability onto a SOCOM network, or onto a SOCOM owned hardware or cloud space, then it would be conducted on your equipment. We will not GFE any equipment under this effort.

- 4. Are contracts going to be multiple contracts by general topic? Just looking at it the range of use cases is pretty wide for any one small company to manage.**

Ideally you would get your first use case and then go from there. Then the company would scale and grow as appropriate, or you would look for teaming partners. The initial contract is a Phase II, up to 1.225 million. A phase three follow on would be a five-year IDIQ, and that's something we could snap off task orders as required.

- 5. Is it preferable for human-machine teaming to be interactive, e.g., "Alexa what is the weather today?"**

Without getting into the specific mission thread that the government has in mind, when you're dealing with a human machine teaming context, let me get into a hypothetical: if the machine presents some information to the operator, maybe it's recommendations 1,2, or 3, and the human says, "I don't like any of those options, give me a fourth option".

So there could be situations like that where there is some interaction between the human and the machine to arrive at an answer that is acceptable to the human.

- 6. Is the off-loading from human to machine need to be just-in-time capable? (edited)**

The machine will have different information feeds to it so it will understand those real-world objects. Those may be positions of other objects in 4D space, or textual data that provides insight into a particular aspect of the mission space. And so the machine would have to be understanding how changes and different information its getting affects the mission thread. Then, providing information or recommendations to the human and those change dynamically as the situation in the real-world changes. The real world is changing, in the context of the mission thread timeline. In that perspective, there are probably aspects that you can consider that are just-in-time.

- 7. So I guess the solution must come within a time frame correct?**

So the time frame again, as we mentioned, the specific use case the government has in mind has a time frame associated with it. The winner will be given that time frame and the particular mission thread that we are dealing with, so then you will be able to apply that time frame associated with that to whatever needs to be accomplished before that. So, from a scaled-out





perspective, maybe that means that you have to apply more compute than you thought that you would have to apply, but that's something that will have to be identified.

**8. Sounds like there's a Q&A component. Is the warfighter actually asking a question?**

Not sure I fully understand the question or statement. There will be end users involved in the topic and we will be asking those end users what part of the given mission thread to they want offloaded to the machine to reduce their cognitive burden. Within that mission thread context and those things that they want offloaded and having machine understand the real-world context of that is kind of the reference frame that the topic deals with.

**9. Will a CAC card be required for misc. activities associated with this effort (e.g. email, chat, etc.)? If so, will SOCOM sponsor the application for one?**

You won't necessarily need a CAC card. We do have most of our Phase II SBIRs on CAC cards for activity, not necessarily the network, just for access to base and other uses. So, if the effort dictates it, then yes we would sponsor a CAC card.

**10. Will you consider prototypes built at private expense?**

If you already have something that has proven out feasibility of a capability that would be included in the feasibility study. "Hey I have this thing I did and it created a feasible product", then you would build off of that to build that prototype SOF specific for the effort of this phase II.

**11. Does a lightweight Augmented Reality headset that can run heavy-duty machine learning algorithms satisfy the criteria of Phase II or feasibility study of PI**

A piece of hardware is less important than the algorithms that it is running. We would be interested in the algorithms that it is running. If it just so happens to run on an augmented reality headset, great, but again software-specific.

**12. Did I hear you say that a great late solution is better than a fair one on time? Please explain.**

Depending on what the mission thread is, if I, as a human, have to do some action within some decision loop and it is too late for the mission context and the machine provides me an optimum answer too late, that doesn't really help the mission decision or mission thread. When you're dealing with machine learning or AI, maybe you can't guarantee that through mathematical proof that this is the optimum answer, but a human expert assessing the performance for the system would agree that it is a good solution to a problem and it's in time, it beats an optimum solution that is too late.

**13. Did I hear you say that a great late solution is better than a fair one on time? Please explain?**

As soon as SOFWERX receives the recording of this Q&A session, we will be posting it on the webpage.

**14. The SBIRs are unclassified, but in this call, use cases haven't been detailed "in this unclassified environment". Will use cases be given to performers once funded?**

Yes.

