

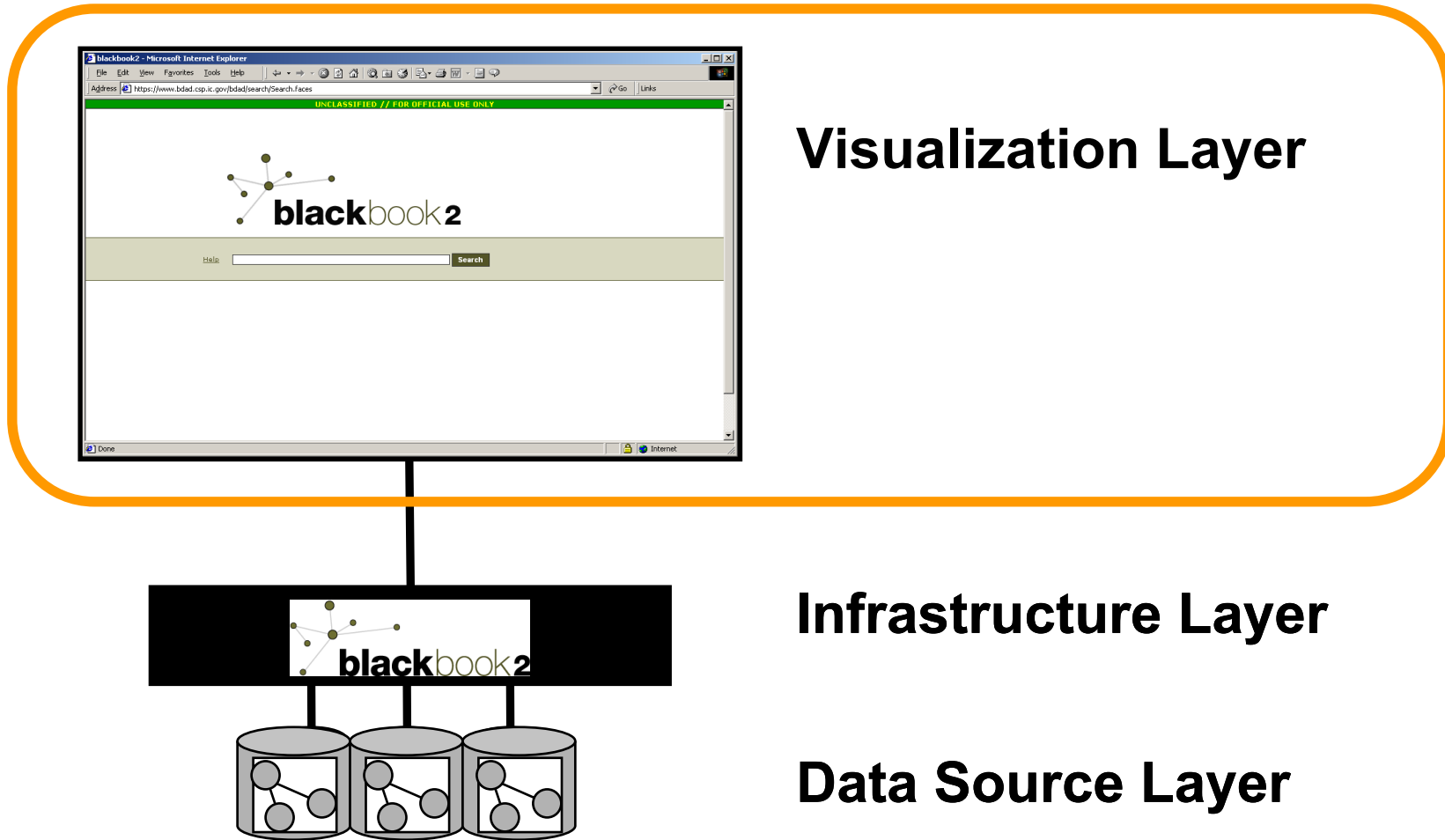


blackbook**2**

Overview

- Blackbook2 is a J2EE server-based data integration framework
- Relies on open standards to promote robustness and interoperability
 - JENA, JUNG, Lucene, JAAS, D2RQ
- Based on semantic web technologies
 - RDF, RDF Schema, OWL, SPARQL
 - Vocabulary agnostic
- Provides a default web application interface, SOAP and RESTful interfaces
- Blackbook2 is PL3 Appendix E certified (PL3+)

Architecture

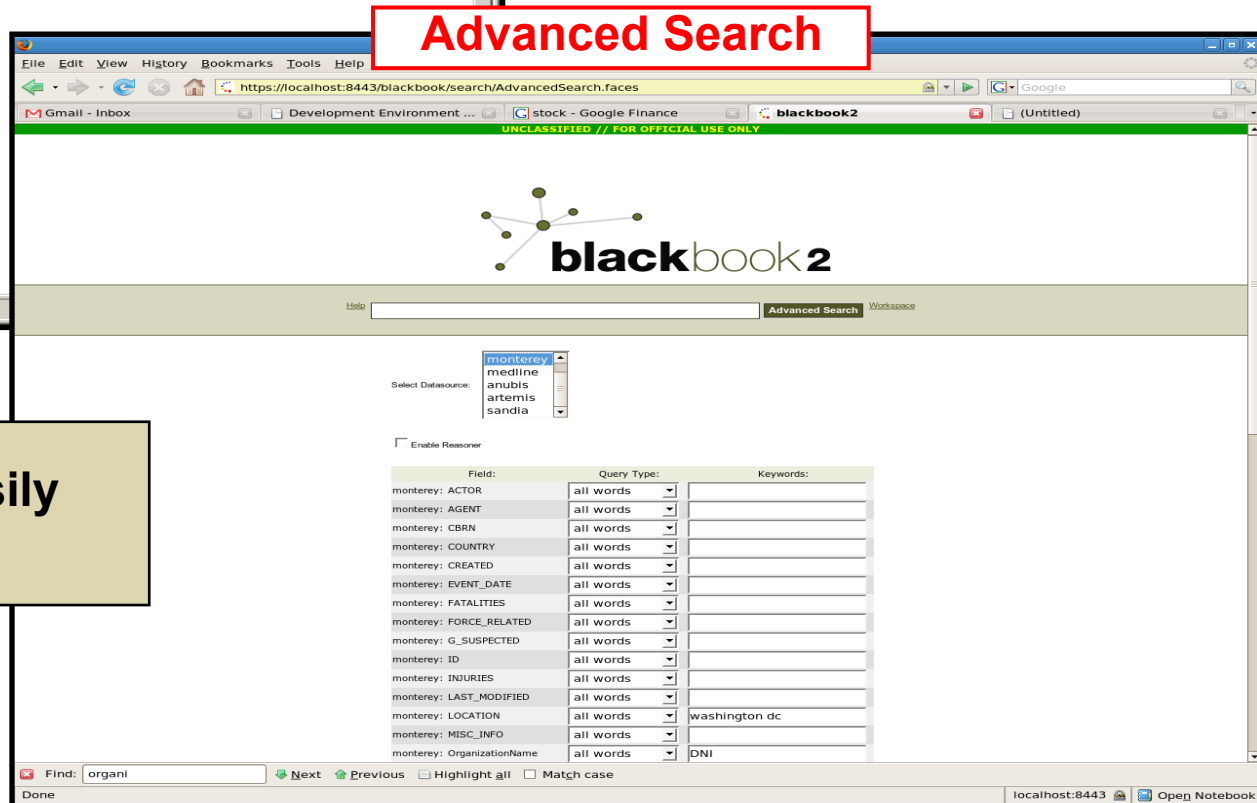
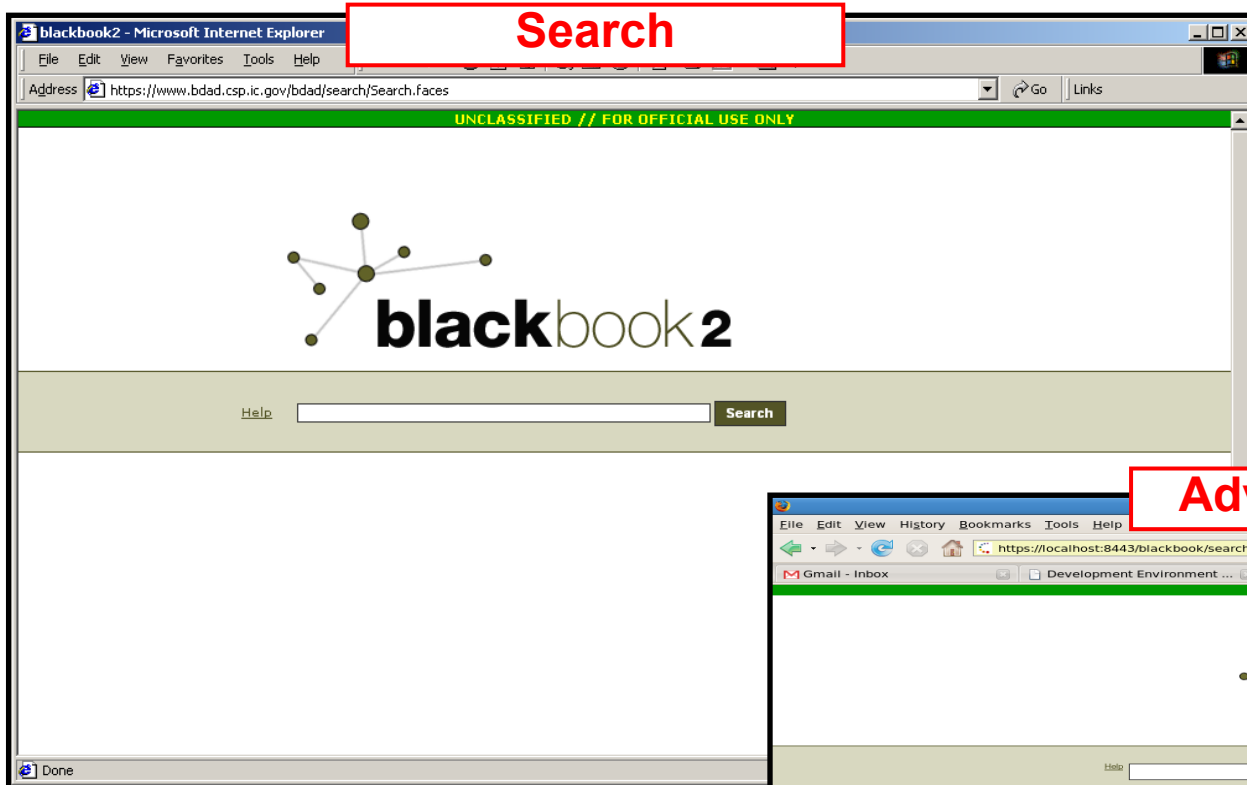


Visualization Layer

Infrastructure Layer

Data Source Layer

User Interface



A front-end "Google-like" user interface allows analysts to easily perform keyword and attribute based searches.

