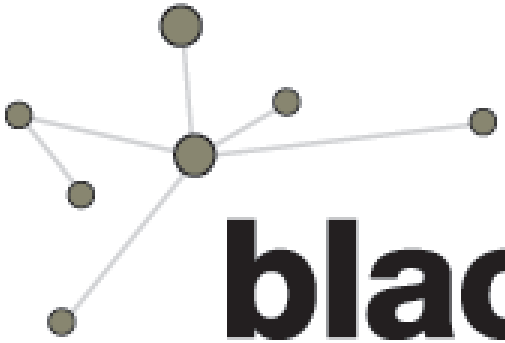


blackbook2 3

Presented by:

Buster Fields
Program Manager

Agenda

- Analytic Modernization
- Linked Data and Semantic Web
- What is Blackbook?
- Blackbook 2.x - Current Capabilities
- Blackbook 3.x - Future Capabilities
- Timeline
- Technology Transfer
- Blackbook wiki
- Q&A

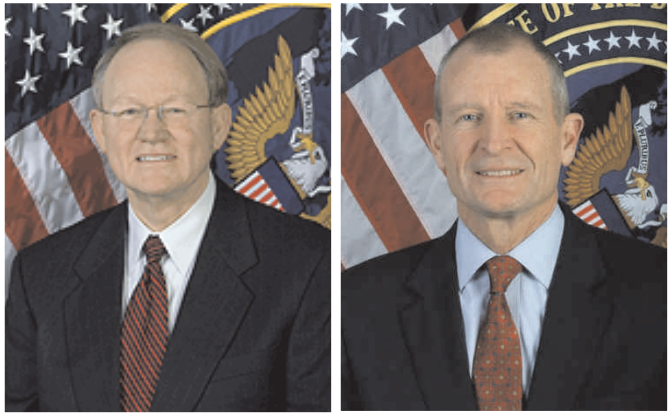
Analytic Modernization

Six Focus Areas:

- Create a Culture of Collaboration
- Accelerate Information Sharing
- Foster Collection and Analytic Transformation
  - A-SPACE – *Collaborative Environment*
  - Catalyst – *"Services of Common Interest"*
  - Library of National Intelligence – *Consolidated repository containing IC-disseminated products*
- Build Acquisition Excellence and Technology Leadership
- Modernize Business Practices
- Clarify and Align DNI's Authorities

Director of National Intelligence

Mike McConnell Dennis Blair



October 10th, 2007

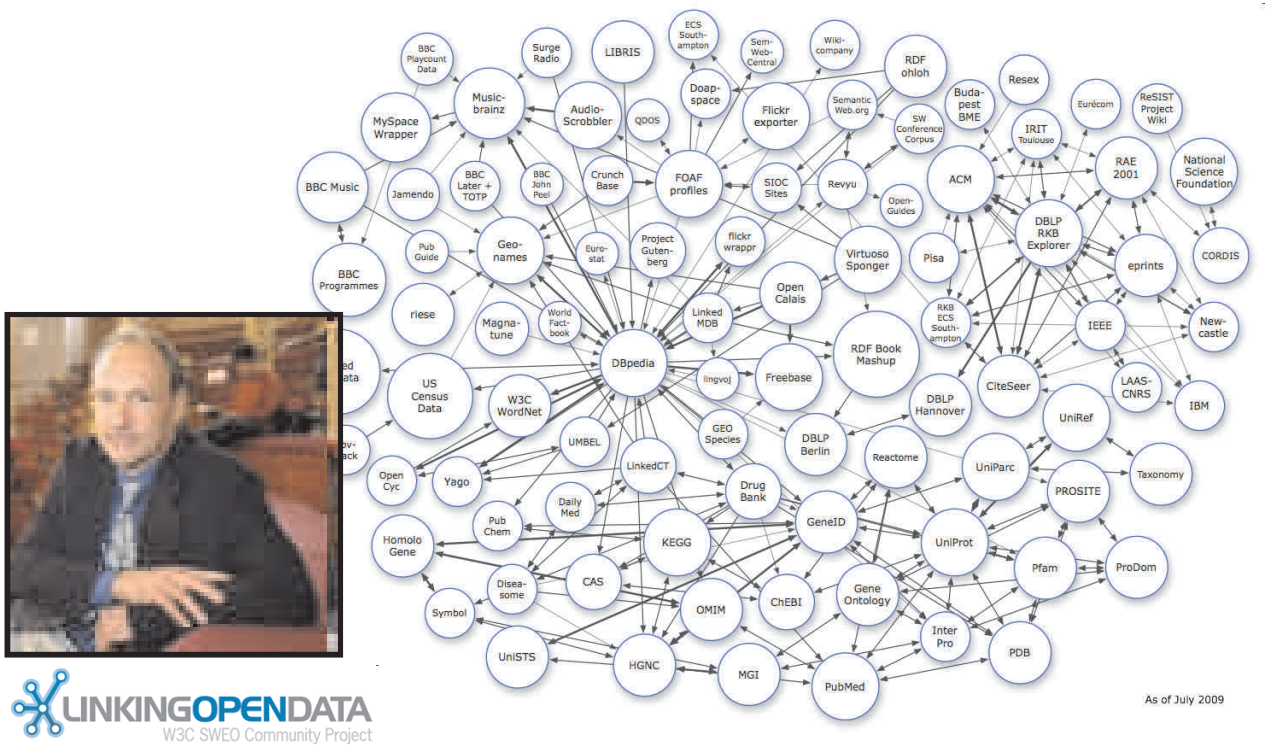
Linked Data

- The term Linked Data refers to a set of best practices for publishing and connecting structured data on the Web

- Key technologies that support Linked Data are:
  - URIs (a generic means to identify entities or concepts in the world)
  - HTTP (a simple yet universal mechanism for retrieving resources, or descriptions of resources)
  - RDF (a generic graph-based data model with which to structure and link data that describes things in the world)

Semantic Web

- The Semantic Web is made up of Linked Data; i.e. the Semantic Web is the whole, while Linked Data is the parts



LINKINGOPENDATA
W3C SWEO Community Project

What is Blackbook?

- Provides a graph analytic processing platform for Semantic Web
- Based on semantic web technologies
  -RDF, OWL, SPARQL, JENA
  -Vocabulary agnostic
- Relies on open standards and “best-of-breed” open source technologies
  -Lucene, JAAS, D2RQ, Hadoop/Map Reduce
- Leverage cloud computing technologies
  -Hadoop/Map Reduce, HBase, Solr
- Plug-and-Play, loosely–coupled architecture
- SOAP & REST interfaces, SPARQL & Linked Data endpoints
- Blackbook can run in secure environments

