

Discovering, Analyzing, and Understanding Improvised Explosive Device Documents

Forrest R. Frank

The Challenge: The Joint Improvised Explosive Device Defeat Organization needed a way to improve its understanding of the whole-of-government IED-related science and technology landscape.

Background

The Joint Improvised Explosive Device Defeat Organization (JIEDDO)¹ was established in February 2006 as a joint activity to “focus (lead, advocate, coordinate) all Department of Defense action in support of the Combatant Commanders’ and their respective Joint Task Forces’ efforts to defeat improvised explosive devices (IED) as weapons of strategic influence” (Department of Defense 2006). To better understand and ultimately defeat adversary IED use against U.S. and coalition forces, JIEDDO sought information from scientific and technical activities overseen by the Joint IED Test Board, individual DoD Components, and other U.S. Government departments and agencies.

In February 2013, the President released his IED strategy, *Countering Improvised Explosive Devices*, which aimed to expand the Administration’s counter-IED focus by building on existing policy and strategy that establish and implement measures to discover, prevent, protect against, respond to, recover from, and mitigate IED attacks and their consequences (Executive Office of the President 2013). The strategy stresses the importance of a whole-of-government approach to countering IEDs.

IDA was asked to help JIEDDO improve its understanding of the whole-of-government IED-related science and technology landscape. The goal of the initial tasking was to identify counter-IED strategy stakeholders based on their production or consumption of scientific and technical information, intelligence, operational lessons learned, and formal responsibilities established in law, regulation, and policy. IDA was able to identify more than 1,000 individuals in 200 departments and agencies who either contributed to or consumed IED-related information in the federal government

To better understand and ultimately defeat adversary IED use against U.S. and coalition forces, JIEDDO sought information from scientific and technical activities overseen by the Joint IED Test Board.

¹ In 2016, JIEDDO was renamed Joint Improvised-Threat Defeat Organization (JIDO) (subsequent to the work described here).

alone. However, aligning JIEDDO-produced or JIEDDO-sponsored information with these individuals or organizations proved more difficult.

Applying IDATA to IDA's Research

JIEDDO's scientific and technical information processes supporting IED technical solution test and evaluation had assumed a wartime mentality. JIEDDO's focus was on delivering counter-IED capability (i.e., materiel solutions and some related tactics, techniques, and procedures (TTP)) to warfighters. Delivering documentation of scientific, engineering, and testing activities to JIEDDO or the Defense Technical Information Center (DTIC) was deferred in favor of fielding counter-IED capabilities from 2006 until 2013. As a result, JIEDDO lacked understanding of its trove of scientific and technical data and reports that could support the whole-of-government approach to countering IEDs.

In late 2013, IDA was asked to capture scientific and technical data and reports distributed to JIEDDO's information systems, as well as accessible data produced by JIEDDO-funded test and evaluation activities. IDA was also asked to develop a process prototype that would make this information and data available for indexing, search, and retrieval. We thought that the IDATA capability, even in its preliminary design, would be a potential partial solution.

The first challenge was locating and retrieving data. We collected more than 7,000 documents from multiple classified and unclassified

JIEDDO information systems. We also collected more than 1,000 documents from information systems operated by the Army and the Navy, where JIEDDO had provided funds, tasking, or data for use in tests, evaluations, and experiments that resulted in the documents stored on those systems.

The next challenge was to identify and characterize the data. We determined that the data would be held on a standalone SECRET//NOFORN system due to its sensitivity when aggregated. Running in a system-high mode on a standalone, classified, power-gaming-type computer allowed us to exercise IDATA in a new environment. The initial version of IDATA was able to characterize documents by source, classification level, and handling requirements. It also automatically generated keywords based on an algorithmic assessment of each document's content. As such, we were able to quickly triage document content and make decisions regarding further processing. In some cases, we ran IDATA against a document multiple times to extract additional keywords or obtain keyword counts to measure importance based on the number of individual keywords, keyword frequency, and trends regarding keyword location within the document.

A second iteration of the IDATA tool set was prepared for this task. This improved capability was equipped to aggregate and count keywords to help generate statistics illustrating the use of certain keywords within a document or set of documents. We could, for example, run IDATA against a set of documents from an information system (e.g., a Navy-housed database) to find

metadata and other content providing insight into test and evaluation processes, successes, and challenges that might have otherwise become apparent only by reading several hundred individual documents.

This ability to look at the frequency and importance of keywords is roughly analogous to the early use of machine translation of Russian and Chinese scientific reports. The data are indicative of content and help subject matter experts select specific documents or data sets for further analysis.

The improved capability also delved deeper into the raw information and extracted text information embedded in analog data sets. For example, IDATA helped identify important embedded scientific and technical data by deciphering text descriptions in graphics; the titles of figures recorded as JPGs, PNGs, or TIFF images; and descriptions accompanying analog audio recordings.

Task Results

IDA researchers demonstrated that IDATA can quickly process and index thousands of documents and discover important keywords without extensive human intervention and laborious document review. IDATA provides a variety of options for collecting a wide array of information and making it available to DoD and other counter-IED organizations if JIDO (Joint Improvised-Threat Defeat Organization, the successor organization to JIEDDO) chooses to build its own information storage, retrieval, and dissemination capability.

Alternatively, if JIDO chooses to rely on DTIC as the repository of all JIEDDO information, including digital and analog test and evaluation data, IDATA's demonstrated ability to automatically generate keywords could be used to facilitate the completion of DTIC's documentation requirements.

References

Department of Defense. 2006. "Joint Improvised Explosive Device Defeat Organization (JIEDDO)." *DoD Directive 2000.19E*. Washington, DC: Department of Defense, February 14, 2006.

Executive Office of the President. February 26, 2013. "Countering Improvised Explosive Devices." Washington, DC: Executive Office of the President.