User Interface

Google-like Results

The screenshot shows the 'blackbook2' web application in Mozilla Firefox. The search term 'anthrax' is entered in the search bar. The results list includes:

- Aum Shinrikyo: Once and Future Threat?** Aum Shinrikyo began its public campaign of terror on June 27, 1994. On that Monday in Matsumoto, a city of 300,000 population 322 kilometers northwest...
- A Poisonous Plot** Watching the police officers come and go, some of them in protective white suits and masks, and seeing the long hours they spent in the top-floor apar...
- The missing pieces** Al-Qaeda commander Abu Mussab al-Zarqawi needed treatment for a shattered leg that was injured, apparently, during the American bombing raids on Afgha...
- La Victoria** Suspected National Liberation Army (ELN) guerrillas killed one policeman and injured another policeman and a...
- Sitra** Arsonists set fire to a store in Sitra, killing a Bangladeshi and injuring another. Shia extremists are suspected.
- Aikimbayev** Aikimbayev
- Atshabar** Atshabar

Network

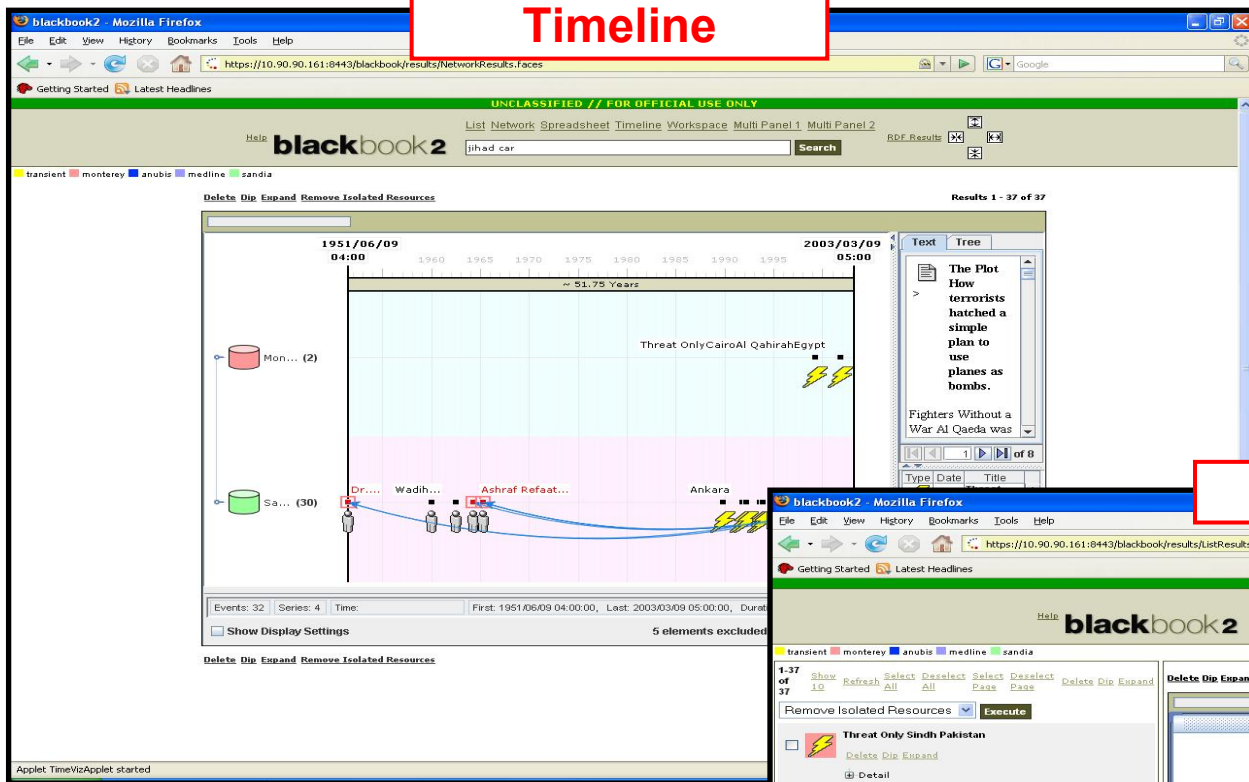
The screenshot shows the 'Network' view of the 'blackbook2' web application. The search term is 'jihad car'. The network diagram displays various entities and their relationships:

- Central Node:** MAJ - Services Office International Islamic Front for Jihad Against the Jews and Crusaders
- Peripheral Nodes:**
 - Ankara
 - Cairo
 - Wadh El Hage
 - Background: Abu Omar
 - Does Bin Laden pose a Threat to Israel?
 - Terrorism against Islamic Organizations and Groups
 - The 'Afghan Alumni'
 - Arab Veterans of the Afghan War
 - Wadiah El Hage
 - The Demise of Radical Islam in Turkey
 - OSIS
 - man to al-Qaida's
 - Ottawa
 - Canada's
 - Mosler
 - Masri
 - The tenants of
 - The Shoe Bomber's World
 - Dr. al Ayman Mohammed Rabih Zawahir
 - Hamball
 - Ashraf Refaat Nabith Henin
 - The CEO of al-Qaada: Khaled Sheikh Mohammed
 - Bold Tracks of Terrorism's Mastermind
 - The Plot How terrorists hatched a simple plan to use planes as bombs.
 - THE MAN BEHIND BIN LADEN: How an Egyptian doctor became a master of terror.
 - attention turns to the other prime suspect

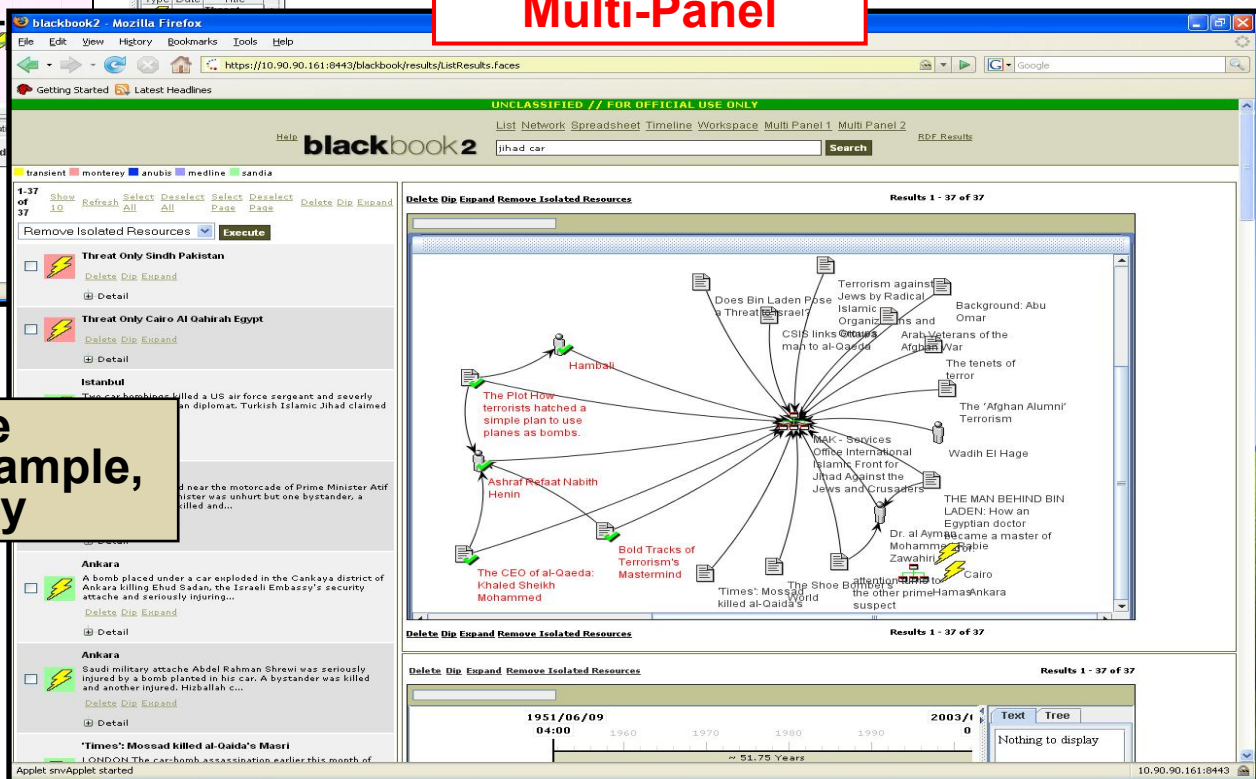
Different ways to view the same information. "Network", for example, displays entities of different types and their relationships to other entities.

User Interface

Timeline

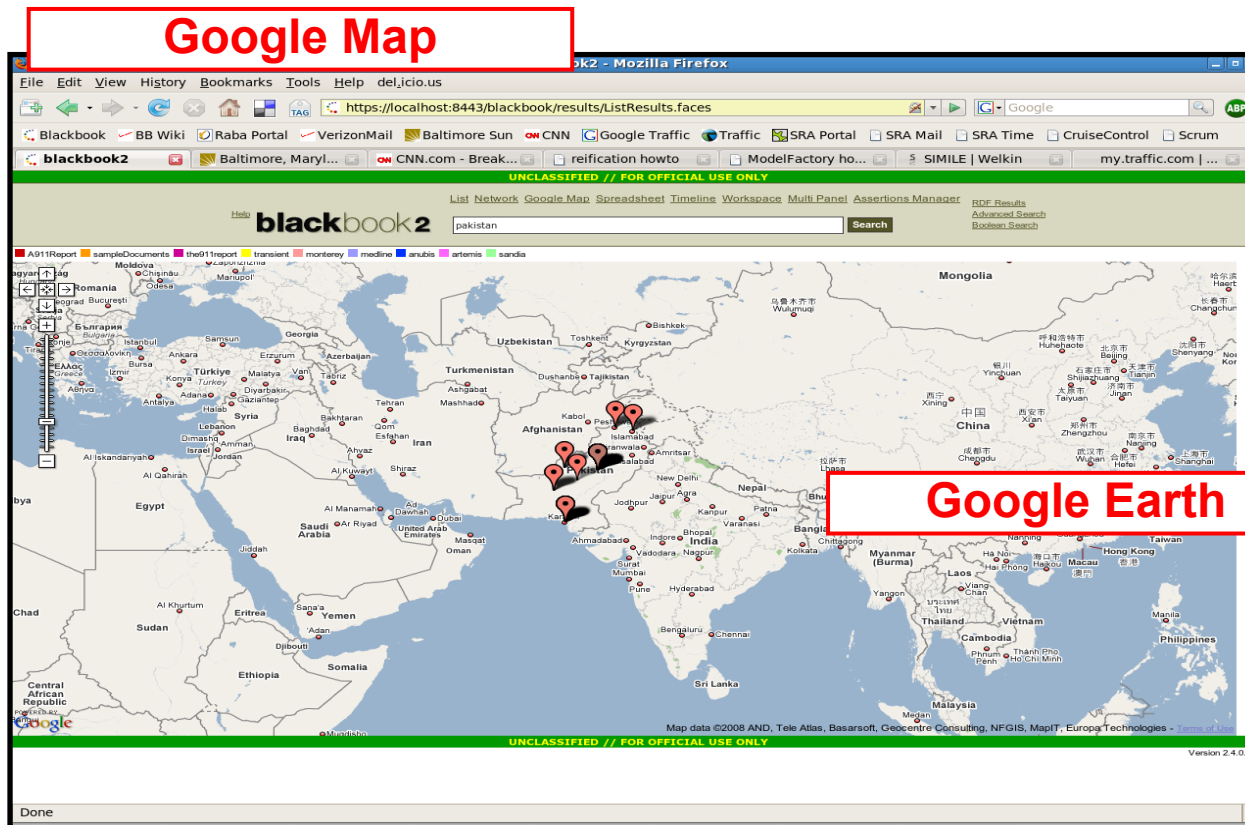


Multi-Panel



Different ways to view the same information. "Timeline", for example, displays entities chronologically