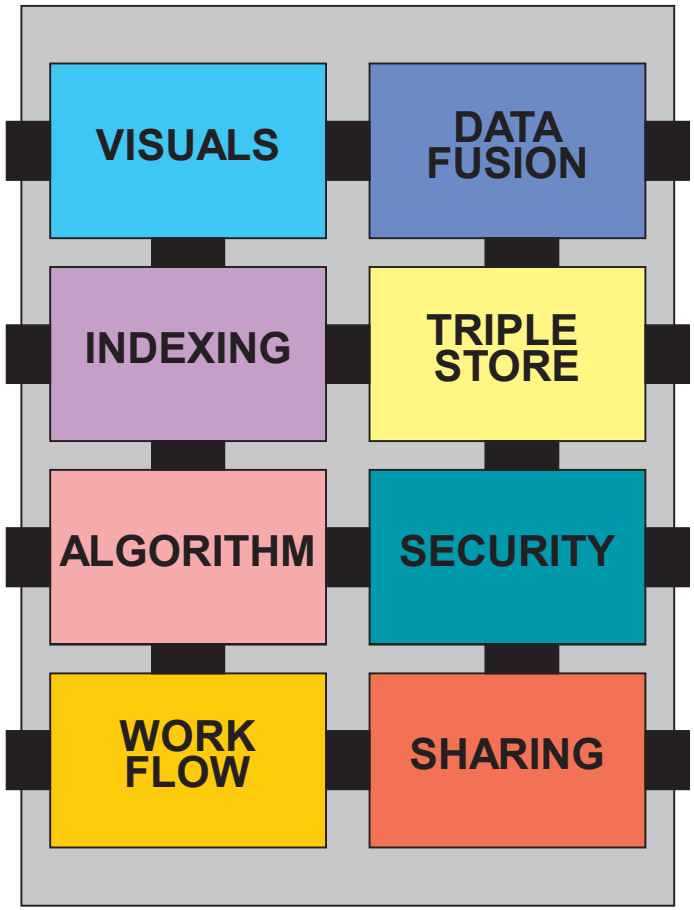
Core Components

Visualization techniques that provide the user a rich perspective on displaying datasets

Rapid search on single keywords, complex phrases, phonetic match

Apply filters, extractors, transformation algorithms

Enable automated and semi-automated control and composition



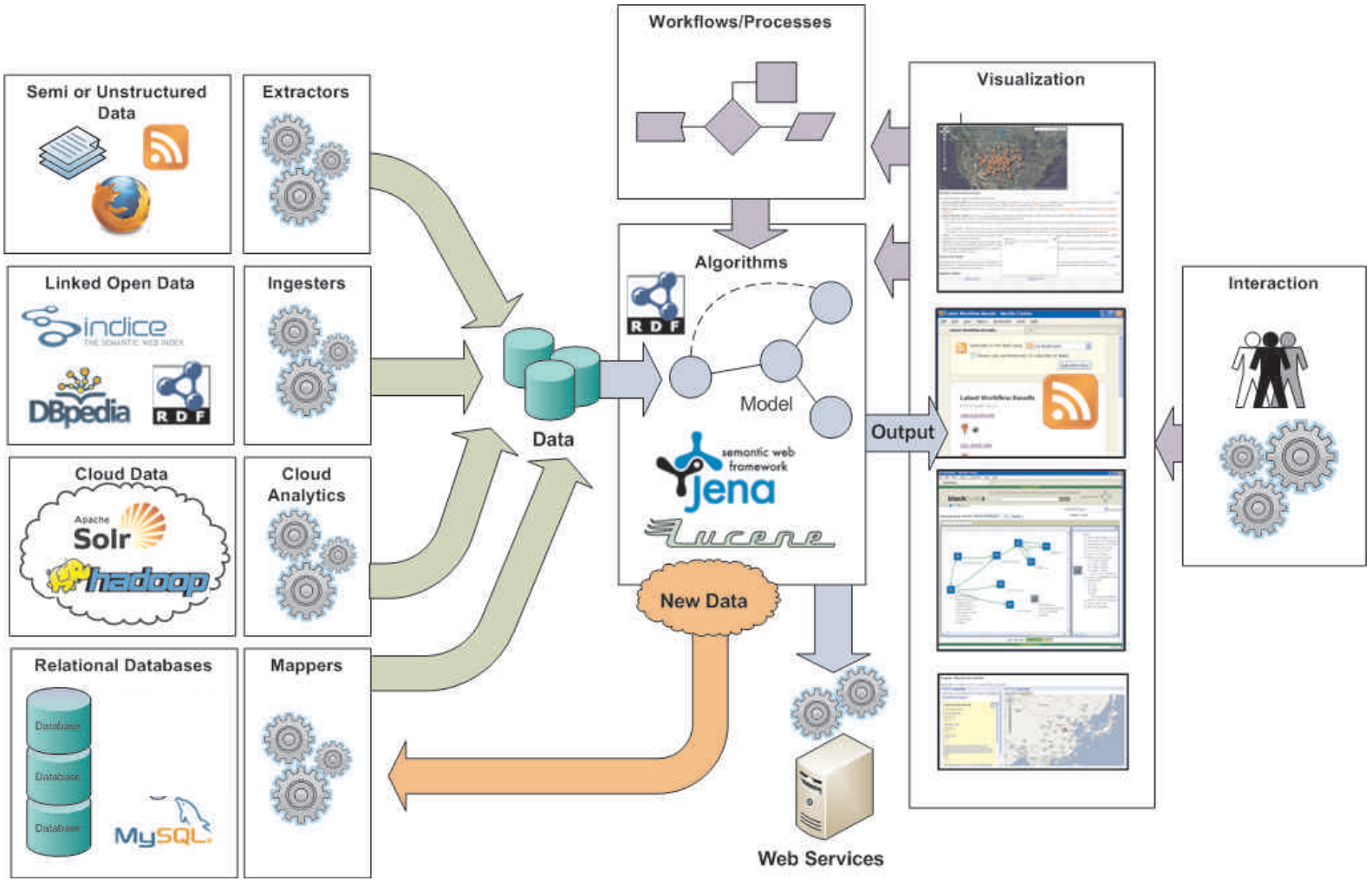
Query and merge data from many different sources, both structured and unstructured

RDF is the core data model; stores triples: Subject, Predicate, Object

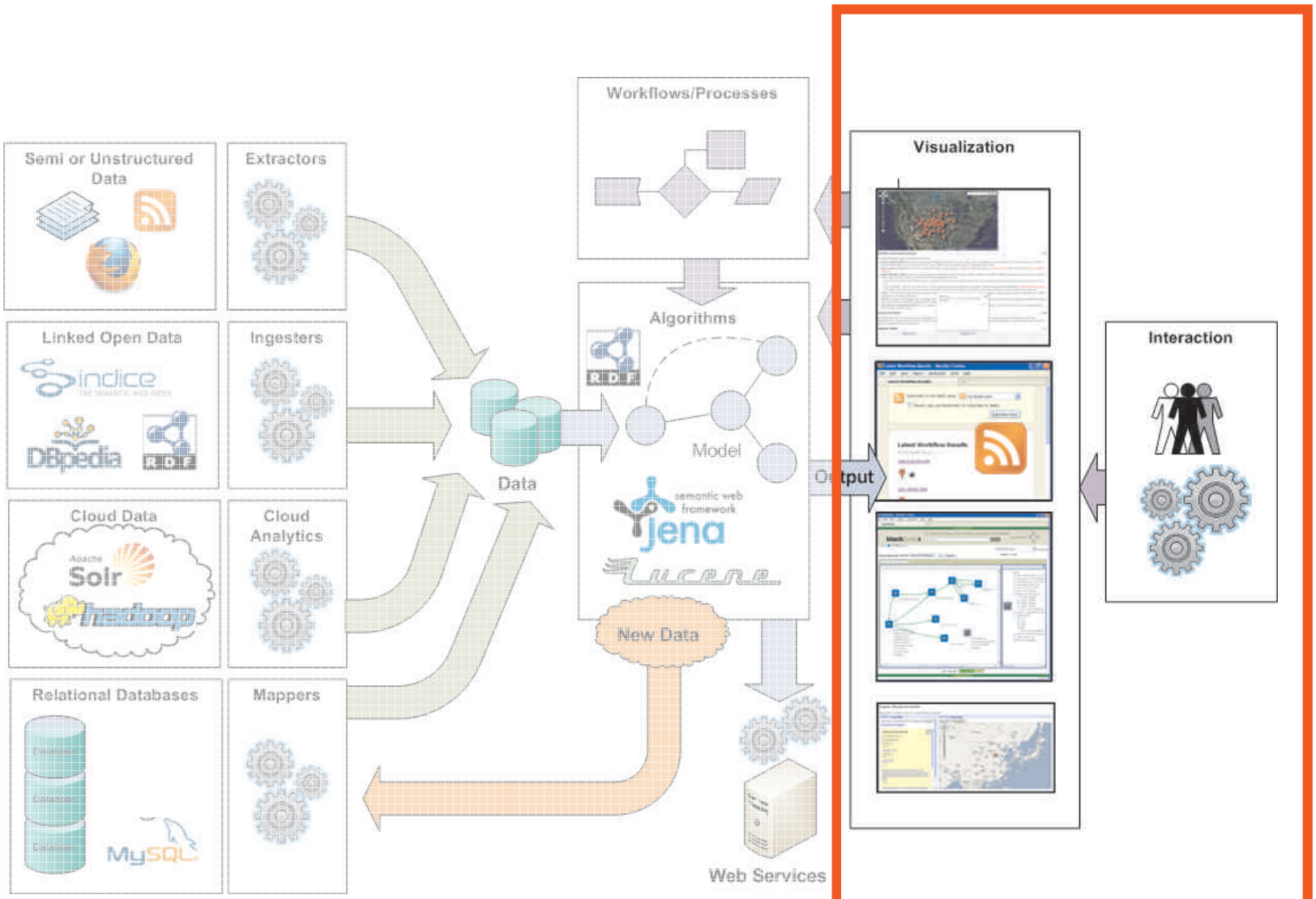
The auditing and adjudication of data as it is accessed and transformed