**15. CDRL A006 description is missing from the documentation on the SOCOM website. Could that be posted? SOO mentions "Epic Hypotheses review" - can you elaborate?**

The approach that we, the program office perspective is taking is scaled agile framework. And epic hypotheses deals with a set of features that defines the overall capability that are posited in the context of a hypothesis and whether, if you do something or not, will that be realized as a business value to the end user, so that's what that terminology means. The earlier reference to DODD3000.09 – Reference to the descriptions of the human in the loop, out of the loop, or on the loop and not necessarily connotating that it's dealing with any weapon systems.

As far as the missing documentation: It's actually in there, we checked, but when we did a combination of the CDRLs into a portfolio, the computer, for some reason, renamed them awkwardly on some of them. We got it, we know we got that fixed and they are there.





**16. Is 004 human machine learning focused on maritime or land environment?**

Answer not going to specifics at this time, it will be based off of the user case that'll be combined with the technical solution presented. So it could be any environment. It's going to be dictated by the proposal that we receive, select, and then identify users to participate in.

**17. Can you please clarify: "prototypes developed with other than SBIR funds that are provided at the end of Phase I feasibility studies will not be considered"**

SBIR funds that are obligated on Phase 1 SBIR contracts are needed to focus on the actual development of doing a feasibility study. We've received many proposals that have come in asking us if we would consider prototypes at the end of the feasibility study. And our answer to that is no because that's not what the intent of the feasibility study is. We want to know what the innovative technologies that can be brought to bear on a particular problem are available and that's what we want the Phase 1 dollars to be spent on. We are not expecting, nor do we want a prototype at the end of Phase 1.

**18. What type of hardware should the solution be hosted on? E.g., PC/laptop, mobile device, specialized/classified network etc?**

Again this comes down to the question of scaling – where we could have an individual operator that has human-machine teaming all the way back to the headquarters, where you have the power of the cloud. We're not going to point you in any one direction. The use case, given by the SOF user will do that plenty. We're going to start off on the strategic side of things where we have better resources, and then move toward tactical as the capability allows.

We won't necessarily be processing on a central processing unit (CPUs) so because of the processing power to handle some of the complexities of problem space, an offeror may decide to offload some of the algorithms onto another processing engine.

**19. Can COTS devices already in use, be a substitute for phase 1 feasibility Phase 1**

No because one of the mandatory requirements for a Direct to Phase 2, in order to be eligible, is that the proposer has done a feasibility study. So the fact that there's COTS devices already in use, would not substitute for that.

**20. Are you looking for hardware that enables easy human-machine teaming, Software for decision making or a combination of both?**

We're not looking for any particular hardware or excluding any particular hardware. We're looking at an overall systems approach, maybe even a system of systems that both the human and the machine can team up to, or the machine provides the decisions part, in offloading whatever the human is asked to be offloaded to the machine.

**21. What type of sources would be providing information on the "real world objects, their interactions, mission goals, legal/policy/doctrinal/physical constraints?"**

I'd say at the unclassified level, what would be a great one would be DoD policy. So that would be a legal or doctrinal policy thing. And then mission goals or interactions as well would be determined by the specific use case. The theater of operations or the kind of mission that the user is performing. Again an awardee will get those types of technical exchange meetings to identify all of those required data points to then move out on the prototype development. Depending on the particular mission thread that you're dealing with, there are numerous doctrinal publications that are even public domain (publicly accessible) that would deal with that whatever particular doctrine you're dealing with. Talking about the rules of engagement and other things like that, there's a Chairman of the Joint Chiefs of Staff Instruction and it is a classified document that talks about rules of engagement, standing rules for use of force, and so there are various legal policy and doctrinal publications that are involved.





**22. can we team with another company to develop it? do we need any letter of agreement to deliver the contracting office?**

I believe that would be called out in your proposal. Any teaming arrangement and a distribution of the funding, you can team with another company to develop it. There are specific SBIR rules that apply to it. It should be 50/50 with the workload from the two companies.

