BAE SYSTEMS

Semantic Interoperability Community Of Practice (SICoP)

Semantic Interoperability Conference 2006

9-10 February, 2006

Knowledge Broker Semantic Integration Prototype

Ray Piasecki
BAE Systems Inc.
E&IS Fellow
Technical Director, CFT/TCS Knowledge Systems
(858) 592-5104
ray.piasecki@baesystems.com

Topics

BAE SYSTEMS

- Research Performed
- Semantic Integration Architecture
- Research Observations
- Prototype Demonstration

Research Outline

BAE SYSTEMS

Knowledge Broker Prototype

- A prototype evaluating the application of knowledge representation technologies to perform systems of systems integration.

1. **Created a Reference Architecture Model that mapped KR, EII and UIM technologies across enterprise platforms and information tiers.**
   - Captured core architectural constructs and key technologies.
2. **Ran several focused technology studies.**
   - Examined functional capability and breadth of several ontology platforms and analytics engines to study elements of the reference architecture.
3. **Developed a large scale prototype to simulate an information enterprise.**
   - Enterprise knowledge portal with advanced search and link analysis tools.
   - Federated Geospatial data systems.
   - Unstructured data stores.
4. **Analyzed phase-1 results and re-factored into spiral-2 prototype.**
   - Examined the practicality of knowledge representation platforms as evident in existing academic and COTS products.
   - Evaluated the robustness and stability of supporting standards.

Technical Studies

BAE SYSTEMS

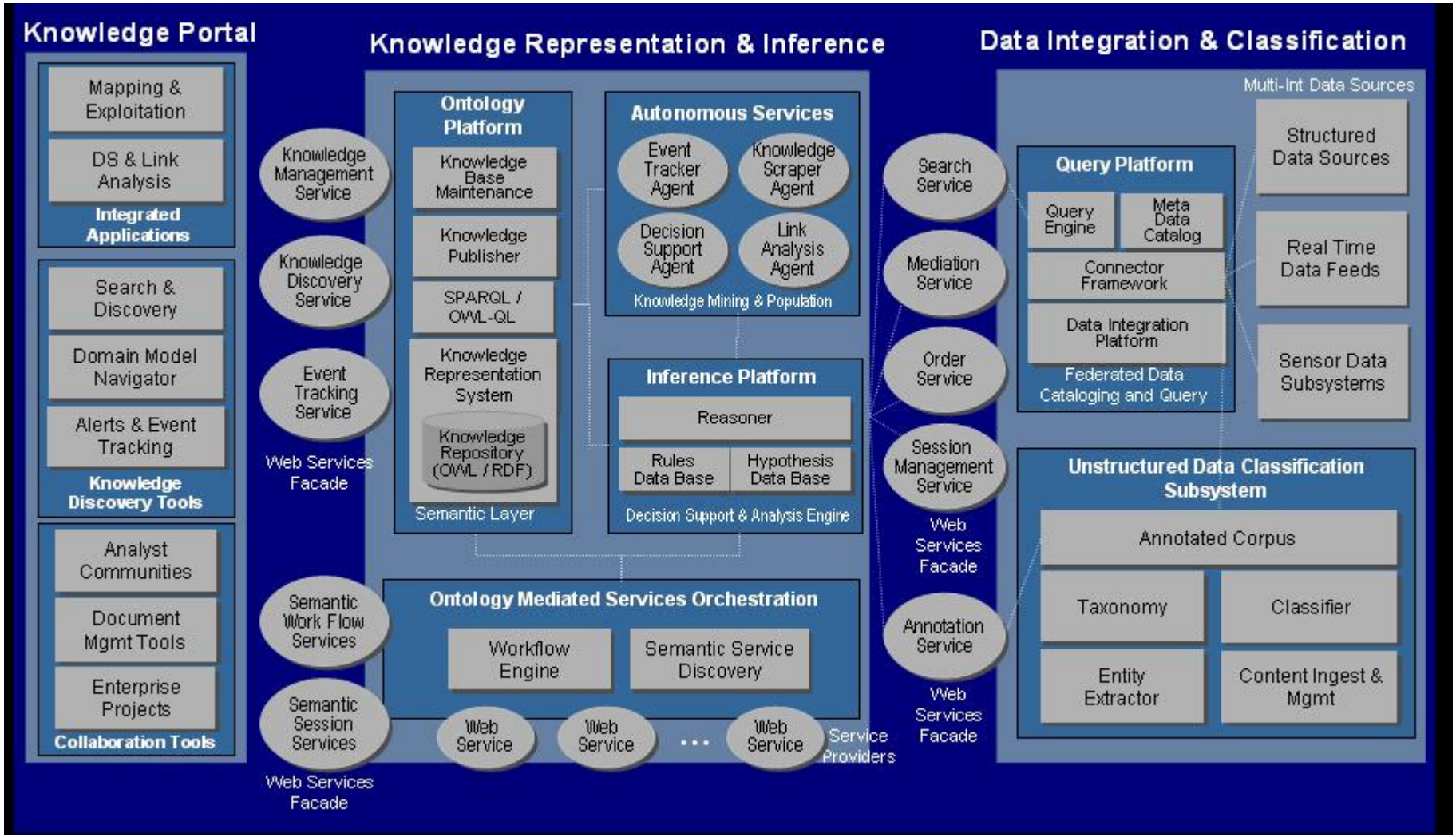
- **Platform Analysis**
  - Ontology: SnoBase, Jena2, TM4J
  - EII: MetaMatrix, LiquidData
  - Analytics: Gate, UIMA SDK