Discovery of web-services, and user workspaces

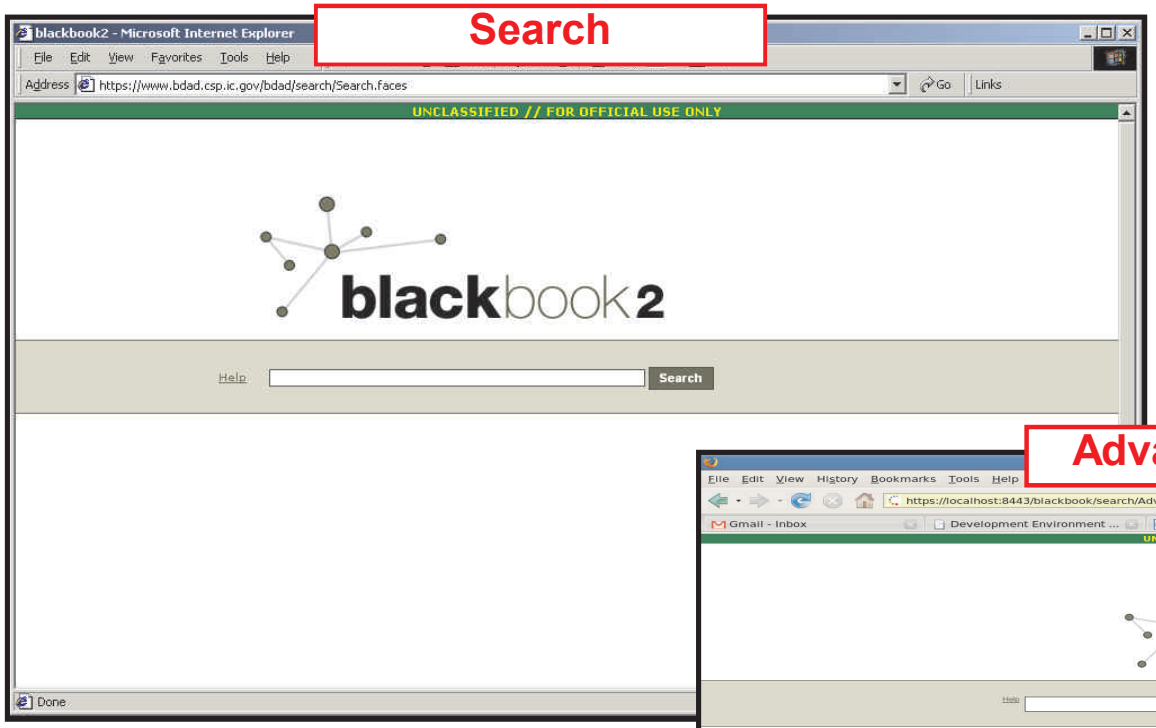
Current Capabilities

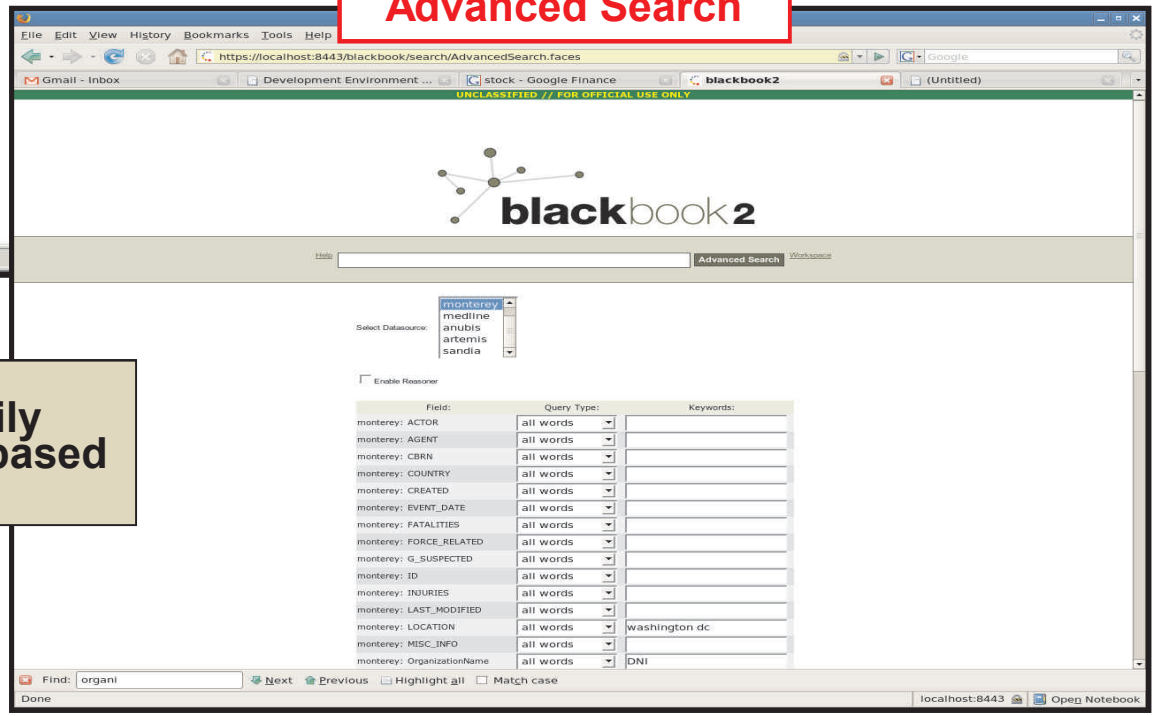


Presentation Tier



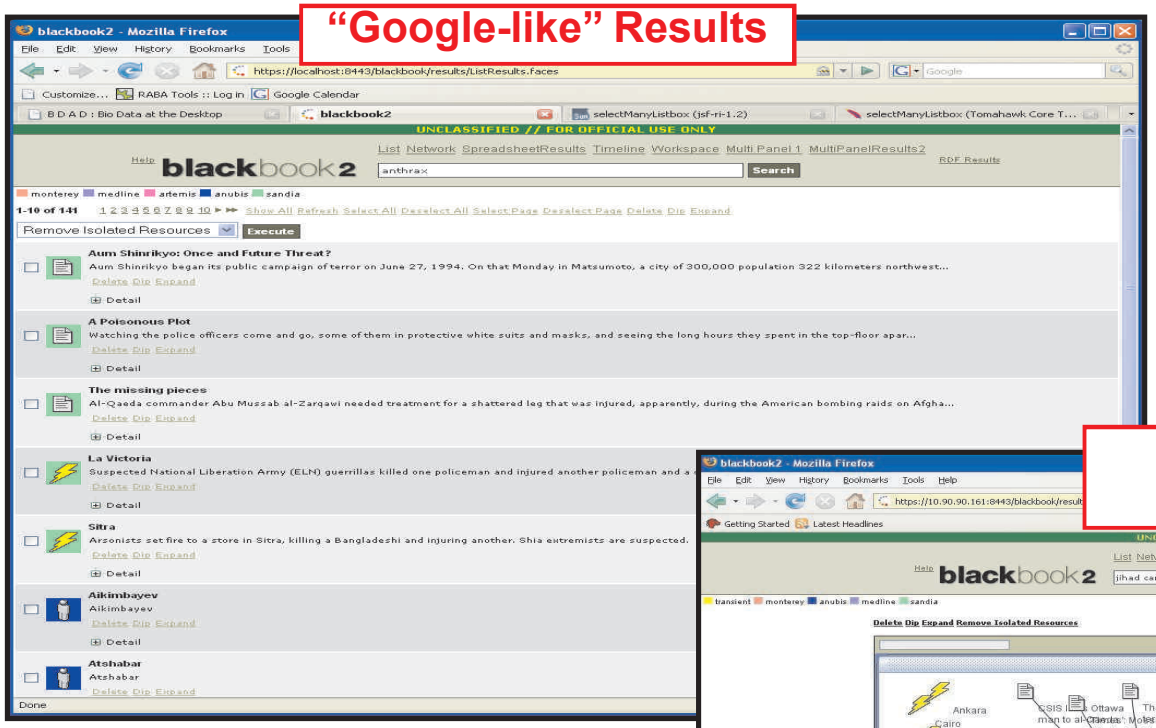
User Interface

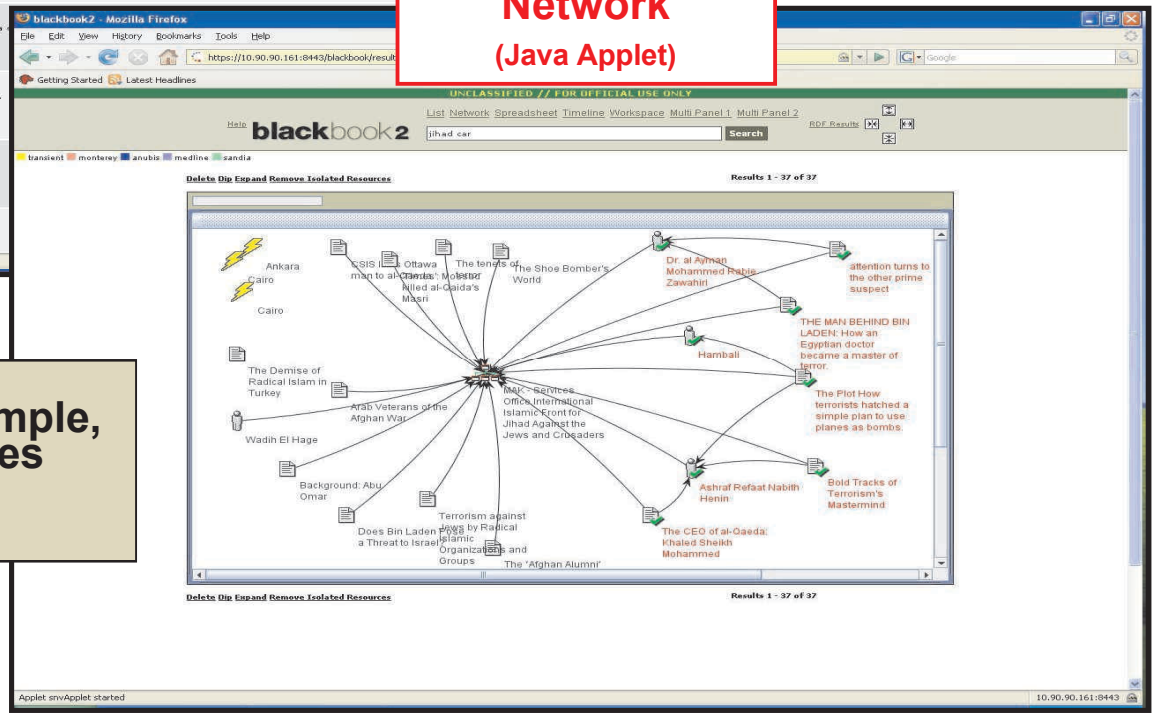




A front-end "Google-like" user interface allows analysts to easily perform keyword and attribute based searches.