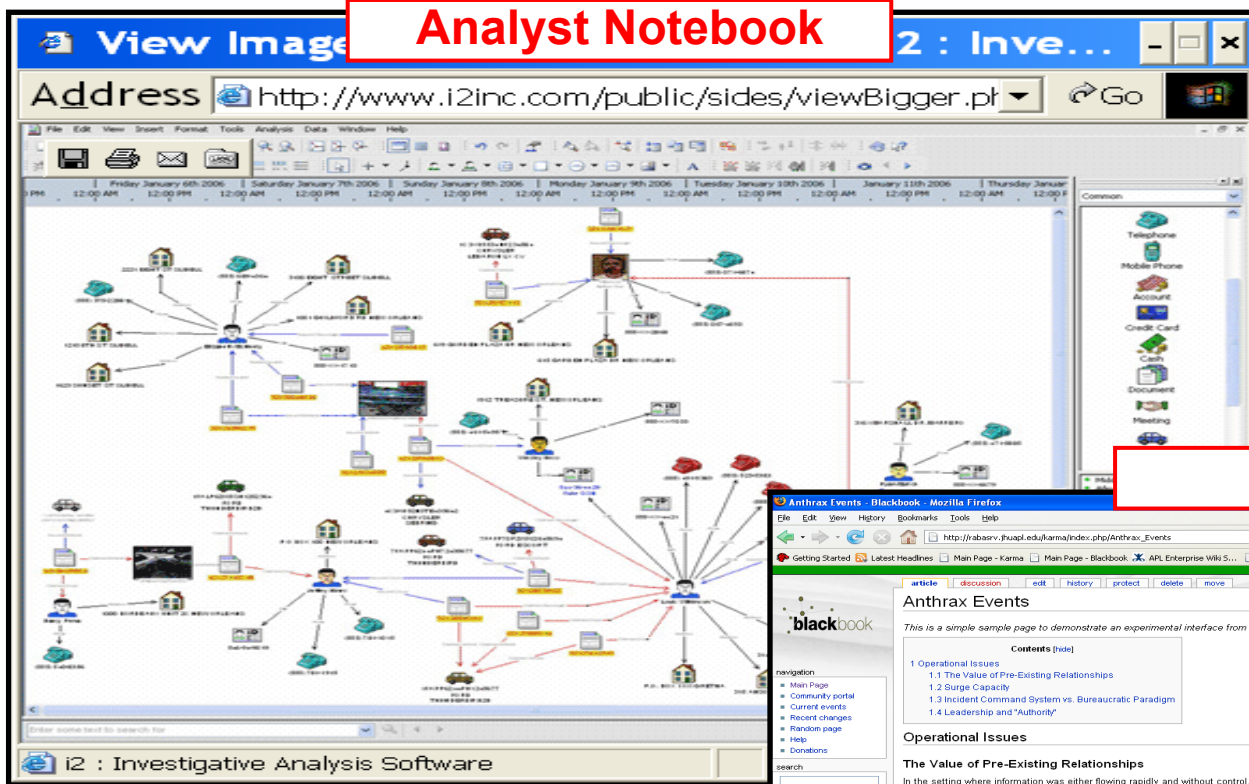
User Interface



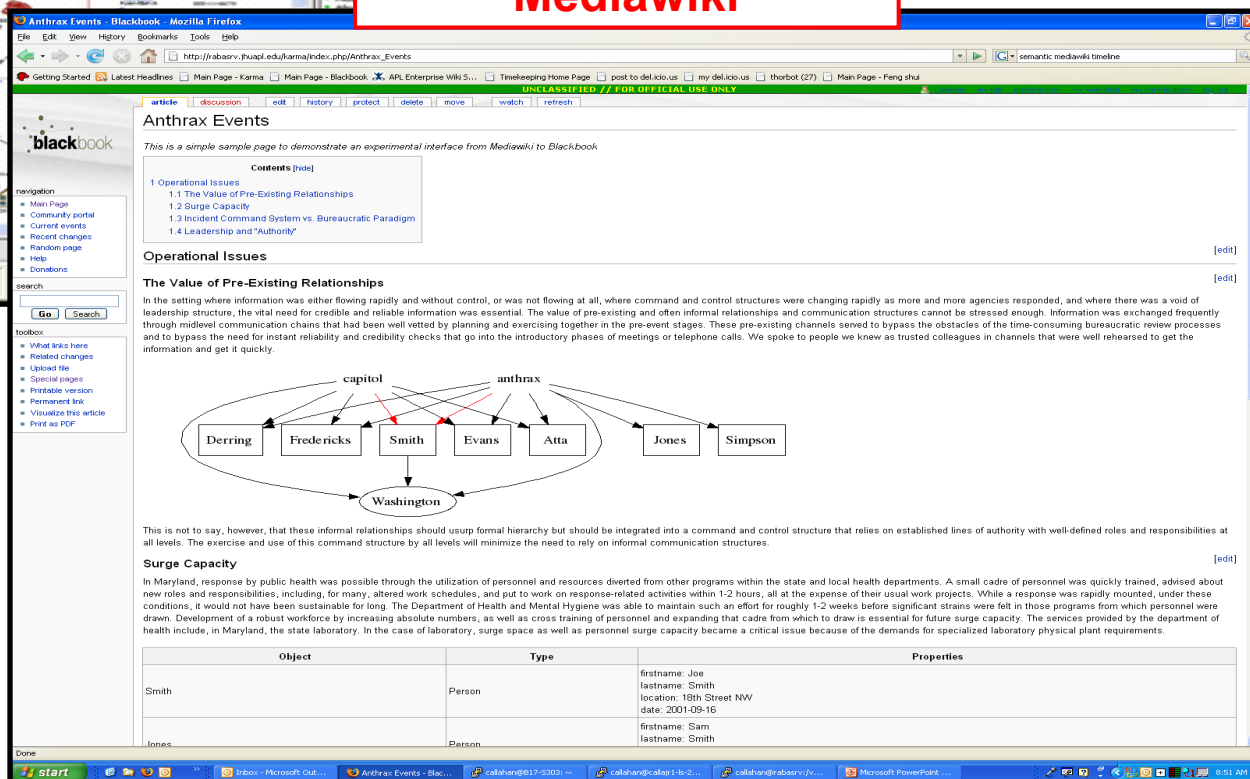
Allows analysts to visualize geospatial content using Google-map and Google Earth.