**23. What are some examples of things that would be included in the Feasibility Study that would not be in the Technical Volume of the proposal?**

The feasibility study is to make a determination of what's in the art of the possible. The technical volume would include a description of what was investigated and what solutions would be the best to satisfy the phase 2 Statement of Objectives.

**24. Where can we find the feasibility study template and/or checklist? Is there a URL?**

The feasibility study itself is the phase 1 topic write up that's included in the BAA. So if we were moving with just a Phase 1 on this topic, that would be our feasibility study that we would be requesting. But this is a Direct to Phase 2, and we're only accepting Phase 2 proposals. So what the firm needs to do is to convince the government that they have done their own internal feasibility study, and that the results of that feasibility study has resulted in a solution that could be used to satisfy what the requirements are in the statement of objectives. That description of what was done by the firm is included as an "Appendix A" to the Phase 2 proposal and that's all clearly described in our USSOCOM specific instructions.

**25. We allot machine to navigate UAV to target site using cognition and radio wave data. Is it something that SOF can trust in such technologies? Thanks.**

I don't think there's distrust in capabilities. But there will have to be a man in the loop to provide direction from the human side. Part of the challenge, ultimately, and it's not any hard stretch to imagine this, but garnering trust from policy makers, decision makers, as well as end users, is part of the challenge with implementing any AI systems.

**26. What type of Human -Automation team tasks the AI-DSS intended to support? Is the focus on tssks like Piloting a UAV, Controlling a robot etc?**

It could be any of those. I'm not going to give any specific use cases. But in regards to any task that a team or individual needs to complete, there could be a human-machine capability. There are several physical systems as well as human-machine teaming. Just from the perspective of I'm getting some answers from a rack of computers, so there's objects that may have smart agents on them in the real world, as well as intelligence agents that are dealing with a decisions support system that are not cyber physical systems and so without talking about specific use case, all that stuff would be interesting in the long run.

**27. What is the degree of keeping a "human in the Loop" that is required?**

The human will be looking at the machine's decisions and recommendations and then providing the input. So there will be no autonomous machine making decisions with this effort. If you look at DOD 3000.09 Autonomy and Weapons Systems that is one example that talks about machine, human in the loop/on the loop/out of the loop. So you can get some context from that instruction. But again, we are not going to get into the particular use case for this.

**28. Is there a cyber component to the solicitation?**

If we are referring to cyber security, yes, not necessarily an ATO or RMF type package in the Phase II. But you would not be the Phase III without that being accounted for. If you're referring to the cyber domain, we are not going to get into the specific things for the use case, but SOF as well as the rest of the DOD is looking to operate in an all domain for the future.

**29. Is the objective to provide the right information in the right time or provide decisions recommendation as well**





Yes. And when you talk about the complexity of the problem space, you don't necessarily have to guarantee the optimum recommendation or decision. A good answer in time beats the best answer late every time.

**30. Is the objective to provide real time decision support during the performance of the task?**

Yes.

**31. What are some of the challenges with the existing approaches? What do you see as the main challenges with integrating AI into this operational context?**

I think mainly it's the complexity of the problem space when you're asking the machine to understand the nuances of the real world in the context of a particular mission thread, and all of the factors that are involved because when you look at the environment it is not only the physical world, it may be spectral space as well – depending on the mission type. So, weather effects, other things like policies, statutes, rules of engagement – when you start overlaying all that complexity on this problem space that's the thing that I see as the most challenging. And being able to reduce that problem space, on the fly, as it changes underneath your feet, and keeping the operator in the know of what's changed, what's going on, and how that's changed the machine's recommendation for the user in real time and dealing with the combinatorics associated with the possible permutations within the problem space.

**32. Would you expect that there would be data scientist involved in this topic (unlike Topic D-003)?**

Not at the command level. There may be access to a data scientist, but that would likely be for configuration of a data set at the headquarters level. And I would say that the application of that use case would be few and far between.

**33. Are you looking for tools that automate functions to reduce data coming at the operator?**

Yes. We have humans with cognitive overload. During the mission thread we'll be identifying specific things that humans want to offload to the machine to reduce cognitive burden, so that machines can do things that the human isn't as good at. The human can retain things that the policy, or statute, or just the operator preference doesn't want the machine doing. That's where the human-machine teaming comes into play.