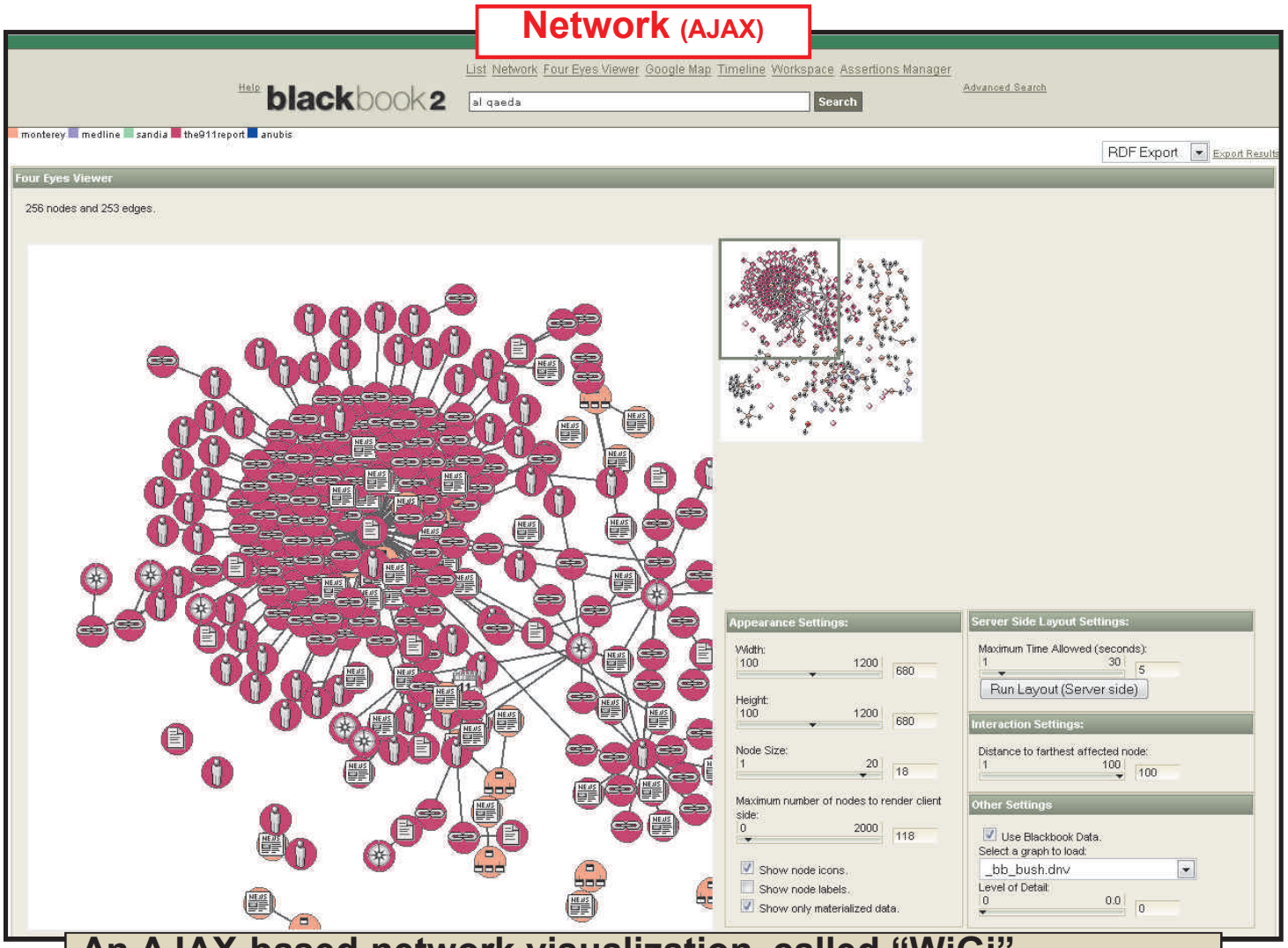
User Interface





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

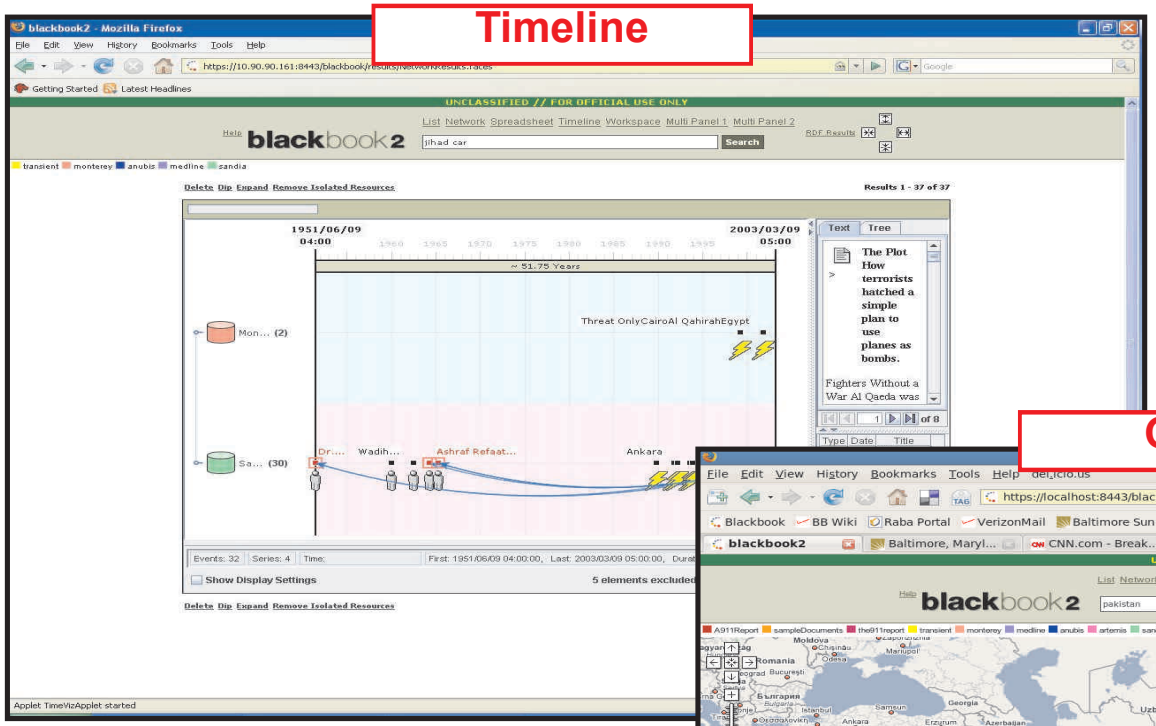
User Interface

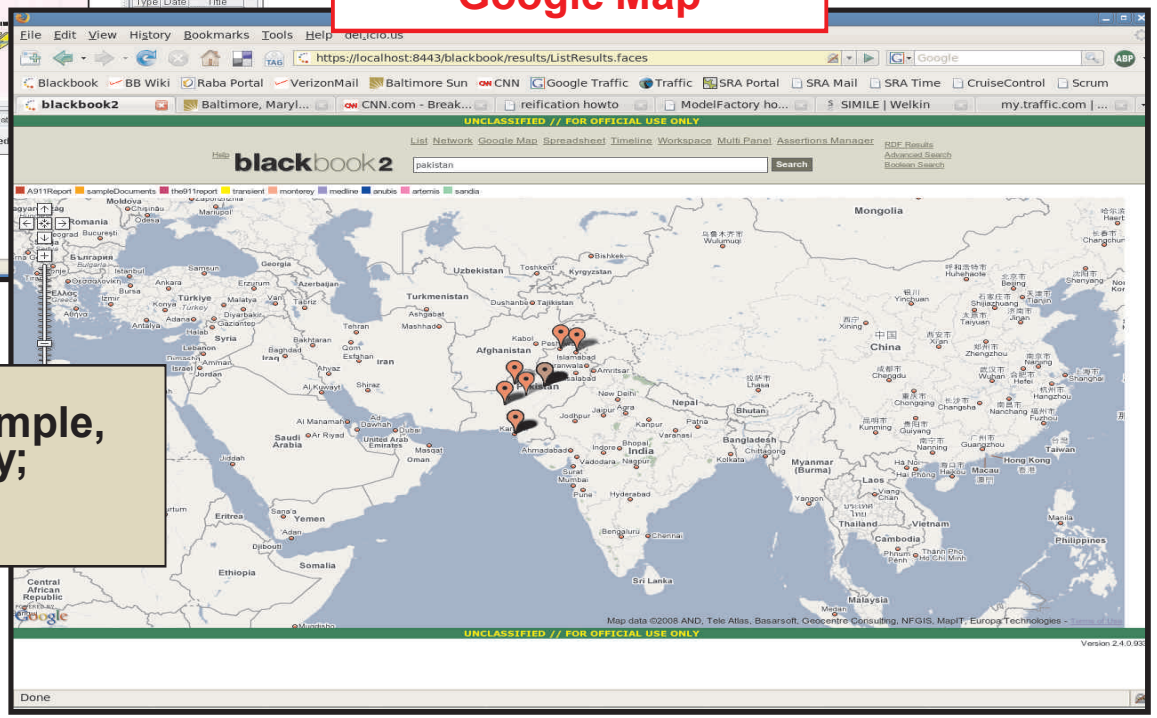


fas

An AJAX-based network visualization, called “WiGi”, optimizes client-server processing for large graphs. Planned to be released as early as Blackbook v3.0 (Nov 2009)