User Interface

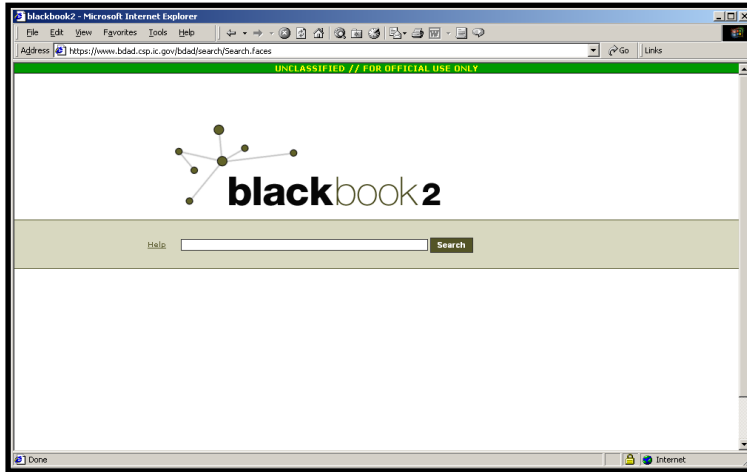
Analyst Notebook



Mediawiki



Architecture



Visualization Layer



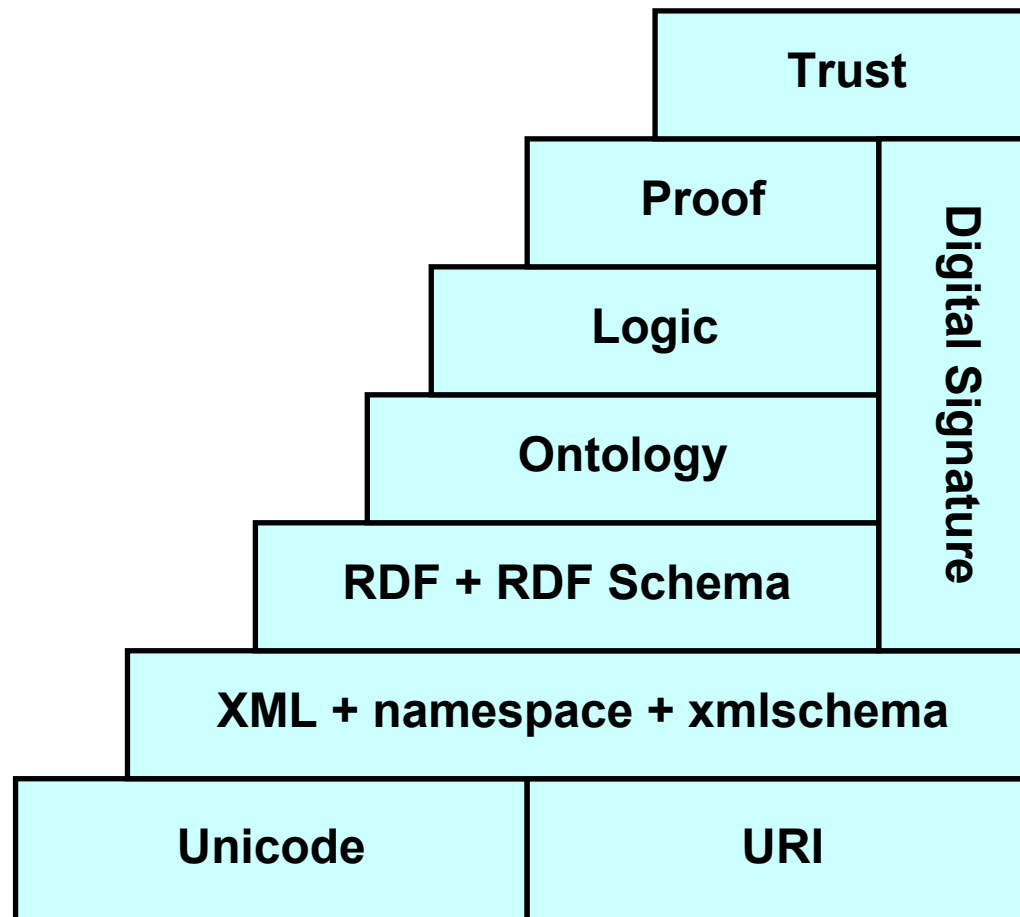
Infrastructure Layer



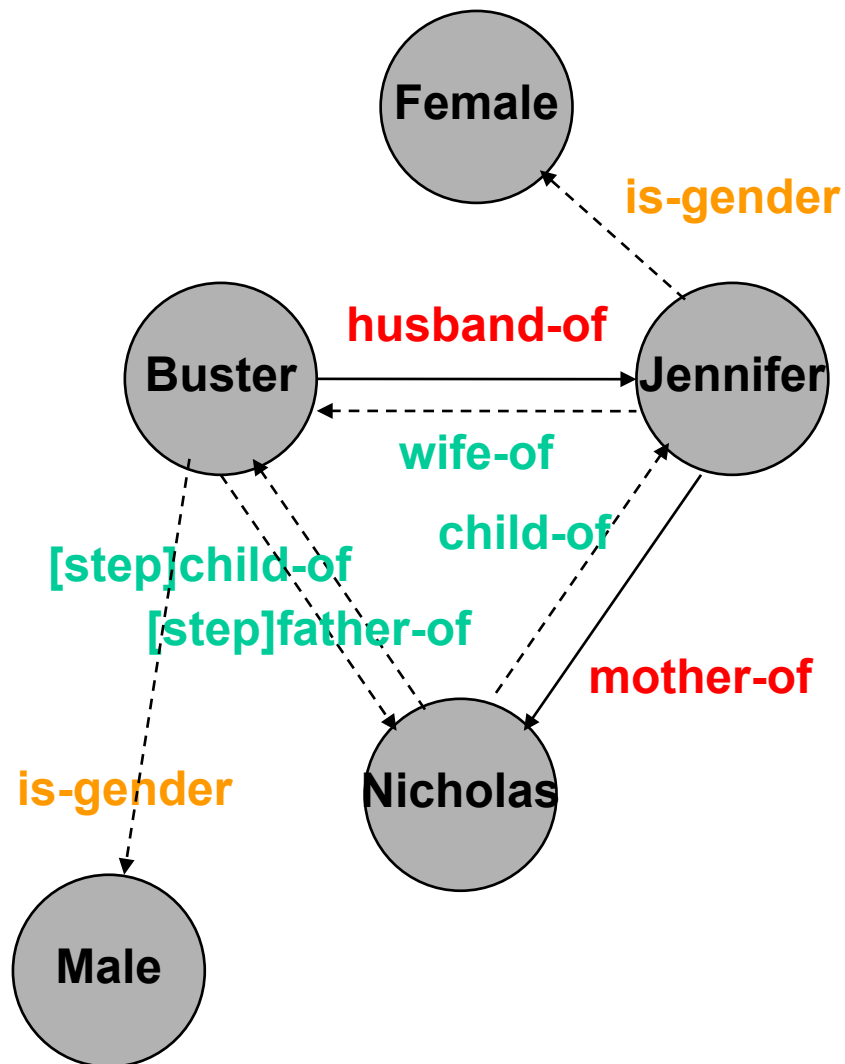
Data Source Layer

Semantic Web

- The Semantic Web is the next generation of the current web in which computers can interpret the meaning of the web content because of explicit semantics provided in markup.



Example 1: Inference



An analyst creates:

- 1) Entity "Buster"
- 2) Entity "Jennifer"
- 3) Entity "Nicholas"

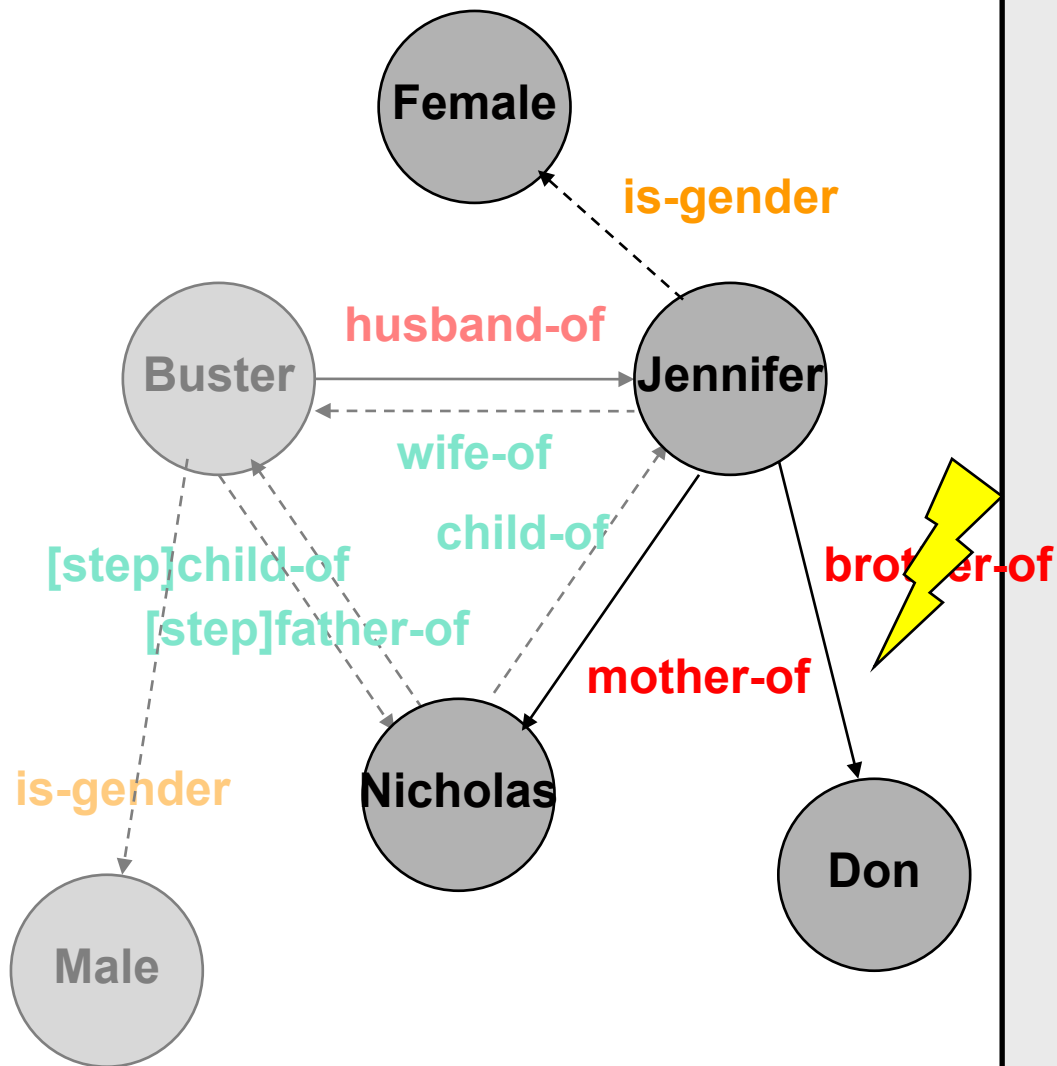
An analyst makes the assertion:

- 4) "Buster **husband-of** Jennifer"
- 5) "Jennifer **mother-of** Nicholas"

Blackbook system can infer:

- 6) "Jennifer **wife-of** Buster"
- 7) "Nicholas **child-of** Jennifer"
- 8) "Buster **[step]father-of** Nicholas"
- 9) "Nicholas **[step]child-of** Buster"
- 10) "Buster **is-gender** Male"
- 11) "Jennifer **is-gender** Female"

Example 2: Invalid Logic Assertion



An analyst creates:

1) Entity "Don"

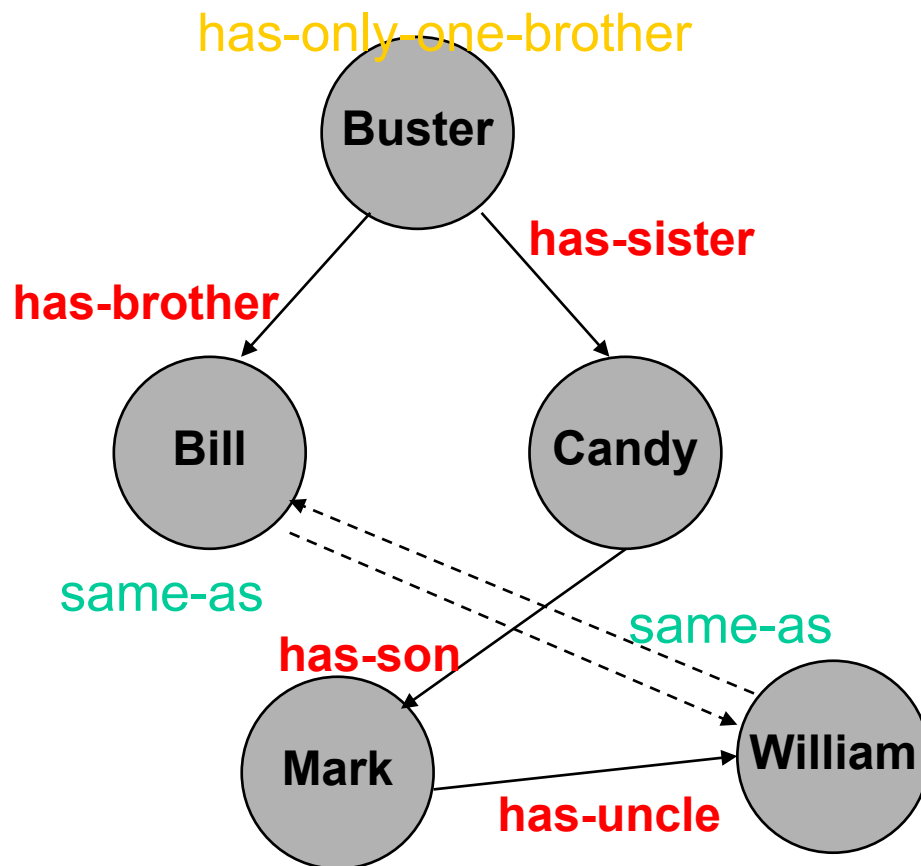
An analyst makes the assertion:

2) "Jennifer **brother-of** Don"

Blackbook system can infer:

3) Invalid Assertion
(Gender conflict)

Example 3: Constraints & same-as



An analyst makes the assertion:

- 1) "Buster **has-brother** Bill`
- 2) "Buster **has-sister** Candy`
- 3) "Candy **has-son** Mark`
- 4) "Mark **has-uncle** William`

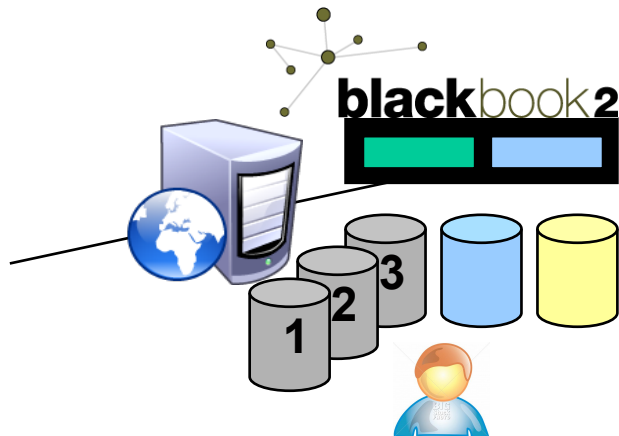
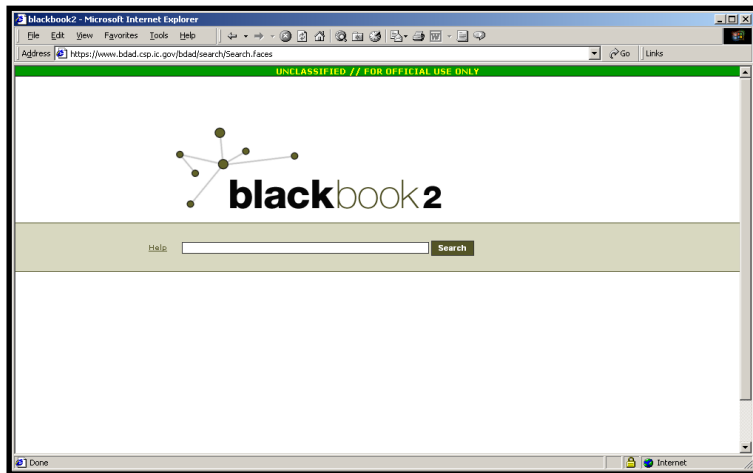
An analyst applies the constraint:

- 5) "Buster **has-only-one-brother** Bill`

Blackbook system can infer:

- 6) "William **same-as** Bill`
- 7) "Bill **same-as** William`

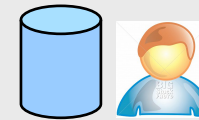
Algorithms, Security, AKB



Algorithm plug-ins can be added



Security PL3+ / User Credentials



CASPORT
Scattered Castles
DIAS

Analysts can store assertions into an Analytic Knowledge Base (AKB)



User Interface

Workflow

The screenshot shows the 'blackbook2' workflow interface. It includes a 'Process Flow' table with the following data:

States	To States	Additional Criteria
0. Expand <input type="checkbox"/> fork	<none>	DataAccess: transient
1. Lucene Keyword <input type="checkbox"/> fork	2. Materialize	DataAccess: transient val:jihad car
2. Materialize <input checked="" type="checkbox"/> fork	0. Expand 1. Lucene Keyword	DataAccess: transient
3. Dip <input type="checkbox"/> fork	<none>	DataAccess: transient

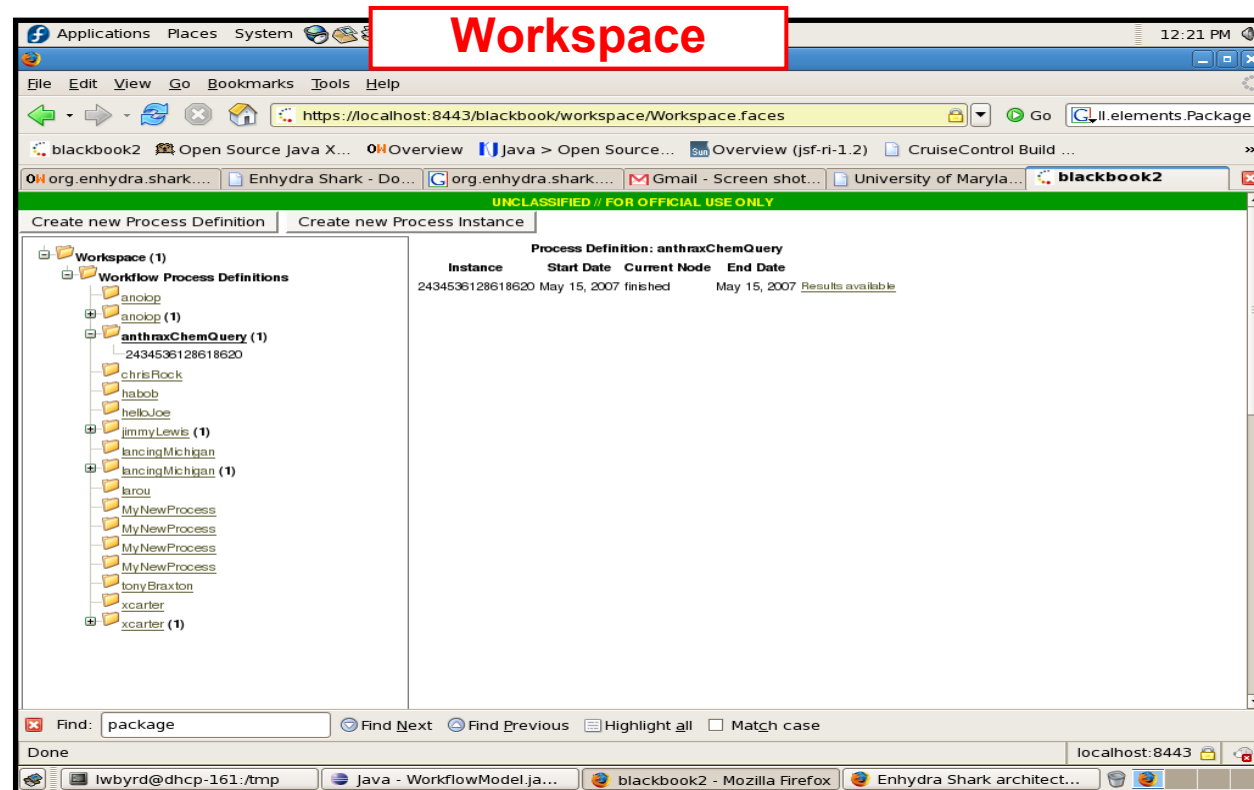
Yahoo Pipes

The screenshot shows the 'Yahoo Pipes' interface. It includes a 'Sources' list and a 'User inputs' section. The pipeline shown consists of the following pipes:

- Where (location)
- URL Builder
- Fetch Feed

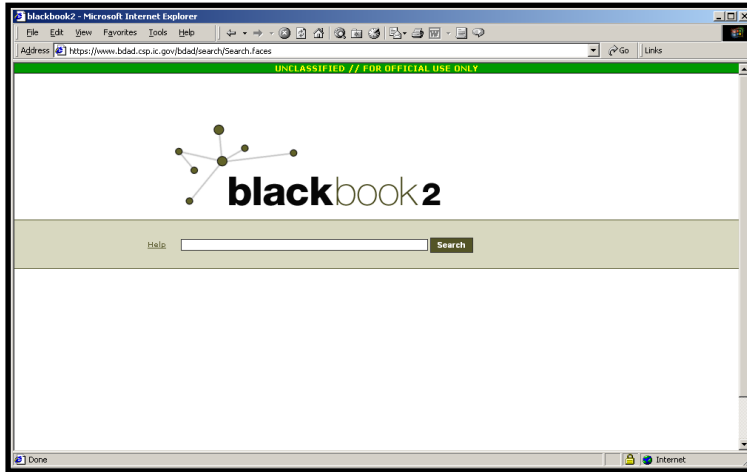
“Workflow` allow analysts to define the order of tasks, configure algorithm parameters, and batch processes concurrently

User Interface



“Workflow` and “Workspace` allow analysts to define the order of tasks, store them in private folders and/or share them publicly with colleagues.

