Big Data and Analytics for National Security

Sean Fahey, JHU/APL
sean.fahey@jhuapl.edu
Questions for the Talk

• What big data challenges are being faced in the national security environment?
• How are federal government agencies using big data approaches for national security challenges?
• What is unique/different about big data in the national security setting?
• What solutions are currently being sought by the government to address national security challenges with big data?
US National Security Strategy

• Our Interests
  – Security
  – Prosperity
  – Values
  – International Order

• Our Approach
  – Defense
  – Diplomacy
  – Economic
  – Development
  – Homeland Security
  – Intelligence
  – Communications
  – People
Major National Security Missions

- Conventional Military Defense
- Counter Nuclear Proliferation
- Counter Chemical/Biological WMD
- Counter Terrorism
- Cybersecurity
- Counter Intelligence
- Counter Narcotics
(Some)National Security Organizations

- Department of Defense
- Defense Intelligence Agency
- Director of National Intelligence
- Homeland Security
- Federal Bureau of Investigation
- NORTHCOM
Defense Positions and Tasks

<table>
<thead>
<tr>
<th>Intelligence (J-2)</th>
<th>Planning (J-5)</th>
<th>Operations (J-3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requirements</td>
<td>Mission Analysis</td>
<td>Shape</td>
</tr>
<tr>
<td>Planning</td>
<td>Course of Action (COA) Development</td>
<td>Deter</td>
</tr>
<tr>
<td>Collection</td>
<td>COA Analysis and Wargaming</td>
<td>Seize Initiative</td>
</tr>
<tr>
<td>Processing</td>
<td>COA Comparison</td>
<td>Dominate</td>
</tr>
<tr>
<td>Exploitation</td>
<td>COA Approval</td>
<td>Stabilize</td>
</tr>
<tr>
<td>Dissemination</td>
<td>Plan Development</td>
<td>Enable Civil Authority</td>
</tr>
<tr>
<td>Archiving and retrieval</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Intelligence Cycle

• The process of developing unrefined data into polished intelligence for the use of policymakers. The intelligence cycle consists of six steps, described below.
Source: DoD JP 2-0
# Intelligence Disciplines

<table>
<thead>
<tr>
<th>SIGINT</th>
<th>IMINT</th>
<th>MASINT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signals intelligence</td>
<td>Image Intelligence</td>
<td>Measurement and Signature Intelligence</td>
</tr>
<tr>
<td>OSINT</td>
<td>HUMINT</td>
<td>GEOINT</td>
</tr>
<tr>
<td>Open Source Intelligence</td>
<td>Human Intelligence</td>
<td>Geospatial Intelligence</td>
</tr>
</tbody>
</table>

Source: [http://www.dni.gov](http://www.dni.gov)
Aerial Reconnaissance: ~2012

- infrared sensor
- color/monochrome daylight TV camera,
- image-intensified TV camera,
- laser designator
- laser illuminator

The full motion video from each of the imaging sensors can be viewed as separate video streams or fused together.

Aerial reconnaissance teams are in direct communication with supported units.
Aerial Reconnaissance: ~2015

Collection → Processing and Exploitation → Analysis and Production → Dissemination

**GORGON STARE**

Current Reaper and Predator sensors offer a single field of view. New wide-area airborne sensors for the Reaper will record video from 12 cameras simultaneously.

- **12 cameras take concurrent images**
- **4 kilometer radius**

Source: Terrain image from Google Earth Pro

John Bretschneider/Staff
Data-to-Decisions Systems
Issues - Personnel

Predator Sensor
Increasing Resolution and Coverage

Analysts
Number of Highly Skilled and Trained Analysts Remains Constant or Decreases

Analysts

National security decision systems span all QDR missions with a focus on finding threats in a specified data volume with limited manpower within a specified time window

Distribution Statement A: Approved for public release; distribution is unlimited.
National Security Big Data Challenges

- Maritime Security
- Cyber Security
- Money Laundering
- Multi-INT Analysis
- Space Situational Awareness
CIA CTO Perspective on Big Data

Next

Beyond Big Data: Riding the Technology Wave

Ira A. (Gus) Hunt
Chief Technology Officer

US Government Big Data Strategy

- White House has acknowledged the importance of Big Data for multiple disciplines.
Big Data Initiative for Research and Development

• Initiative Goals
  – Advance state-of-the-art core technologies needed to collect, store, preserve, manage, analyze, and share huge quantities of data.
  – Harness these technologies to accelerate the pace of discovery in science and engineering, strengthen our national security, and transform teaching and learning; and
  – Expand the workforce needed to develop and use Big Data technologies.
• Data to Decisions Initiative
  – Challenge – The proliferation of sensors and large data sets are overwhelming analysts, as they lack the tools to efficiently process, store, analyze, and retrieve vast amounts of data.
  – Goal - Develop an open-source architecture system that enables rapid integration of existing and future data exploitation tools to achieve a new paradigm in the management and analysis of data.

