

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

Data

(facts, observations...)

Information

(data + content)

Knowledge

(information + judgment)

Actionable Intelligence

Automated

Interactive

Automated



Requirements

- problem
- assumptions
- constraints
- tasking request
- hypothesis formulation
- social
- technology
- technical environment
- mission
- operational

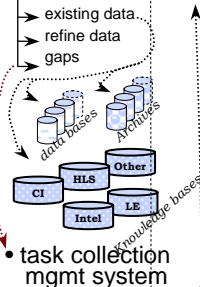
Command & Control

- direct
- manage
- allocate resources
- assign responsibility
- monitor

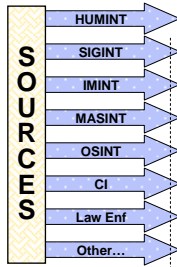
across entire spectrum

IDENTIFY data

- determine criteria
- establish requirement
- id sources
- define query, question & search terms
- query / search



COLLECT data/info



- level 1 analysis
- capture

Transform, Organize & Store Data

- ETL
- tag / index (metadata)
- pattern match
- automatic speech proc.
- machine translation
- ocr
- entity extraction: persons, places, org
- spatial look-up
- Deep extraction: events, relationships
- name – cross lingual
- person/info disambiguation
- alias
- event detection
- spatial disambiguation
- smart data: ontology, taxonomy, thesaurus
- organize
- abstract & categorize
- compile
- correlate
- classify
- summarize

Data Services

- distributed/federated
- OLAP
- OLTP
- Caching
- Replication
- Views

ANALYZE / Create Knowledge

Exploit & Understand

- hypothesize
- explore alternatives
- explain - describe
- detect - discover
- evaluate - assess
- estimate - predict
- construct model
- test
- request new data

Level 2 and 3

Using

- search multi media (semantic, key word, pattern,...)
- browse
- mining (prediction, decision trees, clusters)
- agents
- hypothesis support
- modeling
- Reasoning (GIS, temporal)
- Data analysis
- Visualization
- link analysis
- time line analysis
- collaboration
- event correlation
- situation assessment
- scenario detection

PRODUCE PRODUCT

- review & approve
- Digital Production
- revise
- edit
- adjust
- correct
- update
- populate

DISSEMINATE COLLABORATE

- chat
- email
- instant messaging
- present
- publish
- share
- Asynchronous
- synchronous
- auditing

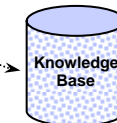
- act



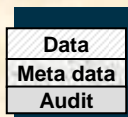
Decision Makers Users



- push
- pull



Data Layer



- Ontology

- acquire
- ingest
- cleanse
- transform
- load

Persistent Data Security Controls

Security Layer

User Identity

Infrastructure Security Controls

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

Relationships	
From: U.S. Relation: Provocation To: Muslims Type: Political Affect: Negative	From: U.S. Relation: Pressure To: Taliban Type: Political Affect: Negative
From: Taliban Relation: Protection To: Bin Laden Type: Political Affect: Positive	

Key:
 From
 Relation
 To
 Type
 Affect

Profiles	
Name: Bin Laden Type: Person Location: Afghanistan Related Organization: Taliban Event1: Turn Over Event2: Indictment Event3: Denial	Name: Muslims Type: Organization Name: Afghanistan Type: Location Subtype: Country Name: Taliban Type: Organization Location: Afghanistan Related Person: Bin Laden Event1: Turn Over
Name: Manhattan Type: Location Subtype: City Modifier: Lower	

Key:
 Person
 Organization
 Location

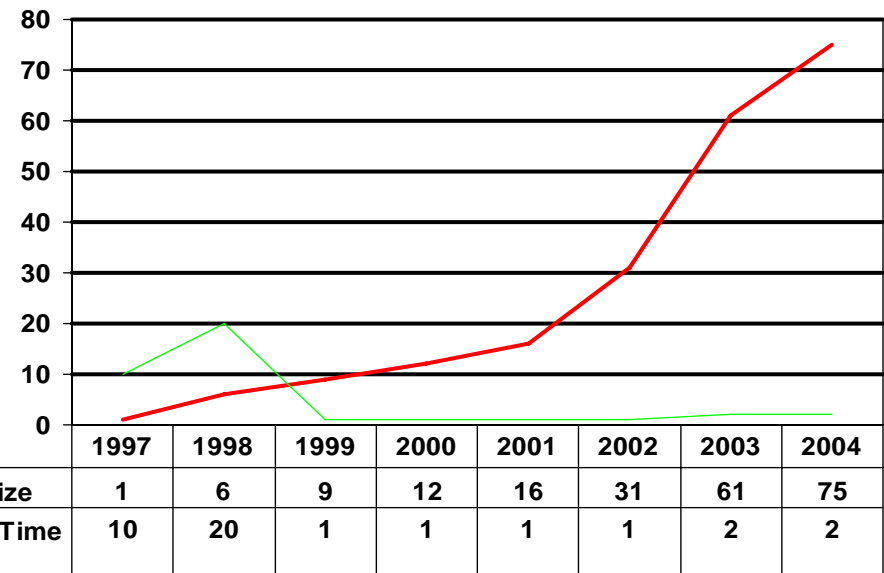
Intermediate Events
Turn Over Who: Taliban Key Verb: to turn over Whom/What: Bin Laden Others Involved: authorities
Indictment Who: Bin Laden Key Verb: was indicted Modifiers: in the embassy bombings
Denial Who: Bin Laden Key Verb: denied Whom/What: any involvement in