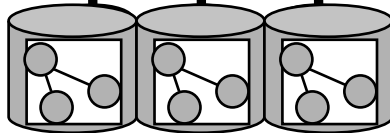
Architecture



Visualization Layer



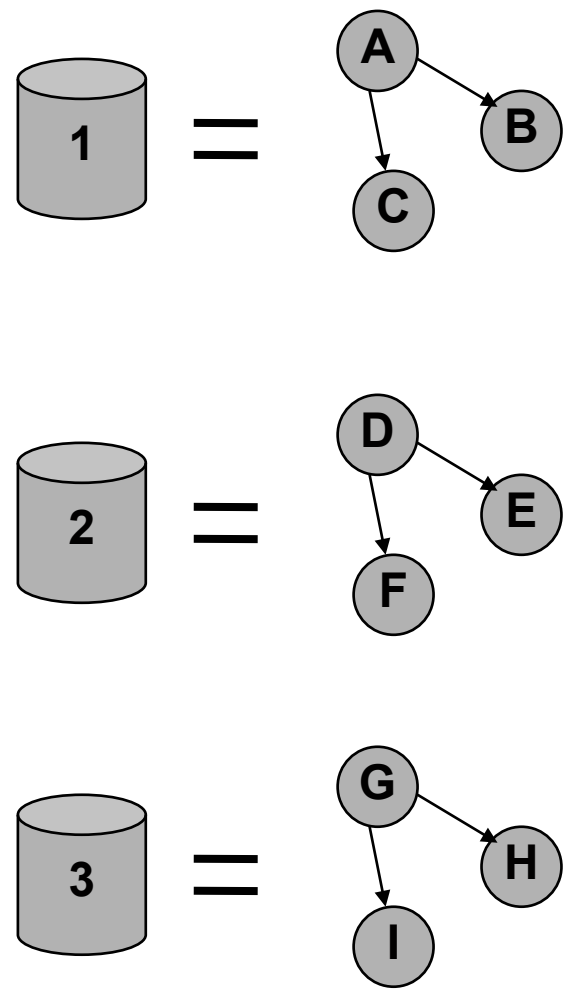
Infrastructure Layer



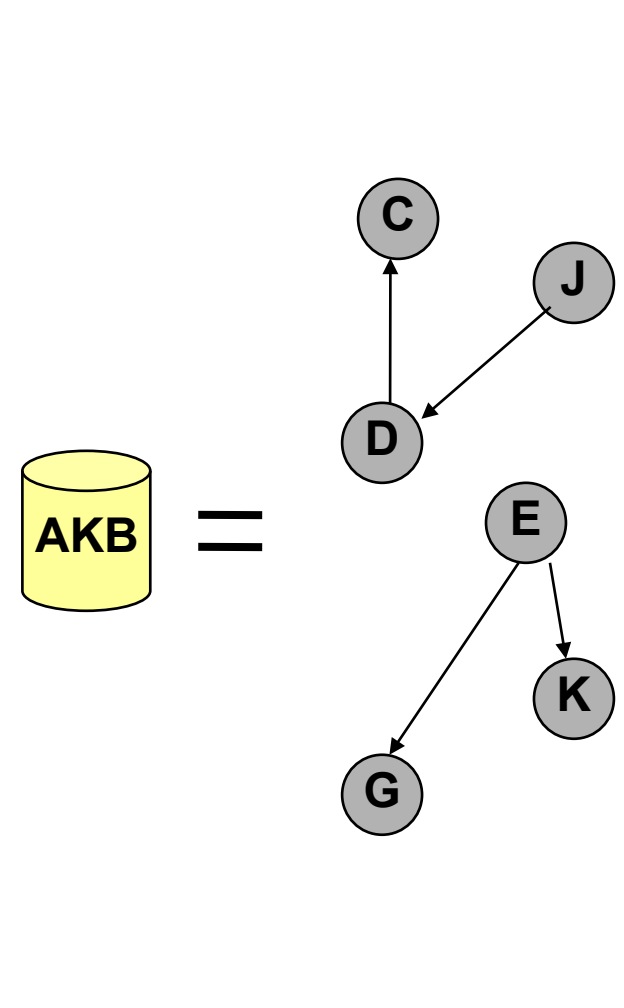
Data Source Layer

Composite Knowledge

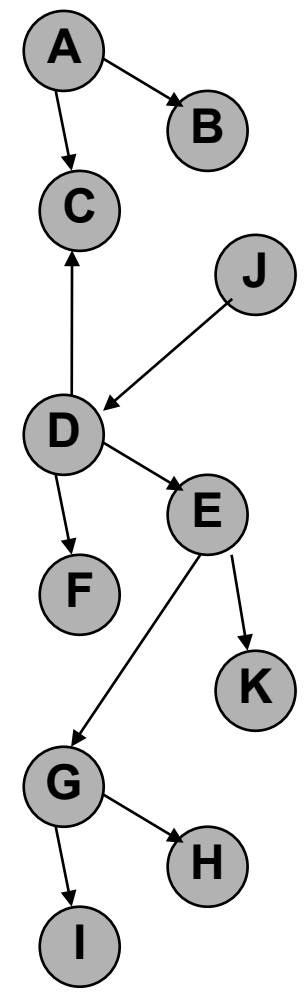
Original Datasource



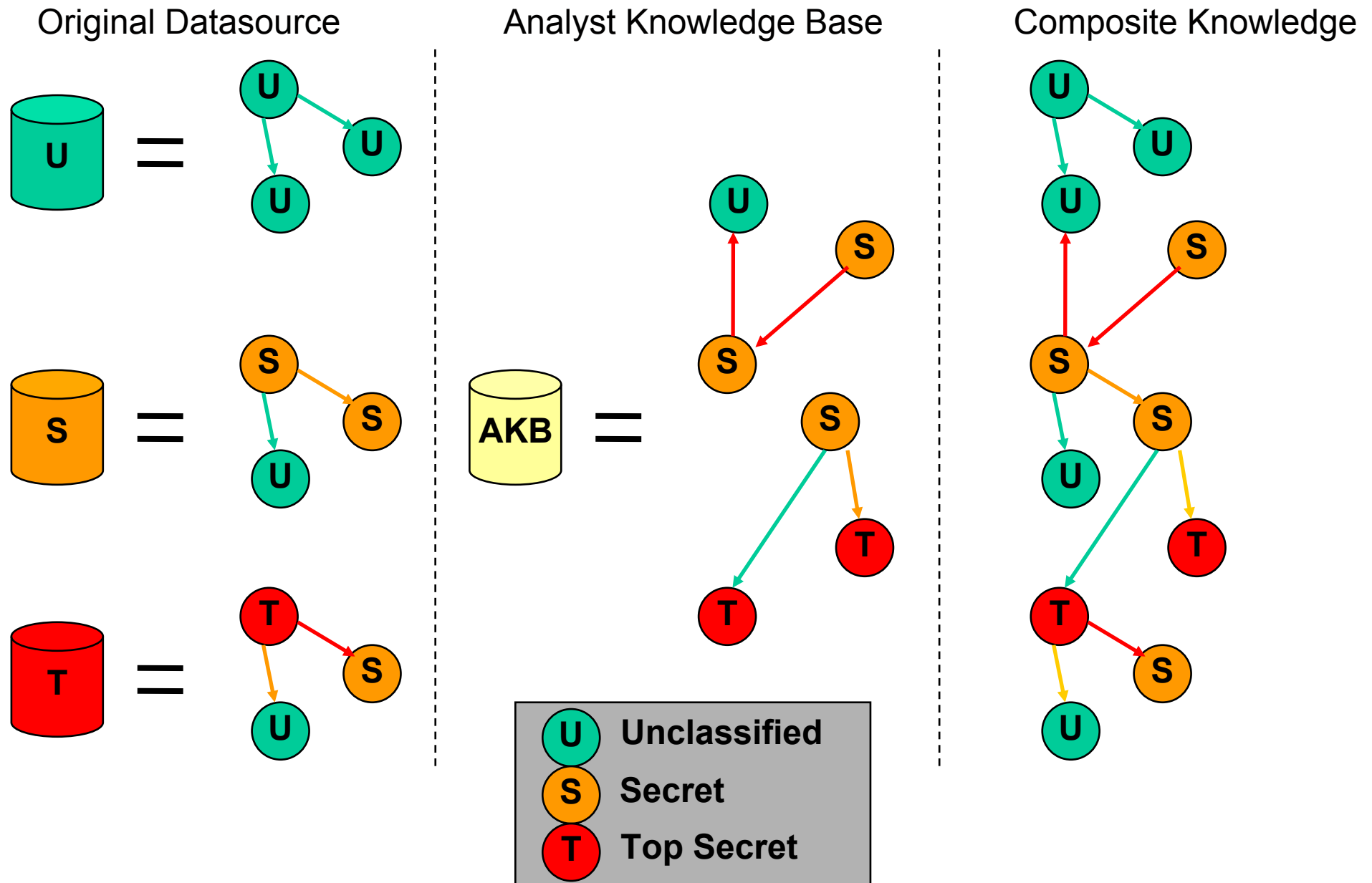
Analyst Knowledge Base



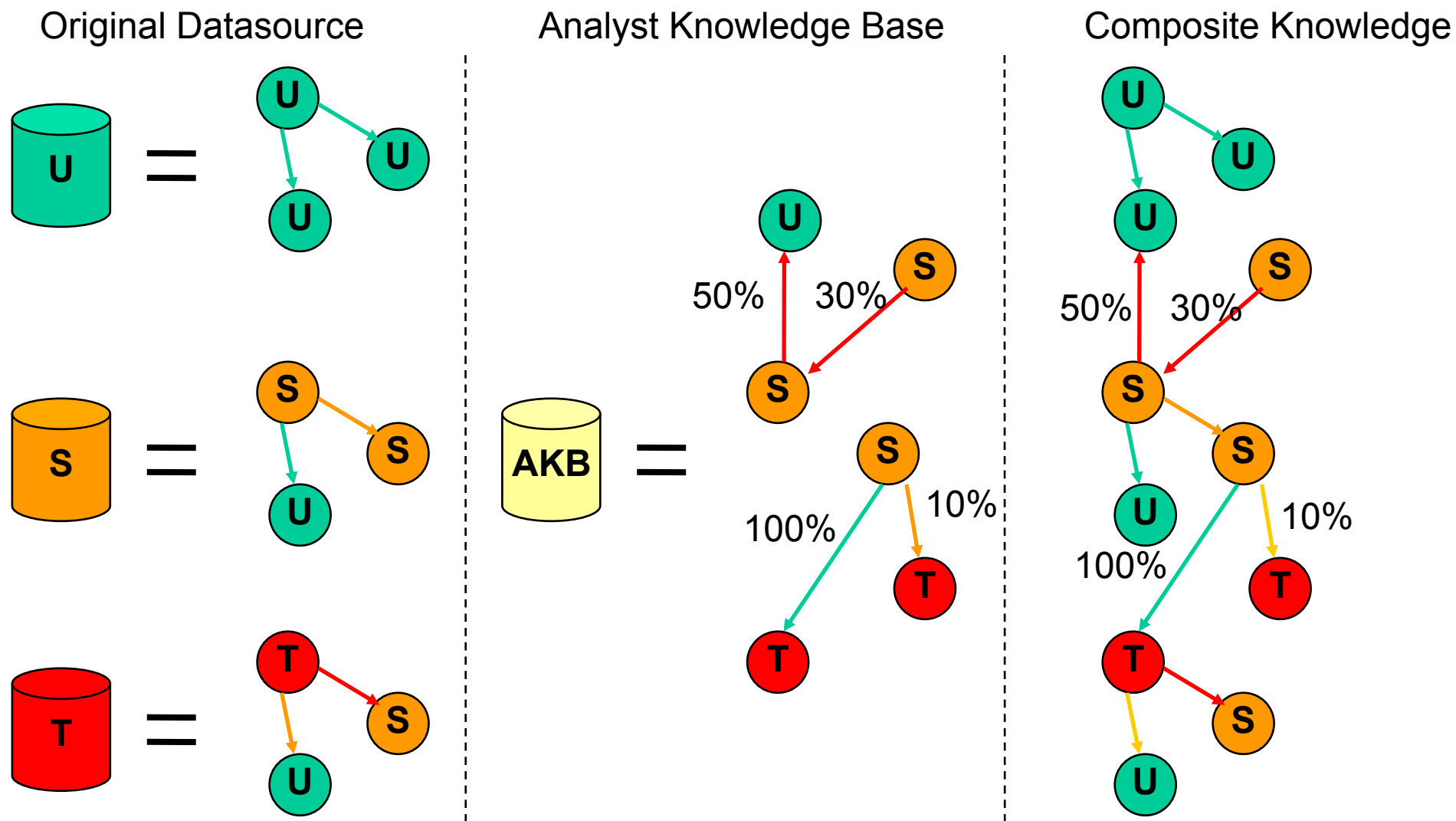
Composite Knowledge



Composite Knowledge with Security



Composite Knowledge with Confidence



User Interface

Relationship Manager

Allows analysts to specify the relationship between two or more entities

Entity Manager

Allows analysts to create entities of different types, and modify attributes

Ontology Import

Allows analysts to upload their own ontology

Unstructured/Structured to RDF

THE
9/11
COMMISSION
REPORT

FINAL REPORT OF THE NATIONAL COMMISSION ON TERRORIST ATTACKS UNDER THE 9/11 COMMISSION

OFFICIAL GOVERNMENT EDITION

Unstructured

THE ATTACK LOOMS

ALS IN CALIFORNIA

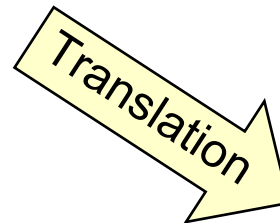
described the Southeast Asia travels of Nawaf al Hazmi, r, and others in January 2000 on the first part of the "planes at chapter we also described how Mihdhar was spotted in rly in January 2000, along with associates who were not en was lost to sight when the group passed through uary 15, Hazmi and Mihdhar arrived in Los Angeles. They weeks there before moving on to San Diego.¹

Los Angeles
Mihdhar came to California, we do not know for certain. Khalid ed (KSM), the organizer of the planes operation, explains that convenient point of entry from Asia and had the added far away from the intended target area.²

har were ill-prepared for a mission in the United States. Their s for this plot were their devotion to Usama Bin Ladin, their and their ability to get valid U.S. visas. Neither had spent any in the West, and neither spoke much, if any, English.³

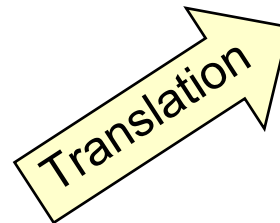
re be plausible that they or KSM would have tried to identify, in advance, a friendly contact for them in the United States. In detention, KSM denies that al Qaeda had any agents in Southern California. We do not credit this denial.⁴ We believe it is unlikely that Hazmi and Mihdhar—neither of whom, in contrast to the Hamburg group, had any prior exposure to life in the West— would have come to the United States without arranging to receive assistance from one or more individuals informed in advance of their arrival.⁵

KSM says that though he told others involved in the conspiracy to stay away from mosques and to avoid establishing personal contacts, he made an