Source: http://www.acq.osd.mil/chieftechnologist/areas/dtd.html
D2D Technology Assessment

- Moderately Mature
  - Driven by IT Industry

- Immature
  - Driven by Defense

- Moderately Mature
  - Driven by IT Industry

Current assessment is that unstructured data analytics is the most challenging and critical component of D2D

ASD D2D program intends to provide representative data of various types that have associated ground truth to support development and evaluation of algorithms and systems in a SOA to be made available.

Distribution Statement A: Approved for public release; distribution is unlimited.
• **Problem Statement:** Existing automation tools do not aid users in finding today’s complex and adaptable threats within mission timelines

• **3-to-5 year timeframe objective**
  - Robust classification to accurately detect, geo-register, classify, and identify surface objects despite difficult environments, configurations and emplacements
  - Robust automation tools to identify relationships, patterns of life and activities of objects on the ground
  - Robust tools to capture, store and retrieve HUMINT-based information to identify and leverage popular support against insurgents
  - Domain-specific tools to capture, search, mine and exploit explicit information on insurgent networks from unstructured textual data sources

• **7-to-10 year timeframe objective**
  - Robust automation tools to identify relationships, patterns of life and activities of dismounts
  - Robust tools to search, mine and exploit open-source data to identify all aspects of insurgent networks
Background

“DoD is swimming in sensors and drowning in data”

Data

- Imperfect
- Distributed
- For diverse missions

XDATA - technology to leverage big-data at all stages from analysis to operations

Approved for Public Release, Distribution Unlimited (DISTAR Case 19136)
Develop computational techniques and software tools for analyzing large volumes of data,
- semi-structured (e.g., tabular, relational, categorical, meta-data), and
- unstructured (e.g., text documents, message traffic).

Challenges
a) developing scalable algorithms for processing imperfect data in distributed data stores, and
b) creating effective human-computer interaction tools for facilitating rapidly customizable visual reasoning for diverse missions.

Open source software toolkits that enable flexible software development supporting users processing large volumes of data in timelines commensurate with mission workflows of targeted defense applications

Approved for Public Release, Distribution Unlimited (DISTAR Case 19136)
Challenges for National Security Big Data

• Government Budget Process
• Government Acquisition Process
• Siloed and Proprietary Systems
• Data Balkanization and Compartmentalization
• Government Privacy Obligations and Data Mining Restrictions
• Personnel System Lag
• Austere Deployment Environments

CAUTION
TRIPPING HAZARD
Government Budget Process

Federal Government agencies have to commit to specific spending programs roughly 18 months before they begin spending.

- **FY2012**
  - Oct: Agency develops investment plans
  - Jul: Agency sends budget request to OMB
  - Apr: OMB revises budget with agency

- **FY2013**
  - Jul: President submits budget to Congress
  - Apr: Congress debates budget and passes appropriations bills

- **FY2014**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2015**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2016**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2017**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2018**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2019**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2020**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2021**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2022**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2023**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2024**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2025**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2026**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2027**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2028**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2029**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend

- **FY2030**
  - Jul: Agencies have funds available to spend
  - Apr: President signs budget
  - Jan: Agencies have funds available to spend
  - Oct: Congress debates budget and passes appropriations bills
  - Jul: President signs budget
  - Apr: Congress debates budget and passes appropriations bills
  - Jan: Agencies have funds available to spend
Siloed and Proprietary Systems

- Government IT systems have historically been procured, developed, accredited and maintained as stand alone systems.
Data Balkanization and Compartmentalization

- Top Secret
- Secret
- SCI
- Coalition
- Unclassified
- US Persons
Government Privacy Obligations and Data Mining Restrictions

- US Constitution Protections
- Privacy Act of 1974
- Privacy Impact Assessments
- US Persons Protections
Recruiting and training systems were designed for...not
Austere Deployment Environments

- Limited Power, Space and Cooling
- Limited or intermittent bandwidth to end users