User Interface

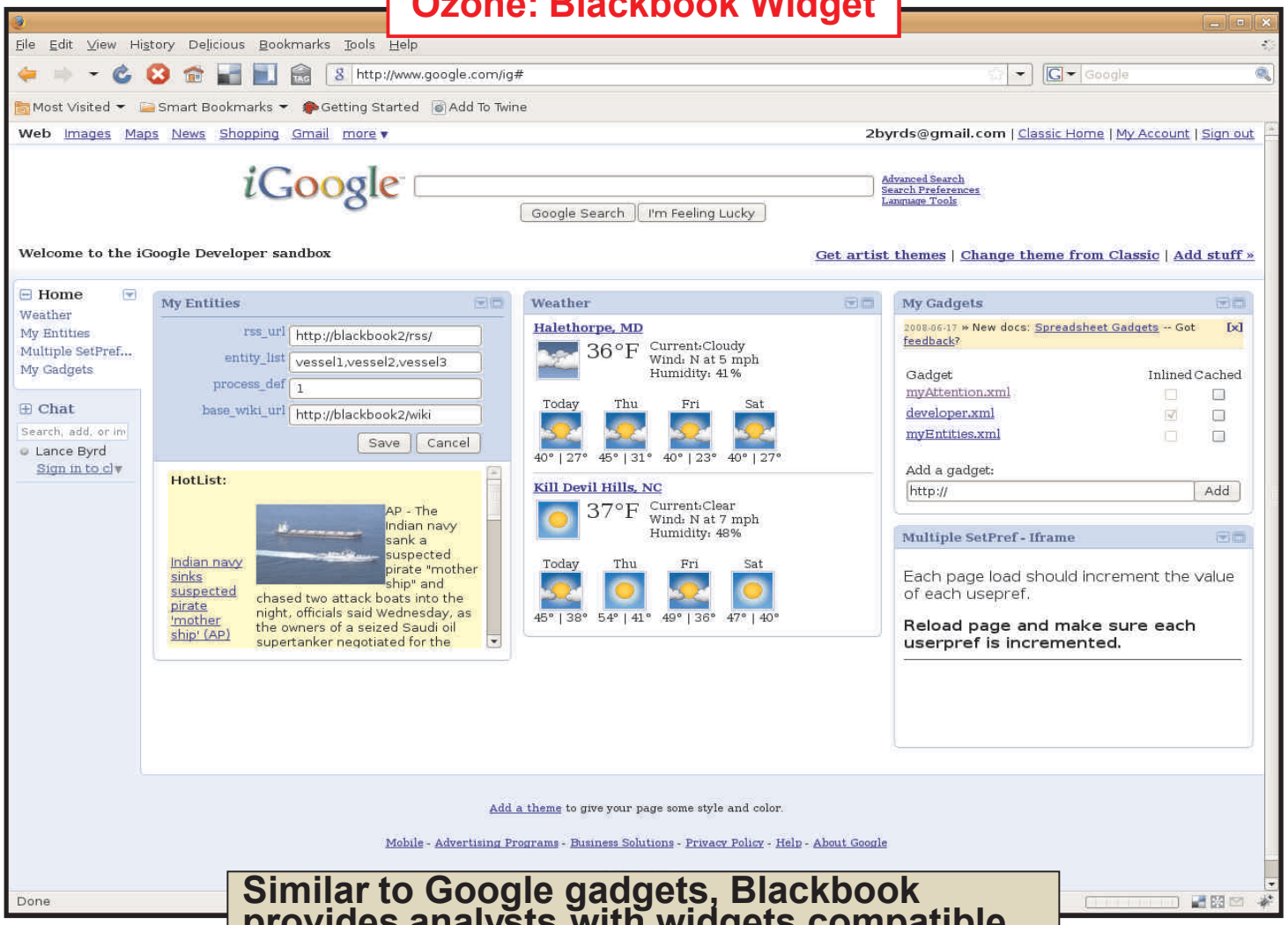




Different ways to view the same information. "Timeline", for example, displays entities chronologically; "Google Map" displays entities geospatially.

