Creating End to End Homeland Security Information Management Environments

Lisa Sokol, Ph.D.
Lisa.Sokol@gd-ais.com

General Dynamics
Knowledge Management
Center of Excellence
Transform Unstructured Data- Text, Audio, OCR into a Structured, Analyzable Form

• Automatically derive relationships, time, events, names, locations
• Disambiguate facts
• Exploit “words” in both storage and search

Profiles

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Location</th>
<th>Related Organization</th>
<th>Event 1</th>
<th>Event 2</th>
<th>Event 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bin Laden</td>
<td>Person</td>
<td>Afghanistan</td>
<td>Taliban</td>
<td>Turn Over</td>
<td>Indictment</td>
<td>Denial</td>
</tr>
</tbody>
</table>

Intermediate Events

- **Turn Over**
  - **Who:** Taliban
  - **Key Verb:** to turn over
  - **Whom/What:** Bin Laden
  - **Others Involved:** authorities

- **Indictment**
  - **Who:** Bin Laden
  - **Key Verb:** was indicted
  - **Modifiers:** in the embassy bombings

- **Denial**
  - **Who:** Bin Laden
  - **Key Verb:** denied
  - **Whom/What:** any involvement in
Design, Implement and Maintain Terabyte Sized Information Environments

- Mixture of distributed and centralized data environments
- Scalable – Fast computation time over terabytes of data
- Mission oriented architectures
- Service oriented architectures

Tens of Terabytes of Data

<table>
<thead>
<tr>
<th>Year</th>
<th>Data Warehouse Size</th>
<th>Median Response Time (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>1998</td>
<td>6</td>
<td>20</td>
</tr>
<tr>
<td>1999</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>2000</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>2001</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>2002</td>
<td>31</td>
<td>1</td>
</tr>
<tr>
<td>2003</td>
<td>61</td>
<td>2</td>
</tr>
<tr>
<td>2004</td>
<td>75</td>
<td>2</td>
</tr>
</tbody>
</table>
Analyst Workbench – Discover Actionable Data, Facilitate Detection, Prevention and Interdiction

- Reason over heterogeneous, multi-source multi-media
- Discover new relevant, timely actionable information, e.g., cues
- Create mission relevant workbenches that help analysts
  - Manipulate large amounts of data
  - Discover new relationships
  - Detect patterns embedded within data, both time and space
- Fusion and Reasoning

Links Between Islamic Banks and Charities
Design Environments to Facilitate Collaboration and Data Sharing

- Implement secure, multi-channel collaboration tools & technologies to accelerate decision-making & innovation
  - Components can include: chat, instant messaging, document management and email, and a collaboration process
- Requirements can make collaboration environments challenging to create
  - Community requirements and attributes
  - Bandwidth availability
  - Asynchronous -- Synchronous
  - Dynamic -- Static
  - Small Group -- Large Scale
  - Security
- Multiple Approaches: Center Based, Peer-to-peer, and Hybrid
Multi-Domain/Cross-Domain Overview

**MSL**
- Expensive hardware—multiple desktops, servers, networks
- Expensive admin/maintenance multiple sets of services
- Reduced software costs-COTS
- Flexible COTS OS
- Limited multi-domain usability

**MILS**
- Reduced hardware—single desktop. Network reduced by use of encryption devices
- Reduced admin/maintenance
- Reduced software costs-COTS
- Flexible COTS OS
- Still limits multi-domain usability and information sharing: Only does partial job

**MLS**
- Lowest hardware cost—single set of hardware
- Lowest admin/maintenance cost
- Higher software integration cost
- Limited OS
- Greatest multi-domain usability
- True Information Sharing

*GENERAL DYNAMICS*
Advanced Information Systems

GDAIS Proprietary
SCIF in a Box

- Novel hardware-based technology that enforces continued originator control over data at rest, in transit, and in use.
  - Protected content through data encryption
  - Flexible access rights granted with electronic “tickets”
  - Access control enforced by software residing between Microsoft Windows and the computer
  - Trusted hardware and tamper detection protect the computer from physical attacks
- Hybrid of trusted hardware and small, accreditable software enables:
  - Highly secure platform
  - Loss prevention
  - Flexible and extensible protection of content
  - Transparent to users
  - Compatible with any application, content type, or storage/transport medium
  - Backwards compatible with legacy content and infrastructure