- **Integration Analysis**
  - Semantic Integration: Used the semantic layer as an integration construct. Mapping a domain vocabulary across the structured and unstructured data sets. (KR to EII connection)
  - Structured Data Systems: Federated queries against enterprise metamodel for access to data stores
  - Unstructured Data Systems: Performed an entity extraction and classification on data sets and posted annotations to populate ontological instances

- **Ontology Analysis**
  - Applied a suite of ontology's built for individual communities of interest (COI) rather then a complex and overly generic cross domain vocabulary
  - Experimented with model size to determine the practicality of reasoning against varied model complexities
  - Compared the end-user “application capability” enabled through a hierarchical taxonomy provided in a topic map against deeper semantic relationships facilitated in OWL/RDF

Domain Knowledge Broker - Architecture Reference Model

BAE SYSTEMS

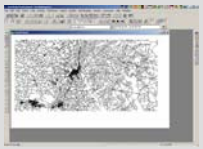


Knowledge Broker Systems Integration Prototype

Spiral-1 Components

BAE SYSTEMS

Semantic Search



Link Analysis

Information Alerts

GeoSpatial Search

Domain Portal

Mission Content Repository

Knowledge Portal

Semantic layer applied as the systems integration construct. Applications searched against the semantic model.

Semantic Search Svc

Event Tracking Svc

Query Builder

Ontology Engine

COI Ontology

COI Ontology

***

COI Ontology

Knowledge Representation Platform

Domain Knowledge Base

Post

Pub/Sub Msg Channel

Federated Access to distributed data systems mapped behind a domain focused semantic layer. Classes in the ontology are mapped to elements in the metadata catalog.

Data Access Svc

EII Platform

Query Engine

Meta Data Catalog

Spatial Mediator

Attribute Mediator

Format Mediator

Query Mediator

Schema Connector

Schema Connector

Schema Connector

Schema Connector

Image

Map

Vector

Target

Post

Federated Data Access & Integration

Documents and sites ingested. Entities were extracted from unstructured data sets and posted to populate model instances

Post

Annotation Svc

NGA Names Server

WordNet

CIA Fact Book

UIMA Platform

Search Engine

Annotated Corpus

Ortho Matcher

Gazeteer

Named Entity Extractor

POS Tagger

English Tokenizer

Sentence Splitter

Ingest Svc

Data Ingest & Entity Extraction

Technical Observations

BAE SYSTEMS

- **Platform technical issues were faced with even pragmatic prototyping**
  - Encountered memory and performance issues persisting OWL statements when trying to create a model robust enough to do some practical level of reasoning.
  - A modest owl model can be 500K statements. Our tests run with 90K OWL statements hosted on the Jena platform took hours to return a query.

