DHS Science and Technology Directorate
Data Analytics Engine – Transforming Data to Decisions

Technology Engines Context

The Department of Homeland Security (DHS) Science and Technology Directorate (S&T) launched a series of high-profile, high-impact Apex programs to look strategically at the nation's security and address future challenges while supporting today's operational needs. S&T Engines were created to meet cross-cutting needs for all Apex programs.

Impact and Vision

The Data Analytics Engine (DA-E) applies leading-edge computational data analytics research and development (R&D) techniques to enable user-focused, data-driven solutions for DHS missions. DA-E is the cross-cutting S&T resource for subject matter expertise and technical capabilities in data storage, security, computation, analysis, and visualization to the Homeland Security Enterprise (HSE), which includes DHS and its public and private sector partners.

Figure 1: DA-E received a Data Impact Award from a panel of Silicon Valley CEOs for demonstrating advanced data analytic applications on a national scale to better understand fire incidents.

Description & Approach

State-of-the-Art Data Analytics Laboratory

DA-E maintains and operates a state-of-the-art data analytics laboratory to support mission-relevant evaluations of emerging technologies, rapid experimentation, and strategic R&D efforts. The lab infrastructure complies with privacy authorities and maintains security accreditation, allowing DA-E to use operational data to help DHS identify enterprise data analytics architectures and solutions. With over two Petabytes of storage capacity, more than 100 servers

and multi-tenant users, the DA-E lab serves as a centralized resource where end-users, technology experts, and other stakeholders can develop and test hypotheses and methods.

The DA-E lab also identifies and supports transitioning computational data analytics technologies to operational use in the HSE. The lab ensures solutions minimally impact end users' existing architectures and work flows, with consideration to costs and scalability. Developing and integrating statistical analysis and knowledge products allows mission outcomes and efficiencies to be measured or traced.
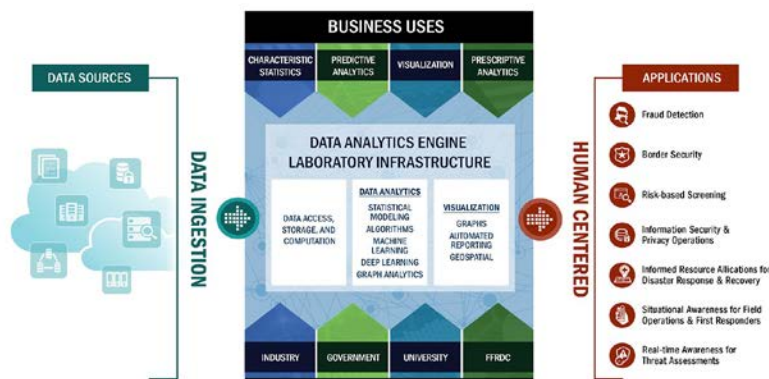
Key Activities

Assessments, Experimentation, and Strategy

DA-E enables program success by conducting:

(1) Emerging Technology Evaluations: DA-E evaluates potential game-changing tools or technologies to determine added-value and the capability to meet DHS mission needs. Evaluations involve requirements identification, market analyses, industry surveys, and hands-on testing and analysis/evaluation of mission-relevant use cases in the DA-E lab.

(2) Rapid Experimentation: DA-E conducts rapid experimentation to review data analytic technologies; provide preliminary results of data analysis on DHS operational data; and prototype solutions. The DA-E lab performs statistical data analysis on ingested data; documents use cases and requirements; evaluates analytic tools and methods against real data; and investigates data enrichment opportunities.

(3) Strategic R&D: DA-E develops and maintains engagements with national, industry, and academic labs. These strategic partnerships enrich DA-E's R&D with cutting edge research in areas such as al-

To learn more about the S&T Data Analytics Engine, contact SandTBigData@hq.dhs.gov

gorithm evaluation for enhanced risk-based screening, improved fraud detection, information sharing, secure cloud computing, automated reporting, and autonomous systems.

The DA-E mission extends beyond data analysis to examining human interfaces between the end-users and the technology and helping document and develop business cases for data-directed investments. DA-E also cultivates and maintains relationships within S&T, the HSE, and partners to ensure successful transition of the best, cost-effective solutions.

Key Successes

The following are examples of key successes and knowledge products resulting from DA-E activities:

➢ Transitioned big data architecture and tools to Immigra- tion and Customs Enforcement Homeland Security In- vestigations that has led to arrests, seizures, new leads, and situational awareness reports increasing counter- proliferation mission effectiveness for ICE offices world-wide. This effort supported United Nations Se- curity Council Resolution 1540, which legally binds member states to have and enforce appropriate and ef- fective measures against the proliferation of nuclear, chemical, and biological weapons.

➢ Transitioned private cloud digital info sharing capability to ICE/HSI/Cyber Crimes Center to improve response time to disrupt child exploitation, decreasing child endangerment.

➢ Transitioned text analytics tools to S&T Chemical Security Analysis Center to search Chemical and Biological documents to enable Chemical Security Analysis Center to keep up-to-date on emerging threats and capabilities.

➢ Transitioned Tessara, an open source data analysis tool, to Domestic Nuclear Detection Office for analyzing Rad/Nuc data

➢ Transitioned requirements gathering capability for Border Situational Apex (BSA) that has 150+ users.

➢ Received a Data Impact Award from a panel of high tech CEOs from Silicon Valley for Independent   data

audit of the U.S. Fire Administration's (USFA) National Fire Incident Reporting System (NFIRS) that collects data from fire departments across the nation. The results of the audit now enable USFA to better understand incidents and equipment failures to protect firefighters. DHS joined Visa, Coupons.com, and others to receive this award recognizing the impacts in their respective organizations, businesses, and/or society at large.

➢ BSA APEX: Piloted capability for real-time interagency collaboration and situational awareness for DHS end users from various agencies across the DC area, Detroit, and Texas during the Presidential inauguration.

➢ Led the DHS Social Media Task Force in the completion of four social media pilots that led to improving methodology and commercial tools for DHS use of open source and social media and developed technical metrics and measures.

➢ Completed and now transitioning the ICE Export Enforcement Coordination Center (E2C2) case deconfliction prototype that worked across 33 agencies and eight Departments.

Knowledge Product Examples:

➢ Entity Resolution Report

➢ Geo-locator Evaluation Report

➢ Social Media Analytics Capability Testing Report, Measures of Effectiveness and Measures of Performance Studies

➢ Apache Spark, SAP-HANA (High-Performance Analytic Appliance), and Teradata Aster Comparison

➢ Cloud Management Tools comparison

➢ Big Data Query Evaluation

➢ USCG Analysis & Production Mission Analysis

➢ Tenable Nessus Audit Script for Red Hat system compliance with DHS Sensitive System Configuration Guidance (SSCG)

➢ Conference Papers (e.g., Autonomy, Robotics, Real-time Analytics)