

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

### **ADDITIONAL INFORMATION**

N/A

### **TECHNOLOGY AREAS:**

Information Systems

### **MODERNIZATION PRIORITIES:**

Artificial Intelligence/ Machine Learning | Control and Communications | General Warfighting Requirements (GWR) | Network Command

### **KEYWORDS:**

Translation, Natural Language Processing, Foreign Language, analytics, machine learning, artificial intelligence, Special Operations Forces

### **OBJECTIVE:**

Special Operations Force (SOF) operations and intel analysis support often need to understand information from data written in foreign languages. Social media posts Collected Exploitable Material (CEM), printed material and signs and other potentially valuable sources of data in a non-native language are a large challenge to those without linguistic specialization in that language. This effort applies natural language processing technology to glean operational relevant information for SOF.

### **ITAR:**

The technology within this topic is restricted under the International Traffic in Arms Regulation (ITAR), 22 CFR Parts 120-130, which controls the export and import of defense-related material and services, including export of sensitive technical data, or the Export Administration Regulation (EAR), 15 CFR Parts 730-774, which controls dual use items. Offerors must disclose any proposed use of foreign nationals (FNs), their country(ies) of origin, the type of visa or work permit possessed, and the statement of work (SOW) tasks intended for accomplishment by the FN(s) in accordance with section 3.5 of the Announcement. Offerors are advised foreign nationals proposed to perform on this topic may be restricted due to the technical data under US Export Control Laws.

### **DESCRIPTION:**

This proposed solution applies Natural Language Processing (NLP) technology to glean operational relevant information. The desired solution will allow users not proficient in a target language to utilize an easy to use user interface(s) (UI) to rapidly glean information from multiple mediums in order to inform intelligence and operational activities. The UI will support a native English speaker yet will perform NLP processing in the native language (before translation to English) to ensure errors induced by translation losses are limited. Current NLP solutions, although good in the English language, have limited foreign language capability. Any foreign language NLP artifacts will be combined with post-translation NLP artifacts in such a way that the English-only user can easily see the results. For example, named entities in the foreign language will be combined with Named Entity Recognition (NER) results after translation and presented to the user in a context where the associations are clear. The NLP need includes NER, relationship extraction/entity linking, sentiment analysis, terminology extraction, coreference resolution, Automatic summarization (text summarization), and any other value-added service available per a vendor's technology needed is a solution that handles [in the colloquialism of the native language] sarcasm, figures of speech, and jargon. It is assumed that some collected exploitable material (CEM) specific component outside systems will handle native language Optical Character Recognition (OCR) and users will be able to supply OCR results to the SDA solution. Air Force Special Operations Command (AFSOC) intelligence analysts who assess the impacts of Information Operations (IO) require the interpretation (machine translation and NLP) and display/visualization of sourced Publicly Available Information (PAI) (foreign textual data). Unlike current programs of record, our tool(s) will allow analysts to quickly establish sentiment analysis baselines and identify adversary disinformation campaigns by using advanced processing techniques to interpret foreign text data.

### **PHASE I:**

Conduct a feasibility study to assess what is in the art of the possible that satisfies the requirements specified in the

above paragraphs entitled “Objective” and “Description.”

The objective of this USSOCOM Phase I SBIR effort is to conduct and document the results of a thorough feasibility study (“Technology Readiness Level 3”) to investigate what is in the art of the possible within the given trade space that will satisfy a needed technology. The feasibility study should investigate all options that meet or exceed the minimum performance parameters specified in this write up. It should also address the risks and potential payoffs of the innovative technology options that are investigated and recommend the option that best achieves the objective of this technology pursuit. The funds obligated on the resulting Phase I SBIR contracts are to be used for the sole purpose of conducting a thorough feasibility study using scientific experiments and laboratory studies as necessary. Operational prototypes will not be developed with USSOCOM SBIR funds during Phase I feasibility studies.

Operational prototypes developed with other than SBIR funds that are provided at the end of Phase I feasibility studies will not be considered in deciding what firm(s) will be selected for Phase II.

NOTE: This topic is accepting Direct to Phase II (DP2) proposals only. Proposers interested in submitting a DP2 proposal must provide documentation to substantiate that the scientific and technical merit and feasibility described above has been met.

#### **PHASE II:**

Develop, install, and demonstrate a prototype system determined to be the most feasible solution during the Phase I feasibility study on a natural language processor for Special Operations Forces.

#### **PHASE III DUAL USE APPLICATIONS:**

This system could be used in a broad range of military applications where collected exploitable material, social media and other data sources are ingested in large quantities but cannot be analyzed due to linguist resource constraints.

#### **REFERENCES:**

1. The Power of Natural Language Processing: <https://hbr.org/2022/04/the-power-of-natural-language-processing>
2. Your Guide to Natural Language Processing (NLP): <https://towardsdatascience.com/your-guide-to-natural-language-processing-nlp-48ea2511f6e1>

#### **TOPIC POINT OF CONTACT (TPOC):**

TPOC-1: SOCOM SBIR

PHONE: N/A

EMAIL: [sbir@socom.mil](mailto:sbir@socom.mil)