- **Established a semantically rich interface to support discovery of enterprise data.**
  - A semantic layer was developed for a domain specific community of interest not as a unifying ontology for cross domain unification or as an enterprise-wide common vocabulary.
  - A universal semantic layer is not seen as practical nor semantically rich.

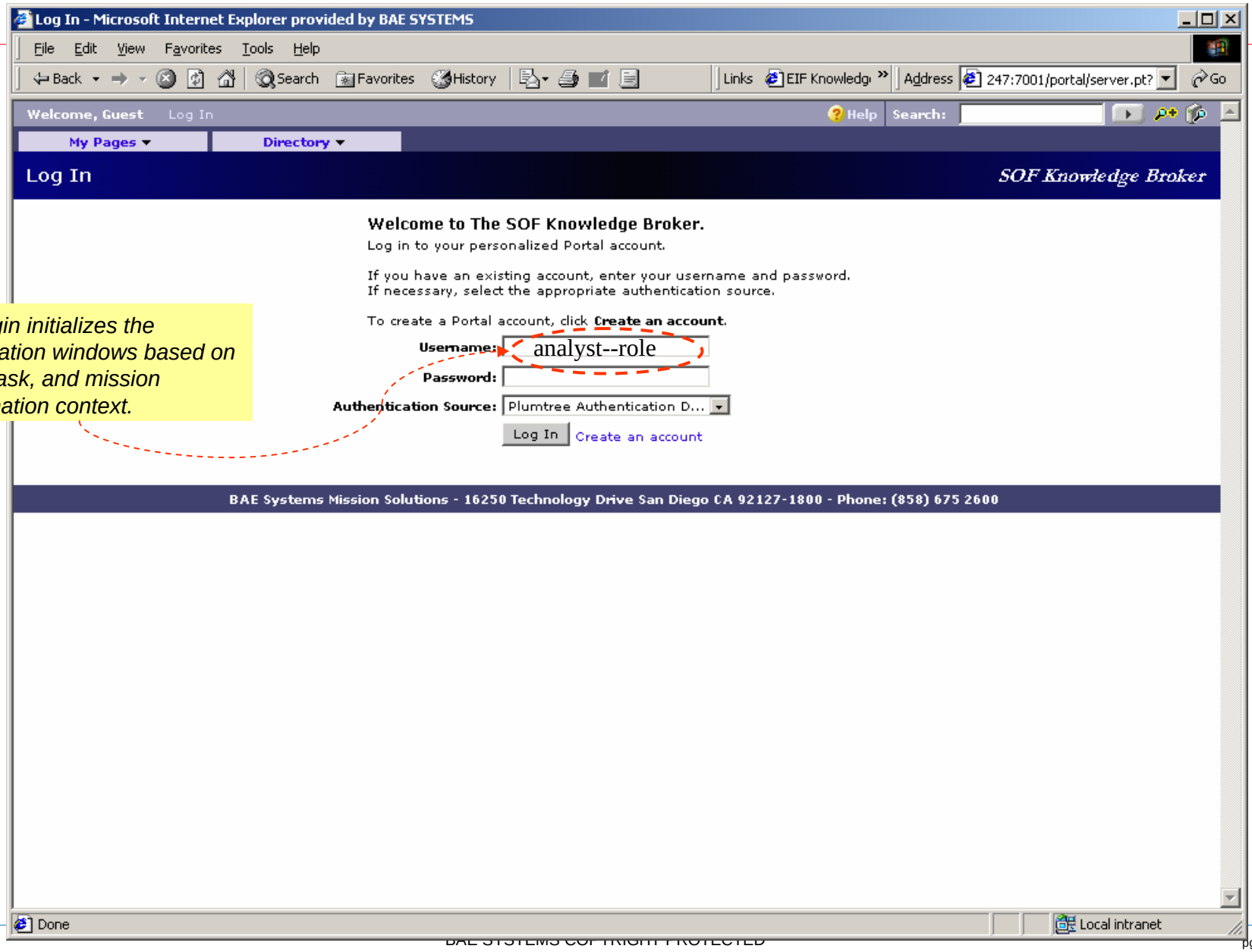
- **Model Instantiation**
  - Populating the ontological model with all instances of information across a large-scale enterprise does not appear to be feasible. The model quickly becomes impractically large.
  - Instead of implicit model population a query approach was used.
  - Classes in the ontology were mapped against entities in the enterprise metamodel. Attributes of the classes were mapped against metamodel properties. Upon request a class was populated using a query against the metamodel.
  - Complex integration code is required to interface the ontology layer and the back end data integration layer.
  - Evolving COTS products that move to incorporate both traditional EII and Semantic Markup engines in a single platform will significantly reduce integration glue code (i.e Metamatrix match-it).
  - Significant complexity is associated with creating an ontological model abstracting Temporal and Geospatial types