**34. Is this topic focused on improving visualization?**

You can't divorce the ability of the machine to convey information to the human from this problem space and so, from that perspective, it is human-machine teaming. So the eye gate is a particular rich way for the machine to convey particular information to the human.

**35. In the context of this topic, is there an expectation that Phase II would include evaluating solution using human subjects?**

So, the human factors and how it effects the human I don't think is part of the scope. If the end users do lead us down that path, we would have to use other entities involved that would focus in on gauging the interactions of the human and seeing their performance gains. It will be more qualitative than quantitative.

**36. Has a front-end analysis been conducted by SOCOM/SOF to identify the specific challenges faced by end-users? If so, will that be shared with awardees?**

Yes.

**37. Will access to end-users be provided based on the specific use case defined?**

Absolutely.

**38. If a team has not performed a Phase I on the topic, but has performed equivalent work to demonstrate technical feasibility, can they apply?**

Yes, you're going to need to do the feasibility study – that is required, it shows that it is technically feasible for you to complete that work.





The team does not have to have performed a Phase I on this topic. As a matter of fact, it's not a direct to Phase II. What the mandatory requirements are is that a feasibility study was conducted by the potential proposer, and that it resulted in a technology that was determined feasible by objective requirements.

**39. Is the goal for this topic to be integrated with the other three SBIR topic solutions D001-3 during the Phase III?**

Absolutely, it's going to be integrated with a lot more SBIR topic solutions than in this proposal. It will be integrated within mission command systems which is a huge program within the command. So it'll be integrated in that environment.

**40. Would the AI/ machine learning be integrated w/in operator equipment (boats/ vehicles) or on the operator's person? What type of environment?**

When we think about defense operations, you have a strategic environment, which you could think of your combatant commands, your theater environment which can include a task force or unit that is forward that is in a specific theatre of operations, and then you can think of a tactical environment. Within that decision space, obviously the easiest to attack with a heavy compute, human machine teaming application is on the strategic side. Probably the place that it would be most valuable is within that tactical environment. I'm not going to push this in a certain direction, I think there's requirements across all levels and it's just dependent upon use case. So we would be interested in both, on vehicles, within the operator's person, or a strategic data center, or mission command platform. If you look at it holistically, in the future, one could envision centralized autonomy logic that understands the capabilities of any edge autonomy logic. Then there's a synergy between those things so that the centralized logic is not having the edge do something that is beyond its capabilities, or where it's not leveraging its capabilities.

**41. Will you be able to provide data? What levels of classification will it include?**

It will most likely be unclassified data, at least initially in Phase II. And you'll get that through the use case technical exchange meetings and that will identify any government information that is required. And then if we look to move into a classified environment for whatever reason within this Phase II, it would require a DDT54 and a couple other things to square that away.

**42. Phase II requires the awardees to install a prototype. Will the prototype be deployed into a SOCOM environment, or will a contractor environment suffice?**

It could be either, probably going to be easier in a contractor environment but we are amenable to pretty much any solution.

**43. Can you elaborate on "real world objects" with example(s) and their importance to this effort?**

If a helicopter was involved in the mission thread, then the machine would need to understand what that helicopter's involvement is as a friendly asset and that there are threat assets in the real-world that can affect that object's ability to perform its mission. For example, wires or surface-to-air missiles. If it is a human, then there may be known data sets of IEDs along routes.

**44. How can a Direct to Phase II proposer show that the unspecified-in-the-topic "minimum performance parameters" are met? Will they be provided (and soon)?**

This is an unclassified environment that we are operating in so direct performance parameters at the classified level will not be provided at this point. The performance parameters with regards to the technology is what we are looking for in the initial proposal. We do have a particular mission thread in mind, but that will not be specified at this time. There are mission timelines associated with that mission thread and so the decision support system would need to be able to process data and inferences within the timeline associated with that mission space.

