CLOUD ANALYTICS:
Empowering the Army Intelligence Core Analytic Enterprise

5 APR 2011
Advanced Analytics
Harnessing Data for the Warfighter

2005...

2006 - Now...

Global Enterprise
"Common Architecture"

Forward Operating Base

Brigade Combat Team

I2E

DCGS

GIG

LandWarNet

Data Silos

Data Access

Data Deluge

Documents
Consolidated Database

Applications

Documents

Entity and Event Services

Advanced Analytics

creates Resolved Entity Database

Entities

Relationships

Events

DB

Entity/Event Centric

• Focuses on Persons, Places, Activities, etc.)
• Resolves and Disambiguates Entities
• Extracts Entity-To-Entity Relationship
• Provides Entity & Event Representation

Documents
Precision Search
Finding Relevant Data Faster

Internet Pages are Ranked by Popularity (Clicks)

"Google" Type Search

Focus of Cloud-enabled Analytics

Precision Search

Present Systems

Enabled By Cloud

Complex Boolean Searches

Recall

Precision
Working with ALL the Data

Lexical Searches, Boolean Search Attributes, or Properties Based Filters

Today’s Analytics Must Filter Data

Product from Filtered Data

Tools / Analysis

With the Cloud, We Do Not Have to Filter to Get a Working Set

Cloud Analytics

Analysis Over the Entire Corpus – Follow the Analytic Threads without Re-Loading Data
Solving the Precision / Recall Conundrum with Semantic Enrichment of the Data

**Concepts/Summarization** (e.g. Terrorist Cell Leaders)

**Resolved Entity** (John with ID xxxxx)

**Entity** (Person, Object, Organization, Location)

**Lemma/Element/Part of Speech** (Noun, Pronoun, Punctuation…)

**Token** (Aggressively Indexed Words)

**De-Anonymization of Large Data Sets**

**Detect / Match Behaviors and Patterns**

**Massive Data Sets for Anomaly / Change Detection**

** Massive Data Aggregation for Machine Analytics, Baselining, and Trend Analysis**

Enabled by fine grain security and compliance enforcement

**Determining that Two Patterns of Life are the Same but Not Necessarily Whose Pattern of Life**

**Indications and Warnings**

**Non-Attributable Aggregate Behavior**

- Determine Avg Traffic Speed by Tracking Cell Movement
- Determine the Sentiment of a Town, City, Region, Country
Simple Search, Graph & Document Viewers
(Entity Search, Entity Resolution, Social Analysis)

- All Searches Are Against *All the Data* (Structured and Unstructured)
  - Unlike Current Systems, *Don’t Need to Filter Data*
- Search Results Returned at Internet Speed
- “Chinese Menu” Provides Rapid Discovery of Unknown Connections
- Immediately Enter Analysis (Seconds vs Minutes/Hours)
- Click to Explore Connections for Relevance *(Snippets)*
- De-clutters Visual Display, Without Filtering
- Adding Context Leads to “Precision Search”

- Associative Memory Index
Simple Functions Combine for Powerful Analytic Capability

Decouples Functionality for Widget Reuse *(Pull Together like LEGOs)*

Ease of Widget Creation

- Lightweight Software Coding (< 100 Lines of Java Script)
- Open Commercial Standards
- Users Can Build Rapidly (Secure Container)

Rapid Response to User Needs and User Feedback

Harnesses the Innovation of our Users

Widget Interaction: Patterns of Life

(Widgets: Time Wheel, Histogram, Map)
Contextual Search / N-Gram
(Enhanced Searching)

- Aggressively Index Every Word
- Data Self-Organizes for Analysts to See Context of Terms
- Internet Speed of Response
- Analysts Form Relevant Queries in Seconds from a Cold Start
- Enables Precision Search
- Foundational Index - Building Block for other Widgets + Micro-Analytics
  - MSR Example: Map the Documents
  - Name Variant Example: Select Common Mis-Spellings to Check ALL
• Use Cloud Analytics Not Just on Red Threat Data, but to Manage BLUE
  • Leverage Knowledge on How the Network is Used
  • Exploit User Profiles and Expertise Across the Enterprise

• Better, Faster Analysis:
  • Link with Other Analysts Working Same Problem (Auto-find SMEs)
  • Give Recommendations Based on Users’ Expertise

• Training: Enables Supervisor Oversight and Management of Analysts
  • Detect Weak Analytic Skills (Avoid “Analytic ADD”)
  • Prevent Too Many Analysts from Working Same Problem

• Detect Anomalous Activity (Counter Insider Threat)
  • Users Accessing Data Beyond Assigned Function
  • Excessive Downloading (WikiLeaks)
CI-HUMINT / CHAMPION
(Integration of “Legacy Systems”)

• Demonstrate the Flexibility of the Cloud to Integrate Legacy “Boxes”
  • Widgetize Functionality: 1/2 Month
  • Complete Integration into Cloud: 2-1/2 Months
  • Total Effort to Integrate Legacy → Cloud: 3 Months
    • Compare with Traditional 18-24 Months to Deploy,
      and Benefit from Less HW / SW / Support Personnel / Footprint

• Decouple Proprietary Services and Reuse Existing Services
• Focus Development Resources on New Capabilities
Cloud-to-Cloud
(Forming the Ubiquitous Cloud)

• Demonstrate Cross-Cloud Functionality
• Allows Data Stewards to Manage Data
  • Others Can *Leverage, not REPLICATE*, Vast Amounts of Data
• Access Others’ Data, Services and Hardware at No Added Cost
• Reduced Bandwidth Required (Only Pass Results, not Underlying Data)
• *Empowered by an IC-common, Non-Proprietary, Open Architecture*
Cloud-enabled Advanced Analytics

- Exponential Improvement
  - Reduces Time for Analysis (Hours → Minutes)
  - Enables Fast Precision Search of Huge Data
  - Every Query Searches ALL Data; No Data Filtering Needed

- Power of IC-Common Cloud HW, SW, Visualization
  - Shared Investments Accelerate Progress
  - Common Standards Allow Widgets to Communicate, Even If Built Separately
  - Development & Fielding Cycle Reduced from Years to Months: Meet Needs Rapidly

Responsive, Agile Innovation to Transform Army Intel