- **OWL & RDF tests**
  - Although OWL adds ability to specify richer semantic relationships , available platforms and query languages don't offer a corresponding semantic query capability to leverage that information (i.e. Jena with RDF query) .
  - Until OWL-QL arrives one is still operating at the level of extracting an element from a taxonomic hierarchy as opposed to querying for richer semantic constructs such as typed class relationships, cardinality, equality, etc.

Prototype Demonstration

Role Based Login

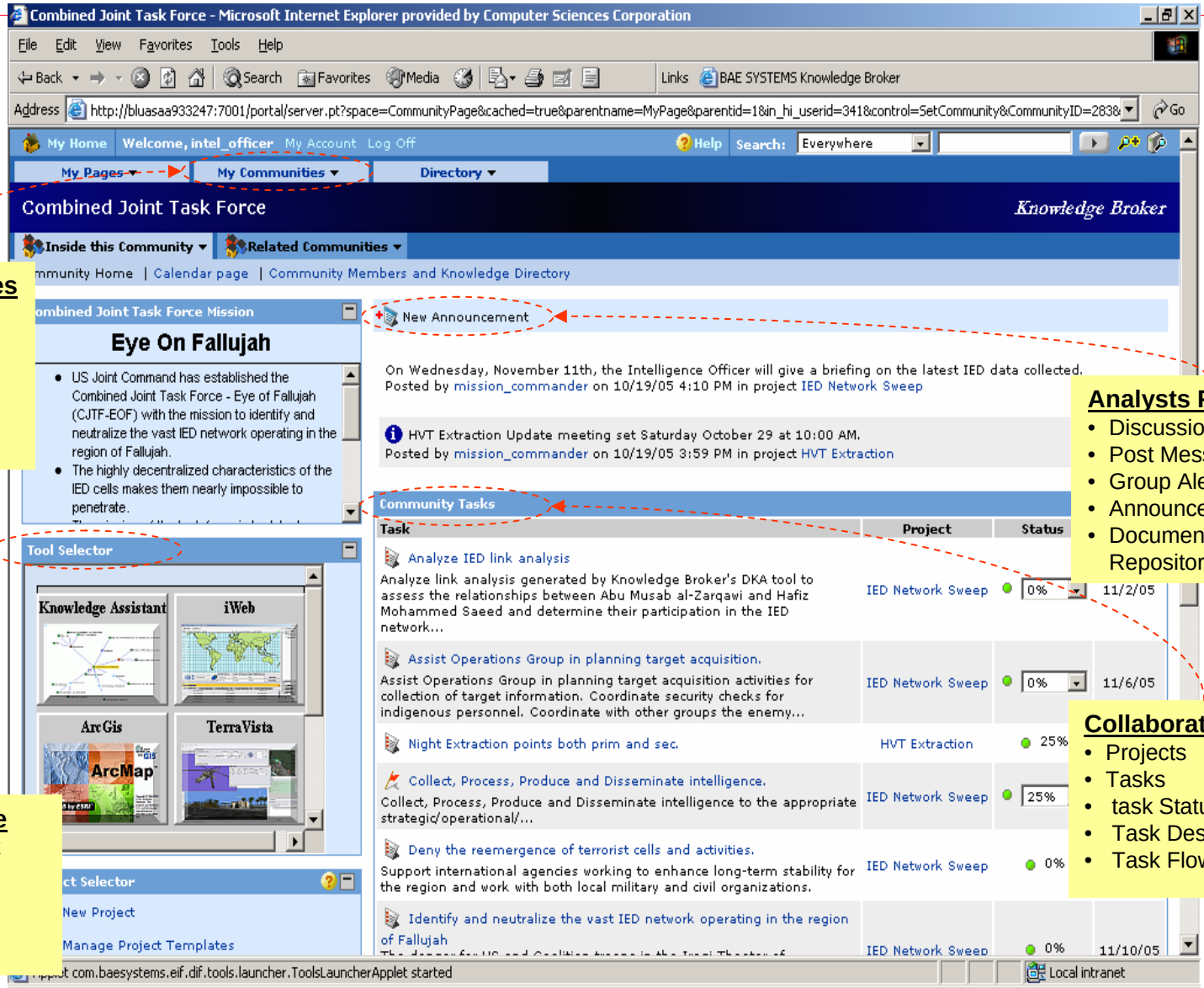
BAE SYSTEMS



User login initializes the application windows based on role, task, and mission information context.

Collaborative Knowledge Portal

BAE SYSTEMS



Analyst Communities
- Enterprise Projects
- Enterprise Calendar
- SME List
- Community Ontology
- COI Announcements
- Discussion Groups

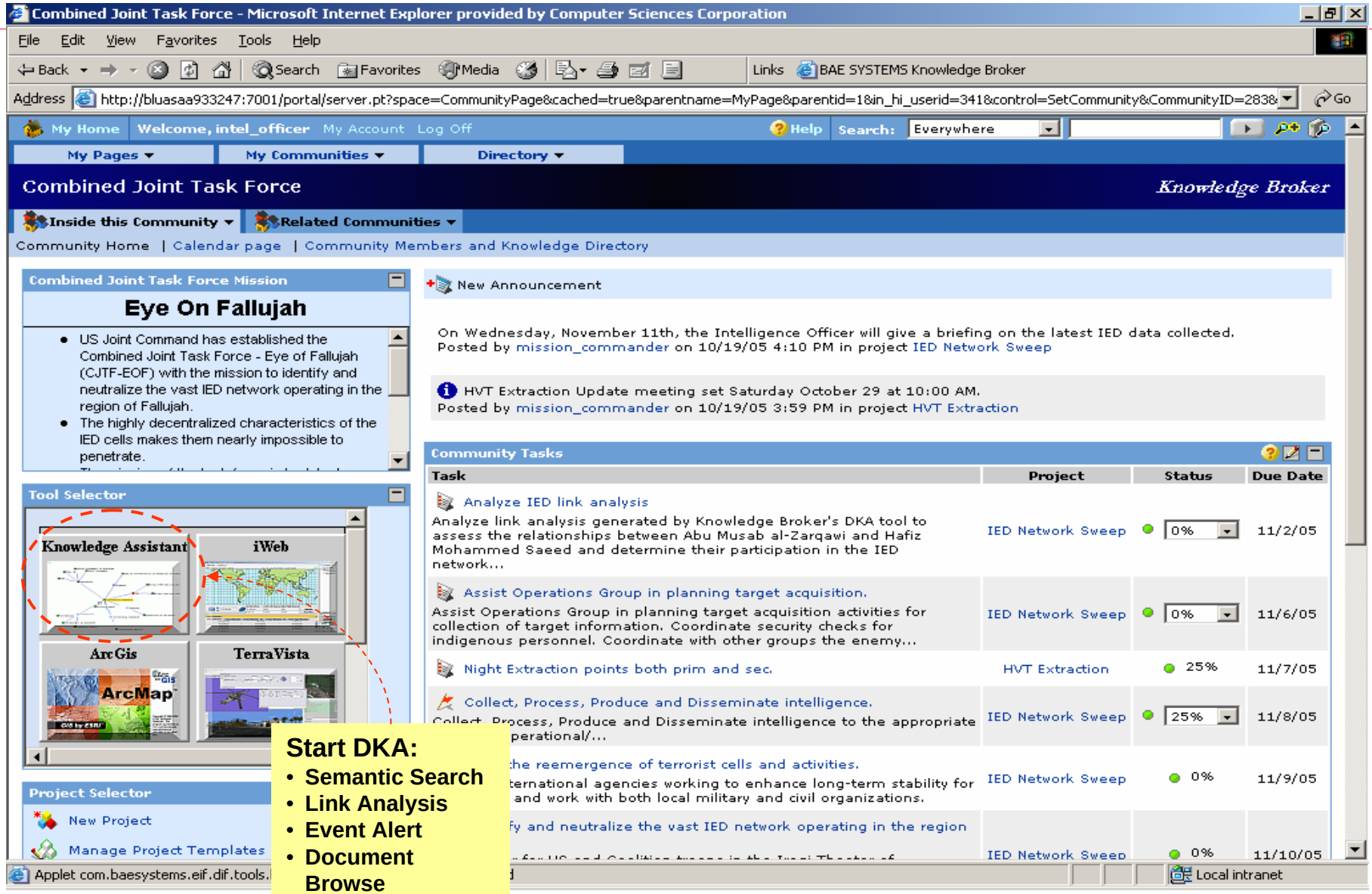
Analysts Forums
- Discussion Threads
- Post Messages
- Group Alerts
- Announcements
- Document Repository