Structured

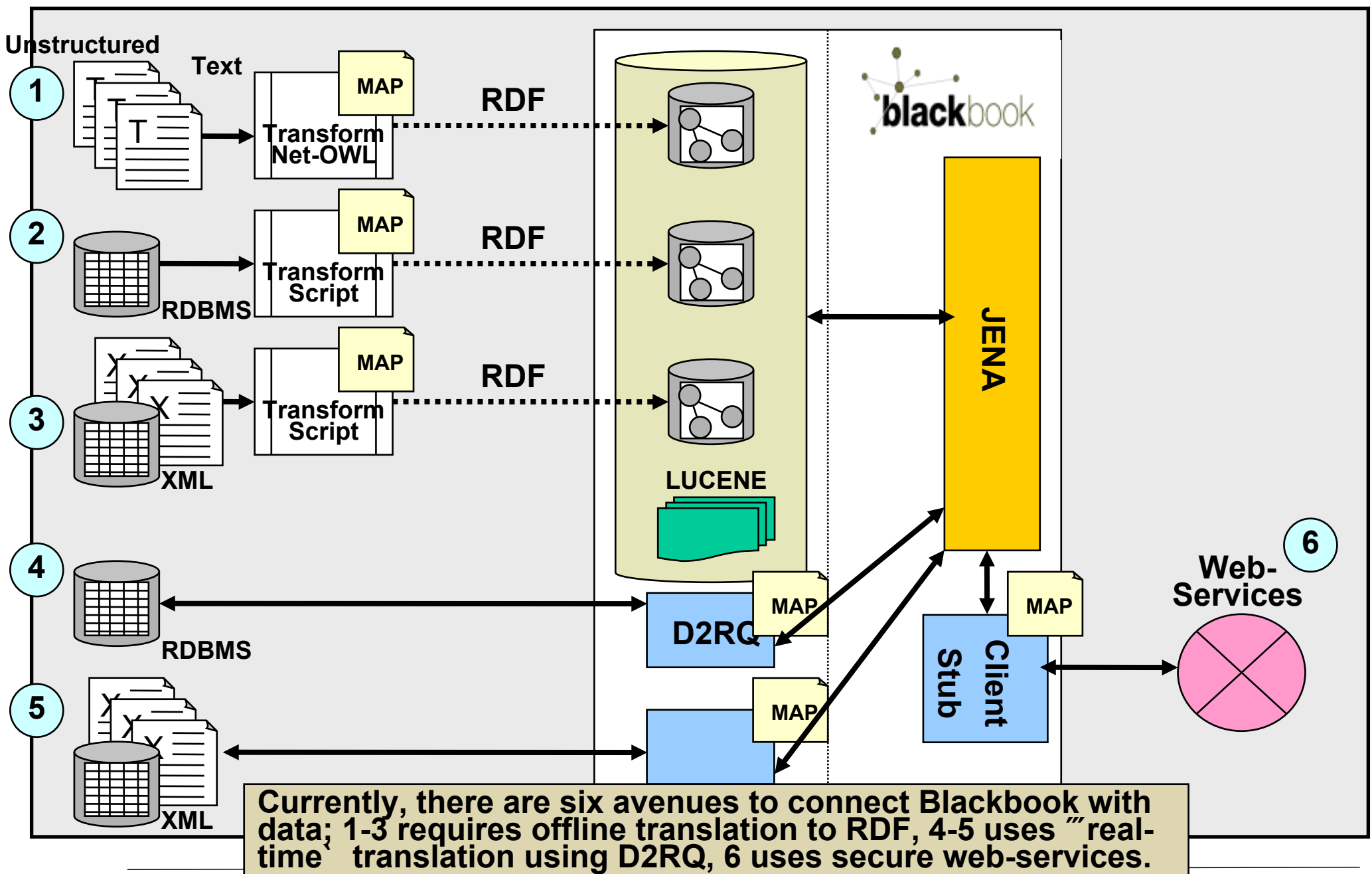
RDMS/XML



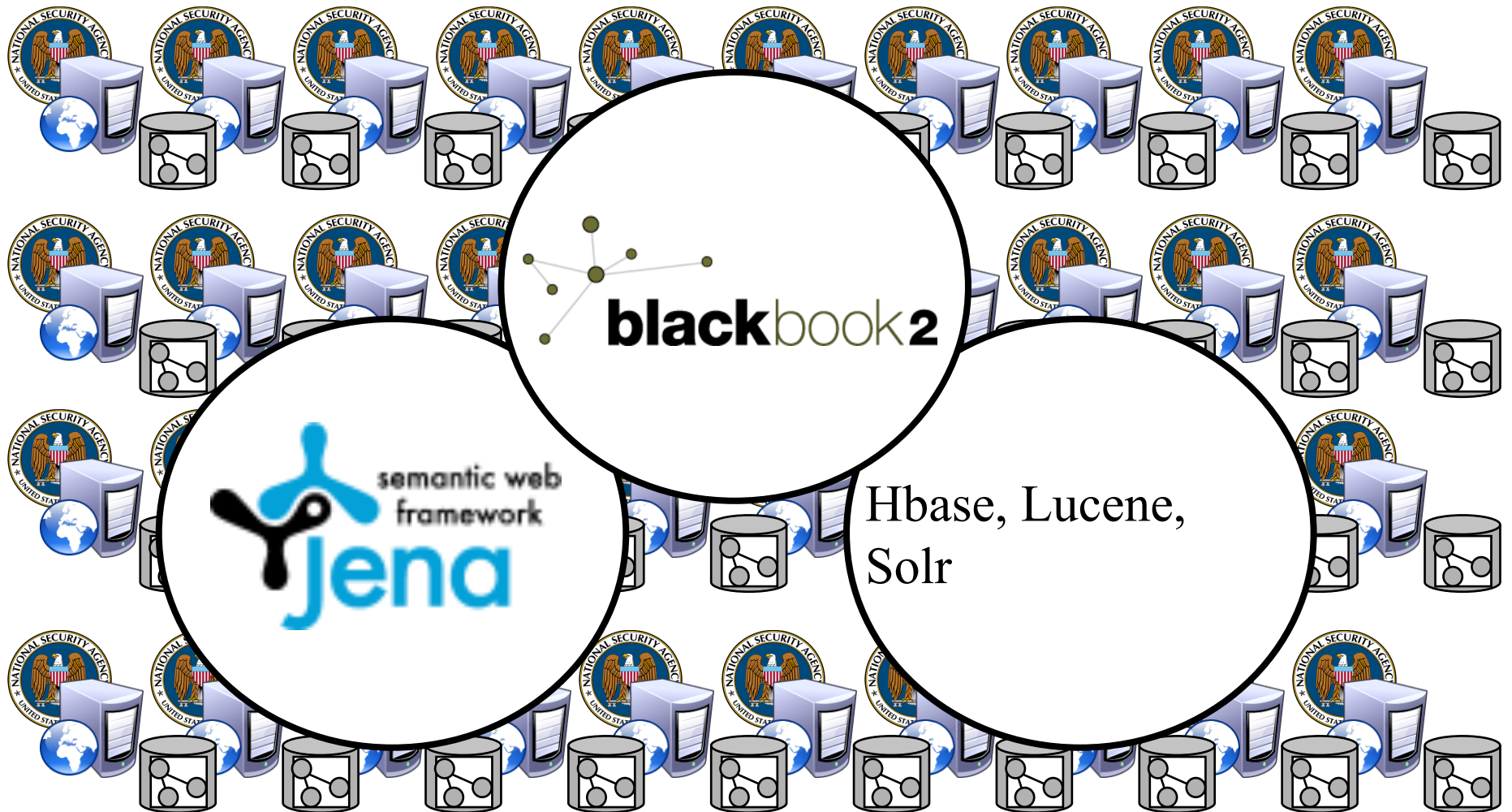
RDF

Unlike most applications, Blackbook performs queries on data in RDF form, not relational form.

Datasource Connectivity



Blackbook and Alternate Stores



Scalability using Hadoop

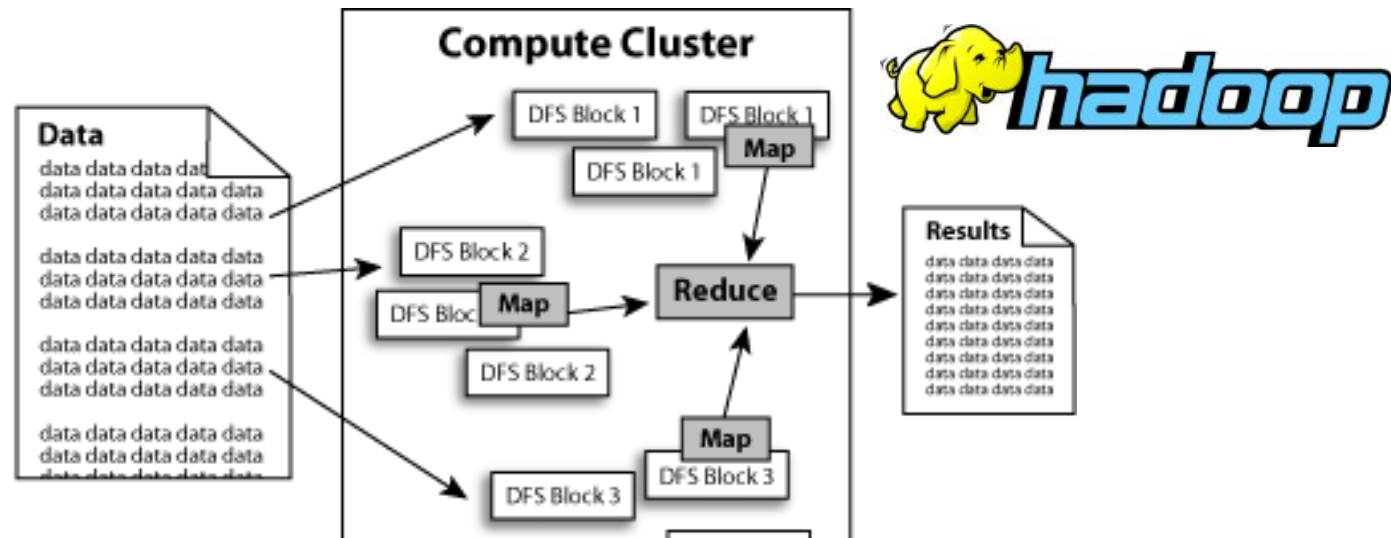
Hadoop implements MapReduce, using the Hadoop Distributed File System (*HDFS*). MapReduce divides applications into many small blocks of work. HDFS creates multiple replicas of data blocks for reliability, placing them on compute nodes around the cluster. MapReduce can then process the data where it is located.

Scalable: Hadoop can reliably store and process petabytes.

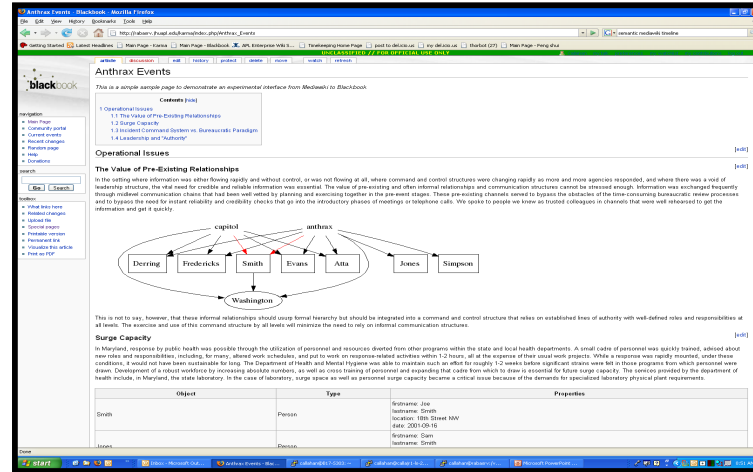
Economical: It distributes the data and processing across clusters of commonly available computers. These clusters can number into the thousands of nodes.

Efficient: By distributing the data, Hadoop can process it in parallel on the nodes where the data is located. This makes it extremely rapid.

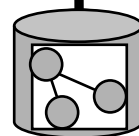
Reliable: Hadoop automatically maintains multiple copies of data and automatically redeploys computing tasks based on failures.