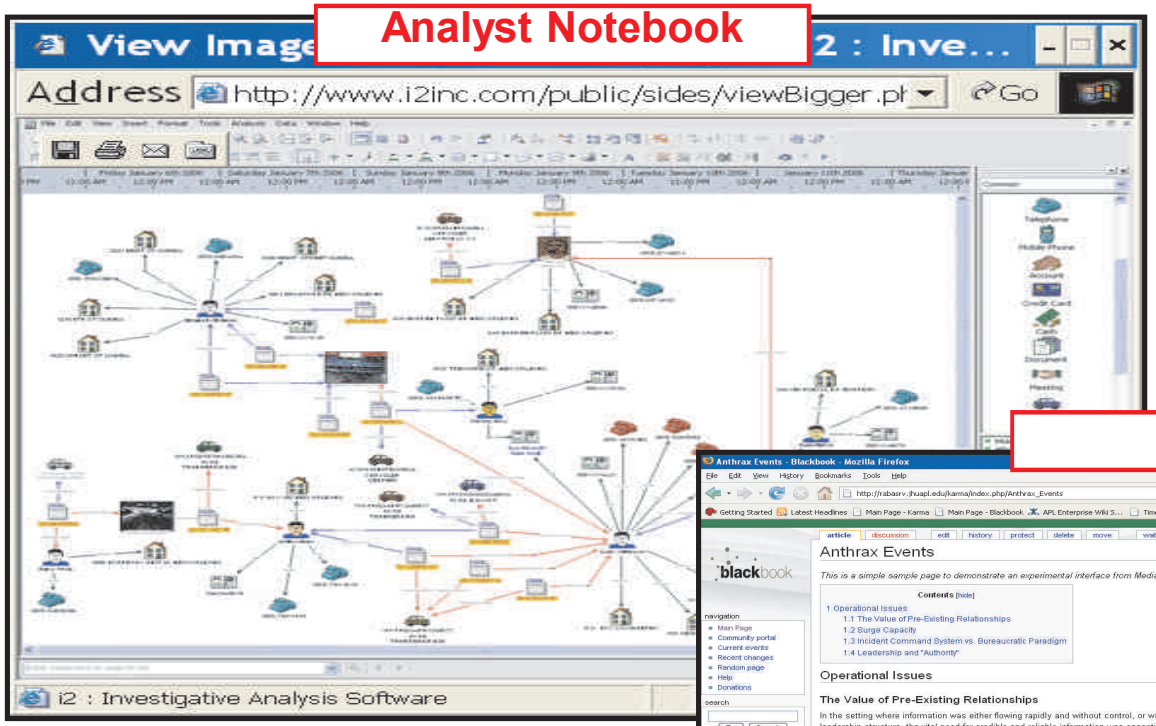
User Interface

Ozone: Blackbook Widget



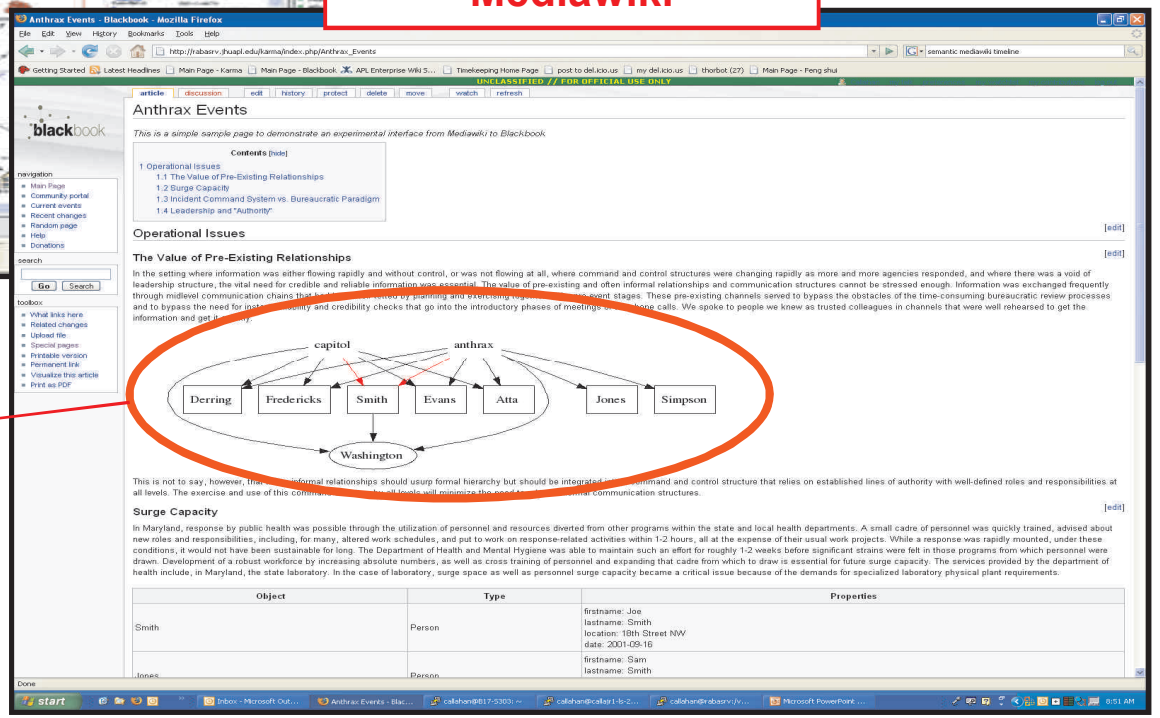
Similar to Google gadgets, Blackbook provides analysts with widgets compatible with the Ozone (an iGoogle-like) framework.