Collaborative Tasks
- Projects
- Tasks
- task Status
- Task Description
- Task Flow

Analyst Tool Palette
- Knowledge Assistant
- Information Search
- Data Access
- Analysis
- Exploitation

Integrated Tool Palette

BAE SYSTEMS

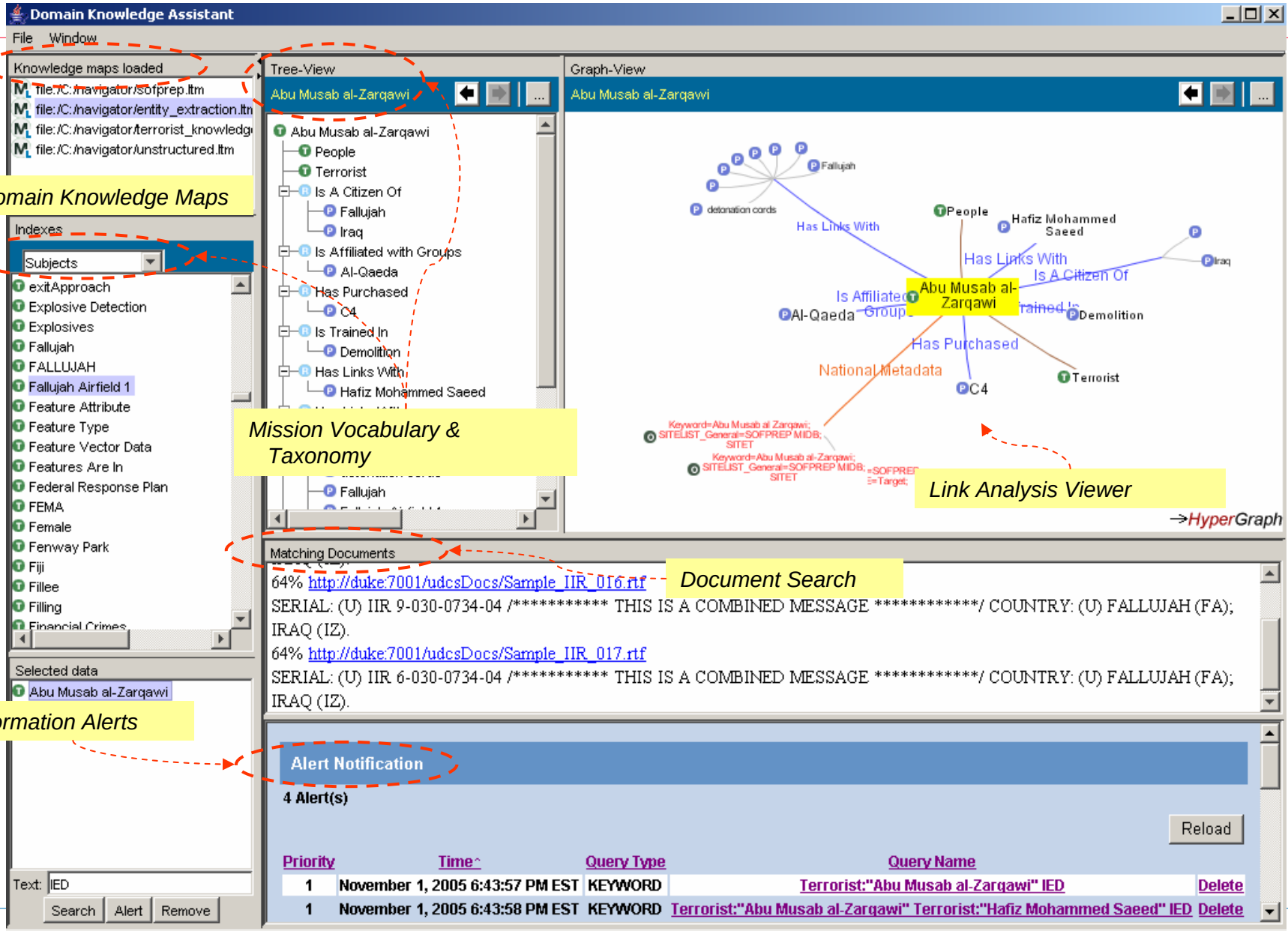


Start DKA:

- **Semantic Search**
- **Link Analysis**
- **Event Alert**
- **Document Browse**

Domain Knowledge Assistant (DKA)

BAE SYSTEMS



Combined Joint Task Force - Microsoft Internet Explorer provided by Computer Sciences Corporation

File Edit View Favorites Tools Help

Back Search Favorites Media Links BAE SYSTEMS Knowledge Broker

Address http://bluasaa933247:7001/portal/server.pt?space=CommunityPage&cached=true&parentname=MyPage&parentid=1&in_hi_userid=341&control=SetCommunity&CommunityID=283& Go

My Home | Welcome, intel_officer | My Account | Log Off | Help | Search: Everywhere

My Pages | My Communities | Directory

Combined Joint Task Force

Knowledge Broker

Inside this Community | Related Communities

Community Home | Calendar page | Community Members and Knowledge Directory

Combined Joint Task Force Mission

Eye On Fallujah

- US Joint Command has established the Combined Joint Task Force - Eye of Fallujah (CJTF-EOF) with the mission to identify and neutralize the vast IED network operating in the region of Fallujah.
- The highly decentralized characteristics of the IED cells makes them nearly impossible to penetrate.

Tool Selector

Start iWeb:
- **Reach Back/ Federated Access**
- **Geospatial data discover, browse and retrieve**

Project Selector

New Project

Manage Project Templates

New Announcement

On Wednesday, November 11th, the Intelligence Officer will give a briefing on the latest IED data collected.
Posted by mission_commander on 10/19/05 4:10 PM in project IED Network Sweep

HVT Extraction Update meeting set Saturday October 29 at 10:00 AM.
Posted by mission_commander on 10/19/05 3:59 PM in project HVT Extraction