Blackbook and Wikis

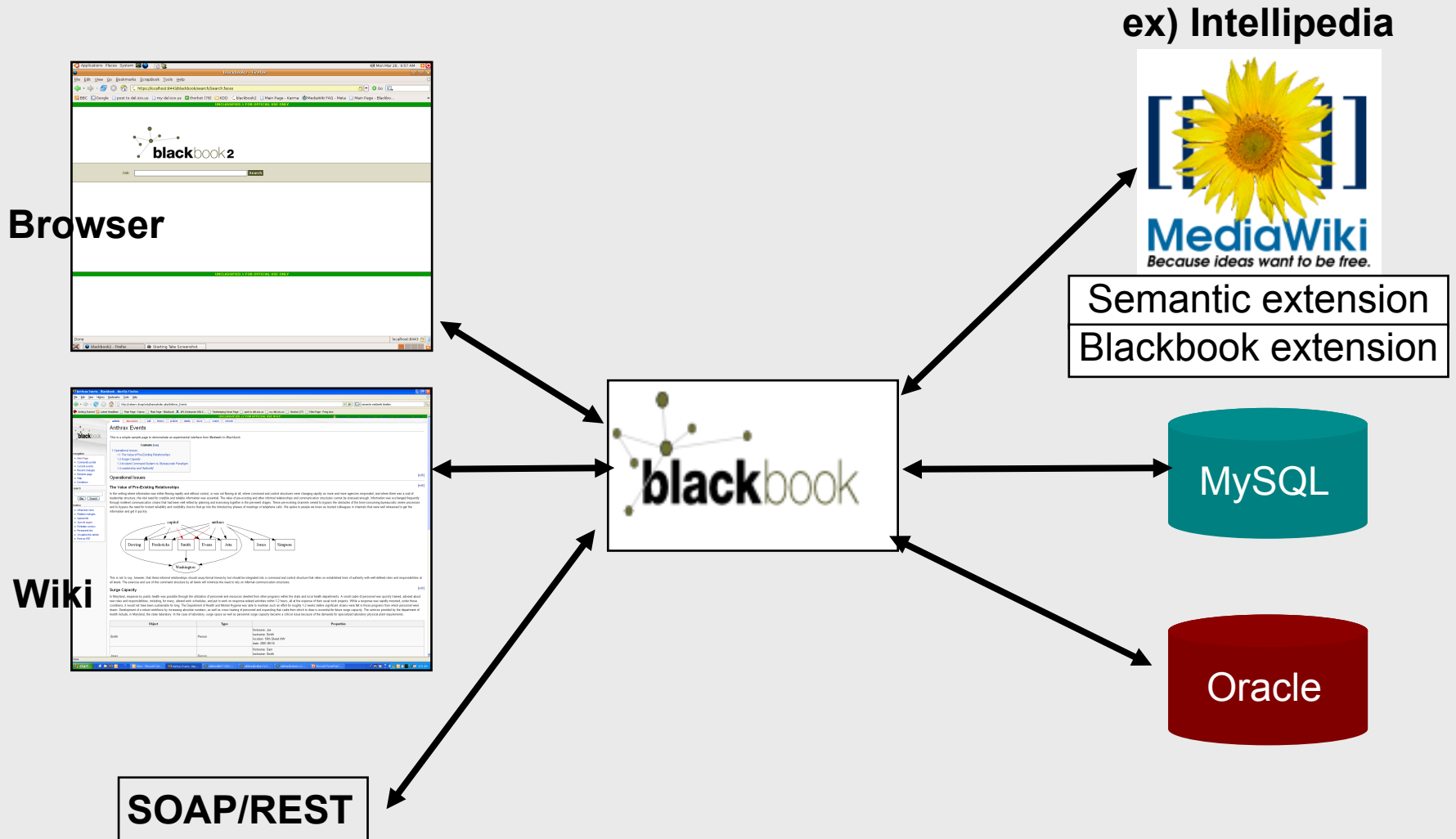


Wiki%100



Wikis

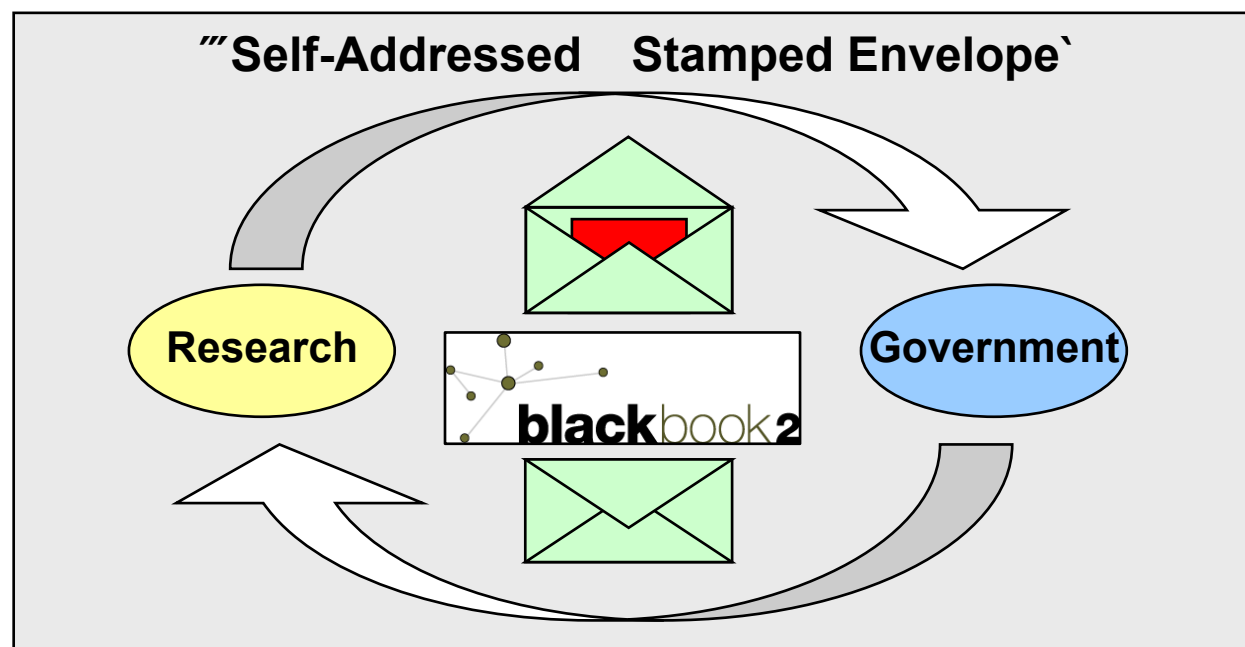
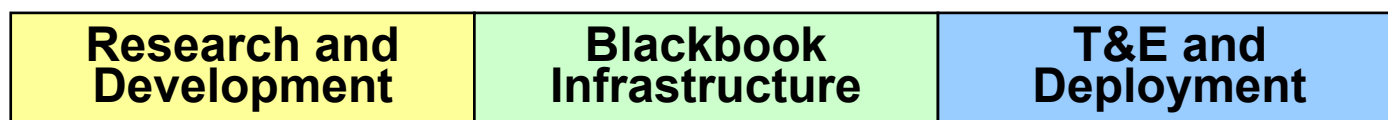
Blackbook and Wikis



Like browsers, "Wiki's can be a front-end to Blackbook. Wiki's can also be a datasource. Wiki extensions can be utilized to enable Semantic and Blackbook features.

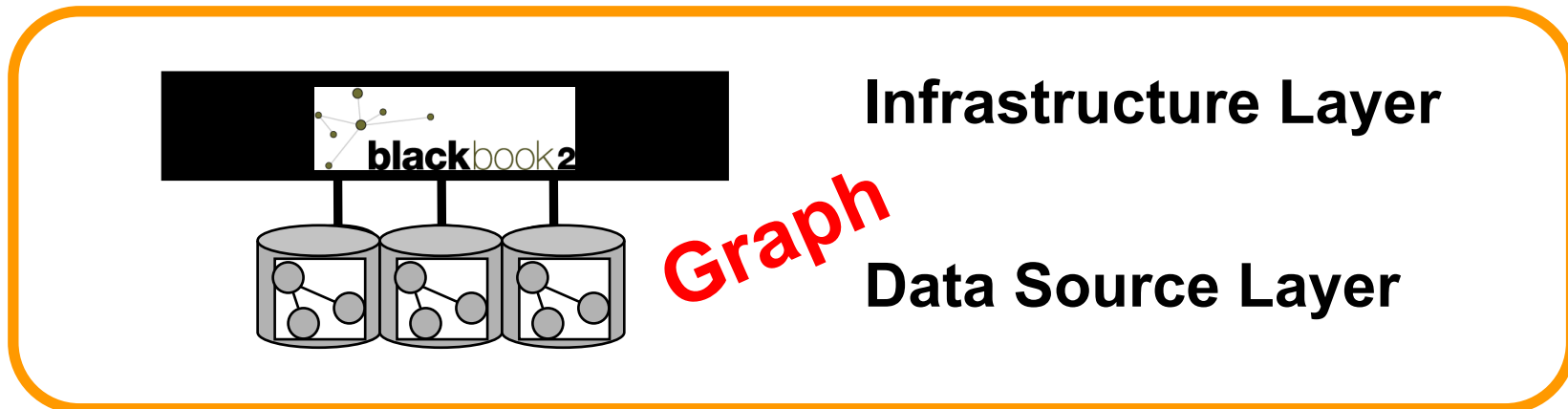
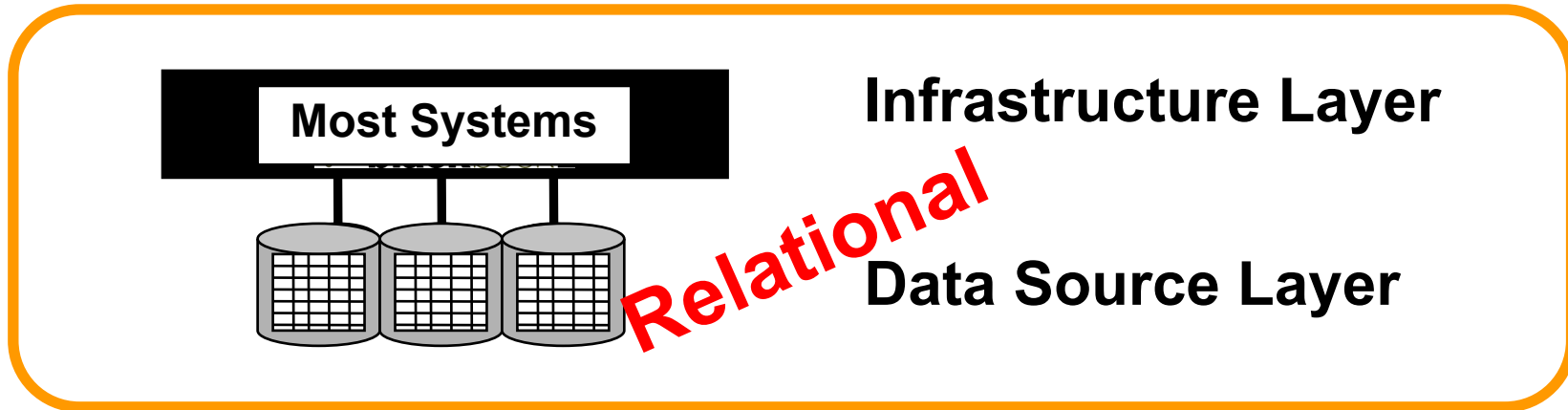
Revolutionize Technology Transfer

Improve Intelligence Analysis by Coordinated Exposition of Multiple Data Sources Across Intelligence Community Agencies



A research product (red), such as a new and improved algorithm or visualization, can easily be transferred from research to government using the Blackbook “envelope” .

Relational vs. Graph-based Systems



Blackbook2 is a JEE server-based RDF processor that provides an asynchronous interface to back-end datasources.