User Interface



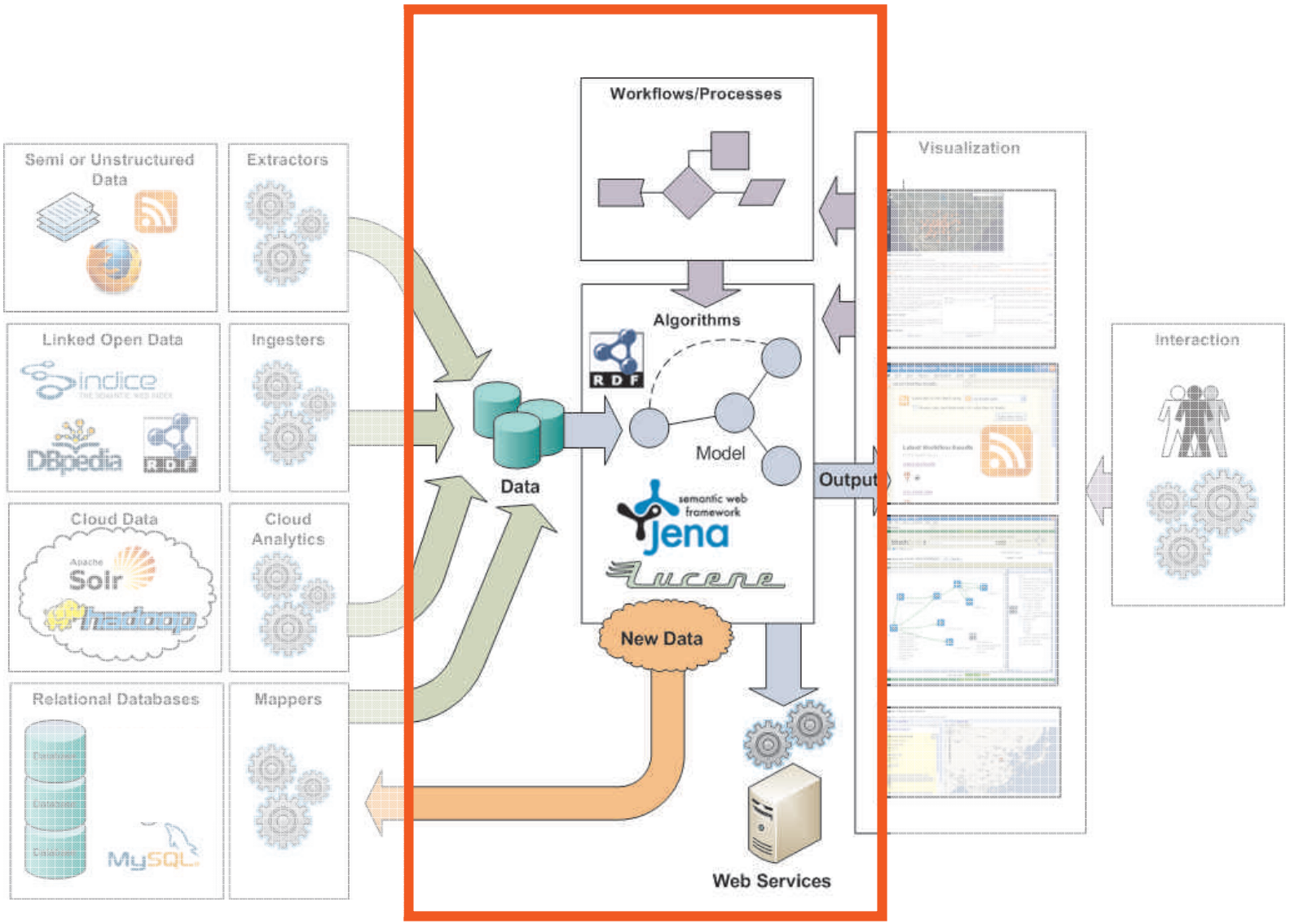
Analyst Notebook

Mediawiki

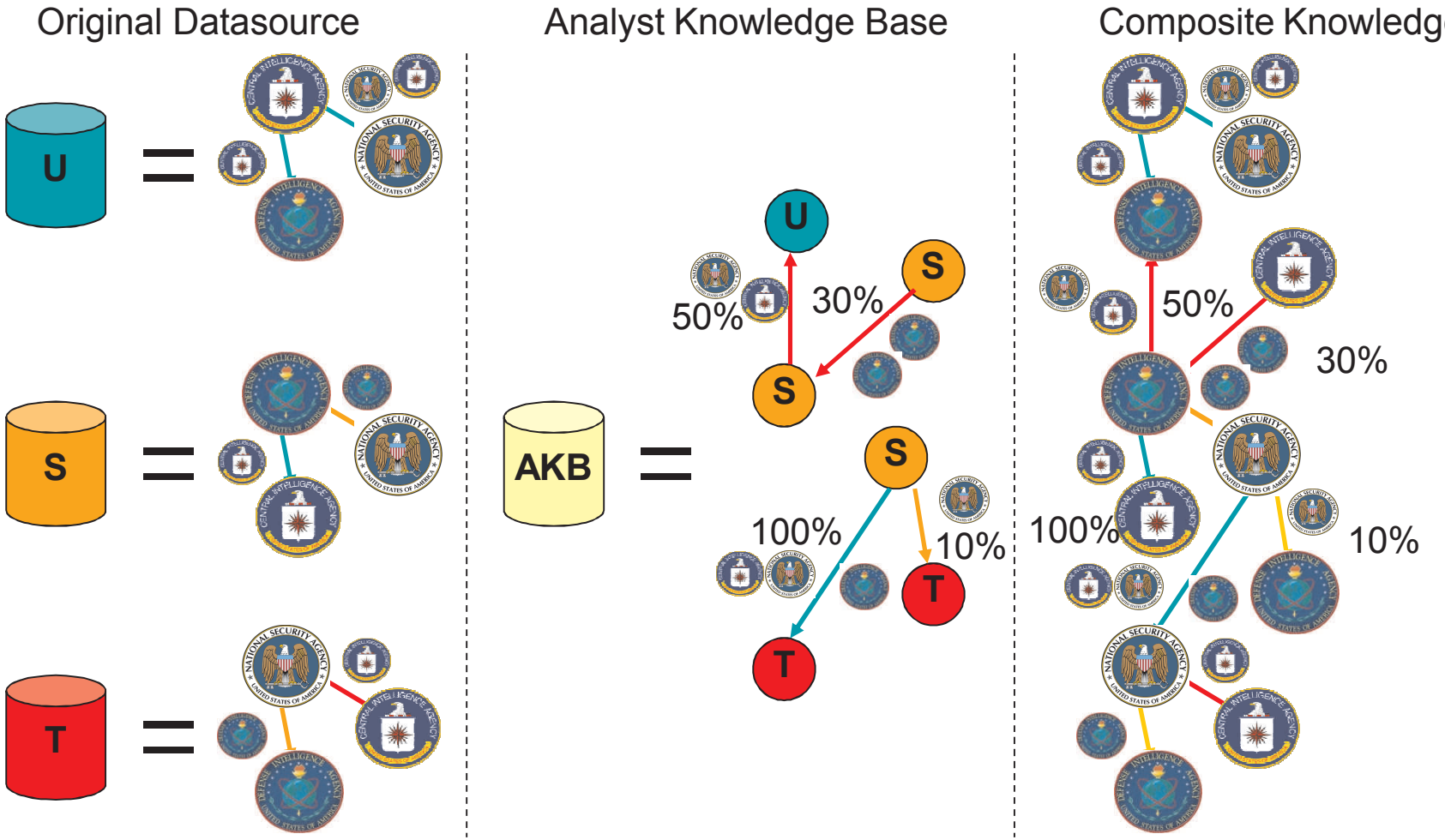


Blackbook is developing a framework called "Aqueduct", allowing interoperability between ozone widgets and wikis.

Middle Tier


Semi or Unstructured Data
Extractors
Linked Open Data
indice
DBpedia
Ingesters
Cloud Data
Apache Solr
hadoop
Cloud Analytics
Relational Databases
MySQL
Mappers
Workflows/Processes
Algorithms
RDF
Model
Data
semantic web framework
jena
lucene
New Data
Web Services
Output
