Topic: AF141-054

DECISIVE ANALYTICS Corporation

Building Open-domain Semantic Search (BOSS)

To date, traditional search techniques such as Information Retrieval (IR) and Information Extraction (IE) have been executed independently providing information that is not analysis ready and is limited in scope. The Building Open-domain Semantic Search (BOSS) technology will combine the strengths of IR and IE into a single capability and address the mutual limitations of both. It will allow users to explore data from any domain through a search that returns structured results. This search capability is semantically-based, allowing users to search text data based on what it means instead of the numerous ways that meaning might be expressed. Decisive Analytics is an Employee-owned company providing data processing capabilities that move beyond simple extraction by automatically organizing, identifying, and extracting activities and relationships from large corpora of text, imagery, video, and audio. We are seeking opportunities to transition this capability to programs where identifying threats from a large corpus of multi-source data set in real time is a difficult challenge.

Technology Category Alignment:

Machine Perception, Reasoning and Intelligence
Advanced Computing/Software Development
Human Computer Interfaces (HCI) for Decision Making
Information Collection/Management
Synthesis/Analytics/Decision Tools

Contact:
James Nolan
jim.nolan@dac.us
(703) 414-5002
http://www.dac.us/
SYSCOM: NAVWAR
Contract: N68335-17-C-0545

Corporate Brochure: https://navystp.com/vtm/open_file?type=brochure&id=N68335-17-C-0545

Department of the Navy SBIR/STTR Transition Program

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.

SPAWAR SBIR PM 28 Dec 2018

Topic # AF141-054

Building Open-domain Semantic Search (BOSS)

DECISIVE ANALYTICS Corporation

WHO

SYSCOM: SPAWAR

Sponsoring Program: Distributed Common Ground System-Navy (DCGS-N)

Transition Target: Distributed Common Ground System-Navy (DCGS-N)

TPOC:
(619)553-5263

Other transition opportunities: DCGS-A, AF DCGS, Securities and Exchange Commission, Special Operations Command, FBI

Notes: BOSS will combine the strengths of Information Retrieval and Information Extraction into a single capability allowing users to explore data from any domain through a search that returns structured results. This search capability is semantically-based, allowing users to search text data based on what it means instead of the numerous ways that meaning might be expressed.

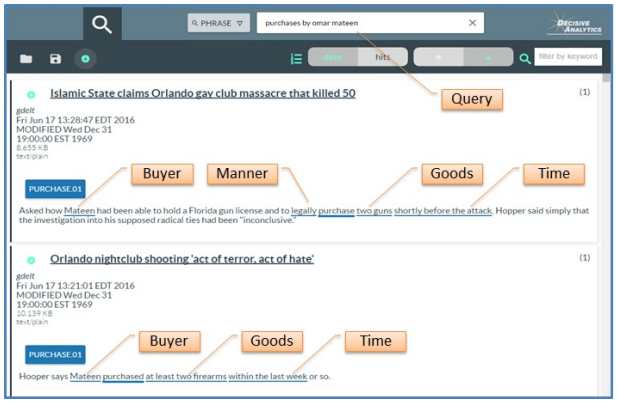

Image Courtesy of Decisive Analytics Corporation, Copyright 2018.

WHAT

Operational Need and Improvement: The quantity of text available to Navy Intelligence analysts has become too large for any single person (or even any organization) to read, identify all relevant information, and fuse the discovered information with the existing knowledge base. In addition, there is a vast array of formats and writing styles being used across the variety of data sources containing unstructured text (e.g. MEPED, IPIRs, USMTF, and Kleig-Lites). Existing search and natural language processing (NLP) technologies can only partially address the problem of exploiting unreadably-large amounts text. Search products currently provide excellent data exploration tools allowing analysts to cover a broad range of information. Natural Language Processing capabilities provide structured results that are analysis-ready, but are only able to accurately operate on a limited fraction of the available data. Combined, these limitations require the user to spend significant time and effort to manually filter and organize the search results for analysis.

Specifications Required: The goal of this topic is to research and develop an advanced indexing and search capability that combines Information Retrieval (IR) and Information Extraction (IE) methods to: (1) dynamically model user information needs, including building models for retrieving entities, events, and relations; (2) rapidly search large (web scale) volumes of textual data to identify relevant information; (3) return relevant information with precision and recall which exceeds the current state-of-the-art; and (4) enable users to refine or change their information needs over time through interacting with the system.

Technology Developed: To address this, the BOSS data discovery capability will be semantically-based to allow users to search based on the meaning of the data rather than having to determine the numerous ways a concept might be expressed.

Warfighter Value: Information Retrieval provides excellent data exploration tools, but does not yield analysis-ready information. Information Extraction provides structured results that are analysis-ready, but the information scope is limited. The BOSS approach to semantic search will allow users to fully explore information based on operational needs, whether those are specific needs (e.g. such as “Chinese ships with ship-to-air missile capability”), general needs (e.g. topics such as “piracy events around the world”), or points in-between (e.g. less constrained event type queries, such as “recent military activity in the South China Sea”).

WHEN

Contract Number: N68335-17-C-0545

Milestone	Risk Level	Measure of Success	Ending TRL	Date
Testing in DCGS-N Capabilities Assessment Lab	Low	Successful tests by DCGS-N Capabilities Assessment Team	6	February 2019
Exercise Testing in DCGS-N Inc 2	Low	Transition to DCGS-N Inc 2	7	June 2019
Operational Testing in DCGS-N Inc 2	Med	Completion of all testing and accreditation requirements	8	November 2019
Deployment within Operational DCGS-N Inc 2	Med	Full capability within the deployed DCGS-N System	9	May 2020

HOW

Projected Business Model: The business model for this effort is a mix of license fees and custom development for various user communities. Through our experience, we know that this capability requires some custom development for different domains to produce effective results.

Company Objectives: Below is a list of transition partners that can benefit from this technology.
DCGS-N / DCGS-A / AF DCGS: Semantic analysis is a powerful and flexible tool to help analysts understand the large corpora of unstructered data. DAC’s automated capability results in a large reduction in the labor required to perform document analysis. Inclusion of these developed methods of processing large scale data repositories of documents will be important to these customers.

The Intelligence Community including the CIA, DIA, NSA, NASIC, MSIC, NGA, and others have requirements for analysts to automatically make senses out of large amounts of unstructured text data.

Potential Commercial Applications: Financial Analysts: Investment management and other financial market analysts have a need to analyze complex data sets about companies and industries. Many of these problems mimic the problems facing intelligence analysts: analysis must draw from huge amounts of data, a variety of expertise is required, and understanding and managing the data is crucial to success. Through a self-funded marketing effort, DAC has past performance and relationships with banks, hedge funds, and investment advisers in the financial markets and has validated this requirement.

Contact: James Nolan, Vice President, Analytic Technologies
jim.nolan@dac.us (703) 414-5002