Key:
 Who
 Key Verb
 Whom/What
 Modifier/Others

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

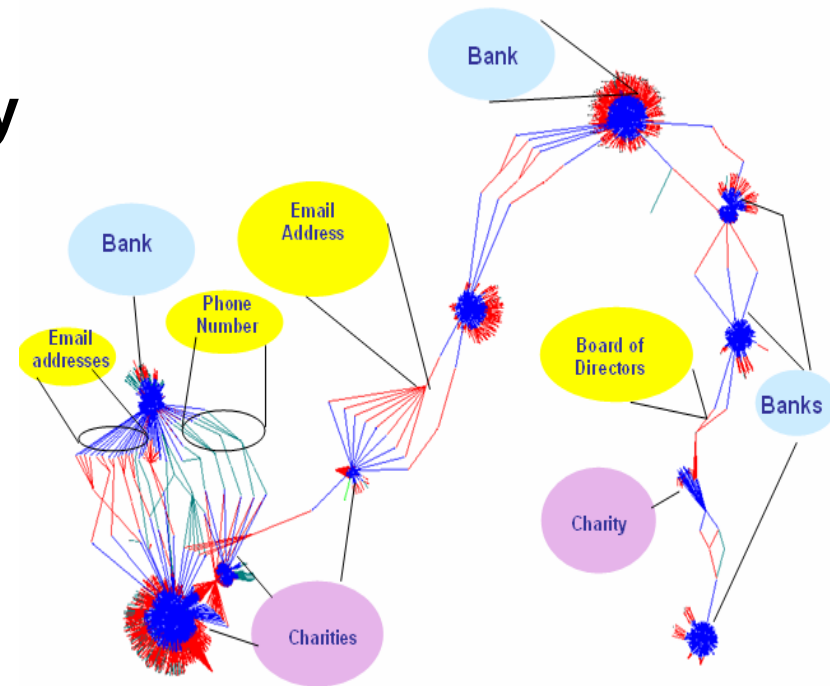
Tens of Terabytes of Data



— Data Warehouse Size	1	6	9	12	16	31	61	75
— Median Response Time (minutes)	10	20	1	1	1	1	2	2

Analyst Workbench – Discover Actionable Data, Facilitate Detection, Prevention and Interdiction

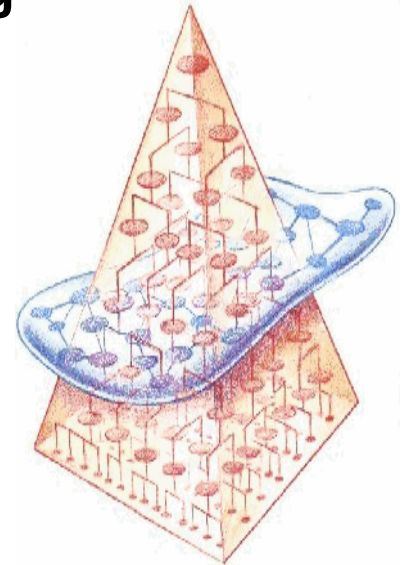
- Reason over heterogeneous, multi-source multi-media
- Discover new relevant, timely actionable information, e.g., cues
- Create mission relevant workbenches that help analysts
 - Manipulate large amounts of data
 - Discover new relationships
 - Detect patterns embedded within data, both time and space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



Links Between Islamic Banks and Charities

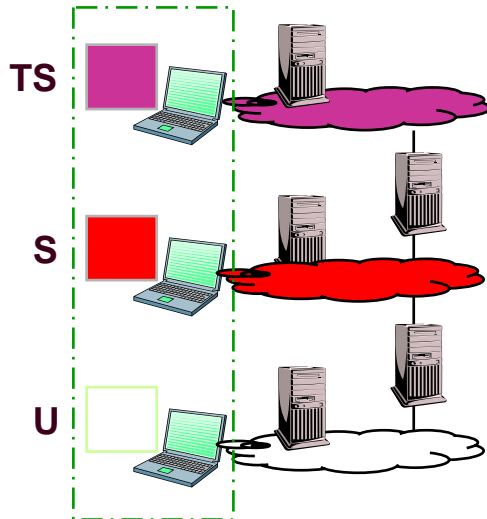
Design Environments to Facilitate Collaboration and Data Sharing

- **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**



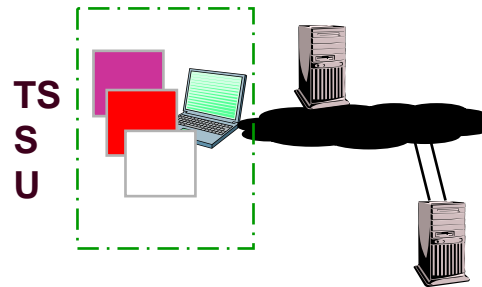
Multi-Domain/Cross-Domain Overview

MSL



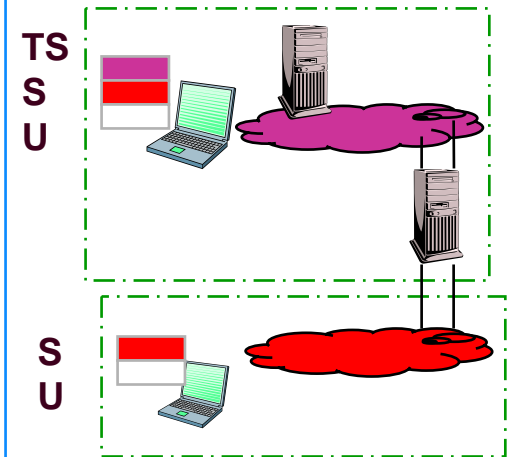
- Expensive hardware-- multiple desktops, servers, networks
- Expensive admin/maintenance multiple sets of services
- Reduced software costs-COTS
- Flexible COTS OS
- Limited multi-domain usability

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

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