**45. How many companies are you planning to award the first/second phase and is there any specific criteria that you are looking for? Threshold vs Objective?**





This is just a Phase II, not a Phase I. We are currently funded to award up to \$1.225 Million. However, we do have additional funding and if technologies come through that propose varying techniques that are applicable and amenable to what we are looking for, then I can see multiple awardees. There is no specific Threshold vs Objective criteria. This is a huge problem set and we are looking at the technical capability and marrying that up against the user to generate a product in the end for the specific use case.

**46. What type of work would suffice as a feasibility study? Does it need to be a Phase I in a similar domain for another DoD Component?**

It needs to be in the technology area of interest. In other words, the topic title of this Phase II effort.

**47. Is the focus of this topic on the AI or is part of this also on the visualization and user interface for interactions across the human/AI?**

It is very difficult to divorce the technology with the human machine interface since this is a human-machine teaming topic. The focus would also certainly be on the visualization of the user interface and how the human and machine team up together.

**48. What should a successful Phase II prototype be able to do? Are there minimum requirements that need to be met?**

There are no minimum requirements that need to be met. This is a Phase II effort, so it is past the initial research stage, but certainly something that will want to see matured as much as possible, given that we are going to give you a specific user to engage with and develop the use case. It is going to have a lot of background parameters, such as weather and other things that they are looking at and reducing the cognitive load for that individual. There are no minimum requirements, but a successful Phase II prototype is something that gives the user and the Program Office the confidence to move it on to Phase III and to invest more funding into it. Then it will be applied to multiple SOF use cases.

**49. Is part of the desired solution an AI system capable of performing causal inference/reasoning, to "understand" how people and objects interact in the world?**

Overall, we want the machine to understand enough of the real world as well as the machine thread that is applicable to the user at the time, to be able to provide the inference. We will start a few things out there. The machine will need to understand weather effects to the mission, the rules of engagement, how real-world objects effect the people or equipment that are involved in the mission, and what the goals of the mission are. From that perspective, I would say yes.

**50. Is a machine performing computer vision to assist humans of interest? 2. Can you provide an example of an AI-DSS that is of interest to SOF?**

There are a lot of computer vision solutions out there. When we are talking about AI-ML in this context of human-machine teaming it goes beyond the scope of computer vision. That may be one of technologies that are involved. We are focusing on decision support systems and so that may be where the machine can provide recommendations within the mission thread about courses of action. I do not want to get into any specific use cases, but there are a variety of decisions that the human needs to make about a particular mission thread, whether that be a strategic operation or a tactical level of war. We will be assessing what human cognitive load we want to offload to the machine later when we get the end users involved. That will be a part of the process - identifying what cognitive loads they want to offload onto the machine. It will be within allowable policies and statutes associated with the machines making the decisions.

**51. Since this is a Direct to Phase II, should we do the Phase I feasibility study before even submitting a proposal?**

Yes, that is required.





**52. What level of decision making is SOCOM looking for? The cited works go from Brigade-level C2 decisions, down to individual-level decisions in the field.**

This will be based on use case. If you are potentially looking at a strategic level of decision-making use that will have commander level decisions - that would be at the O6, O5 level, down to an individual level user if they are controlling a platform that you are looking to have inference on.

**53. Is the primary focus the underlying tech, which is applied to a particular use following award- Or do submissions require specific use cases prior to submission?**

It is the former. No particular use cases are required prior to submission. We are more interested in the technological capability and innovation being pitched from an agnostic point of view.

**54. Does the feasibility study need to demonstrate an innovation in human-machine communication, or is it feasibility of AI to perform in a novel military domain?**

The solution does not have to have any innovative human-machine communication. Recognizing human command is not necessarily the interface that is required. Any time that you are dealing with human-machine teaming, a challenging part of the problem is the machine conveying to the human how the machine came up with a particular recommendation, answer, etc. That is part of the problem – How does a machine provide provenance and other aspects of its answers to the human.

**55. Are there particular mission use case examples?**

Yes, but we are not going to provide them at this time. However, I can tell you that this is not restricted to a single domain: Air, land, sea, cyber, or space. It's important to note that this will be primarily for SOF Operations.