Community Tasks

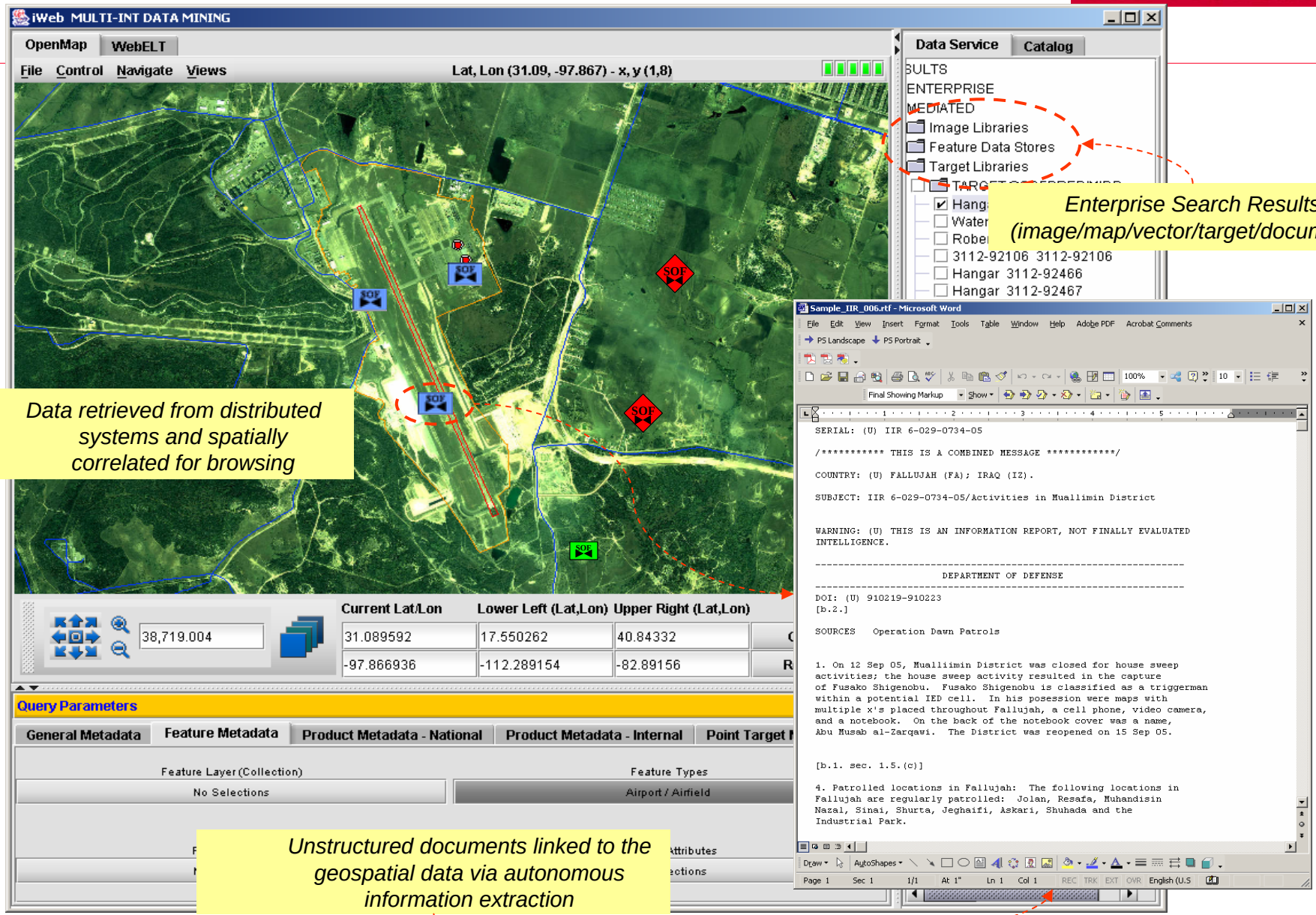
Task	Project	Status	Due Date
Analyze IED link analysis Analyze link analysis generated by Knowledge Broker's DKA tool to assess the relationships between Abu Musab al-Zarqawi and Hafiz Mohammed Saeed and determine their participation in the IED network...	IED Network Sweep	0%	11/2/05
Assist Operations Group in planning target acquisition. Assist Operations Group in planning target acquisition activities for collection of target information. Coordinate security checks for indigenous personnel. Coordinate with other groups the enemy...	IED Network Sweep	0%	11/6/05
Night Extraction points both prim and sec.	HVT Extraction	25%	11/7/05
Collect, Process, Produce and Disseminate intelligence. Collect, Process, Produce and Disseminate intelligence to the appropriate	IED Network Sweep	25%	11/8/05
cells and activities. to enhance long-term stability for litary and civil organizations.	IED Network Sweep	0%	11/9/05
ED network operating in the region of Fallujah	IED Network Sweep	0%	11/10/05

Applet com.baesystems.eif.dif.tools.launcher.ToolsLauncherApplet started

Local intranet

Federated Access To GeoSpatial Data Products

BAE SYSTEMS



BAE SYSTEMS

Questions ?