Visualization
Interaction

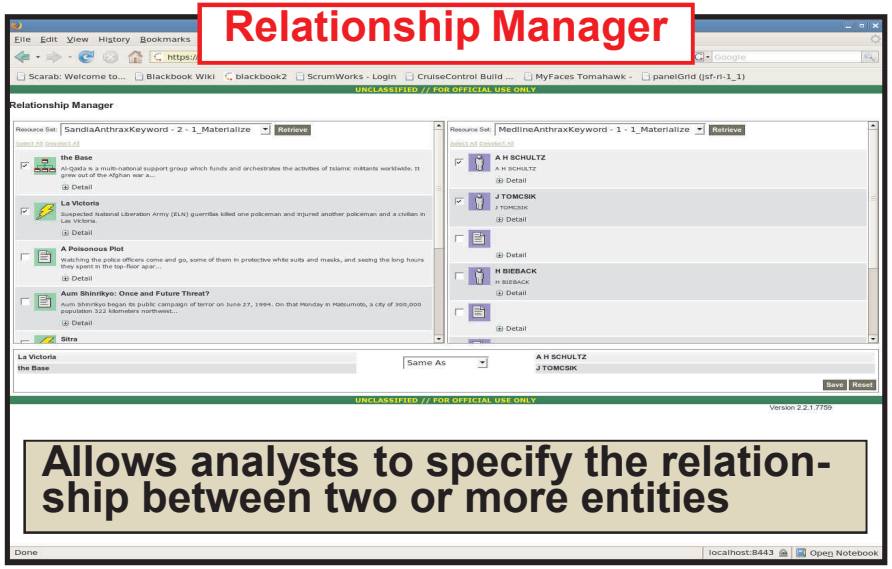
Security, Confidence, Affiliation



Blackbook uses reification for classification markings, confidence values, and affiliation. Original datasources are read-only, AKB's are read-write.

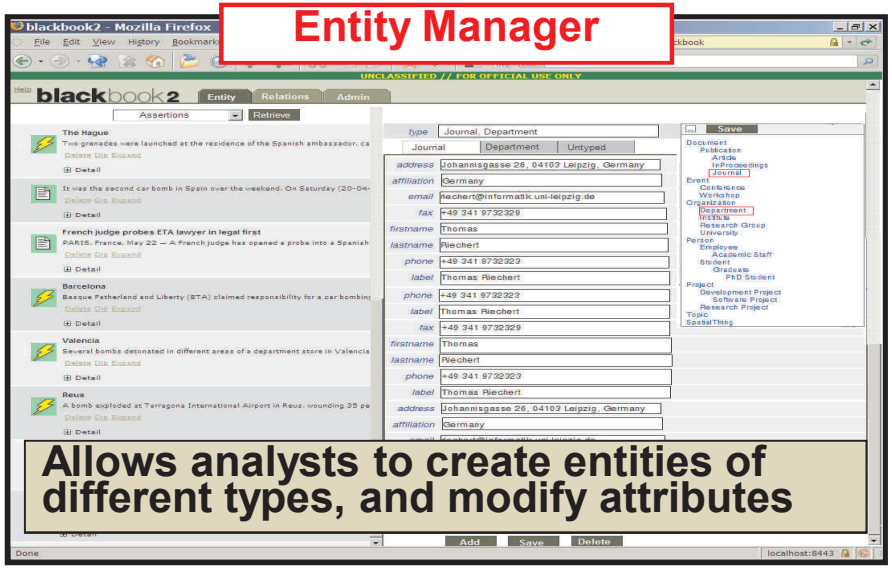
User Interface

Relationship Manager



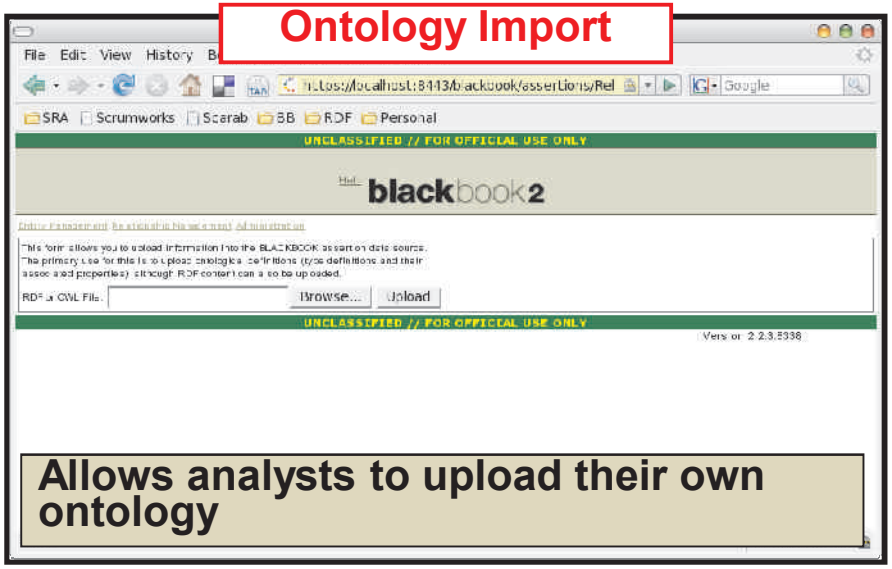
Allows analysts to specify the relation-ship between two or more entities

Entity Manager



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

Ontology Import

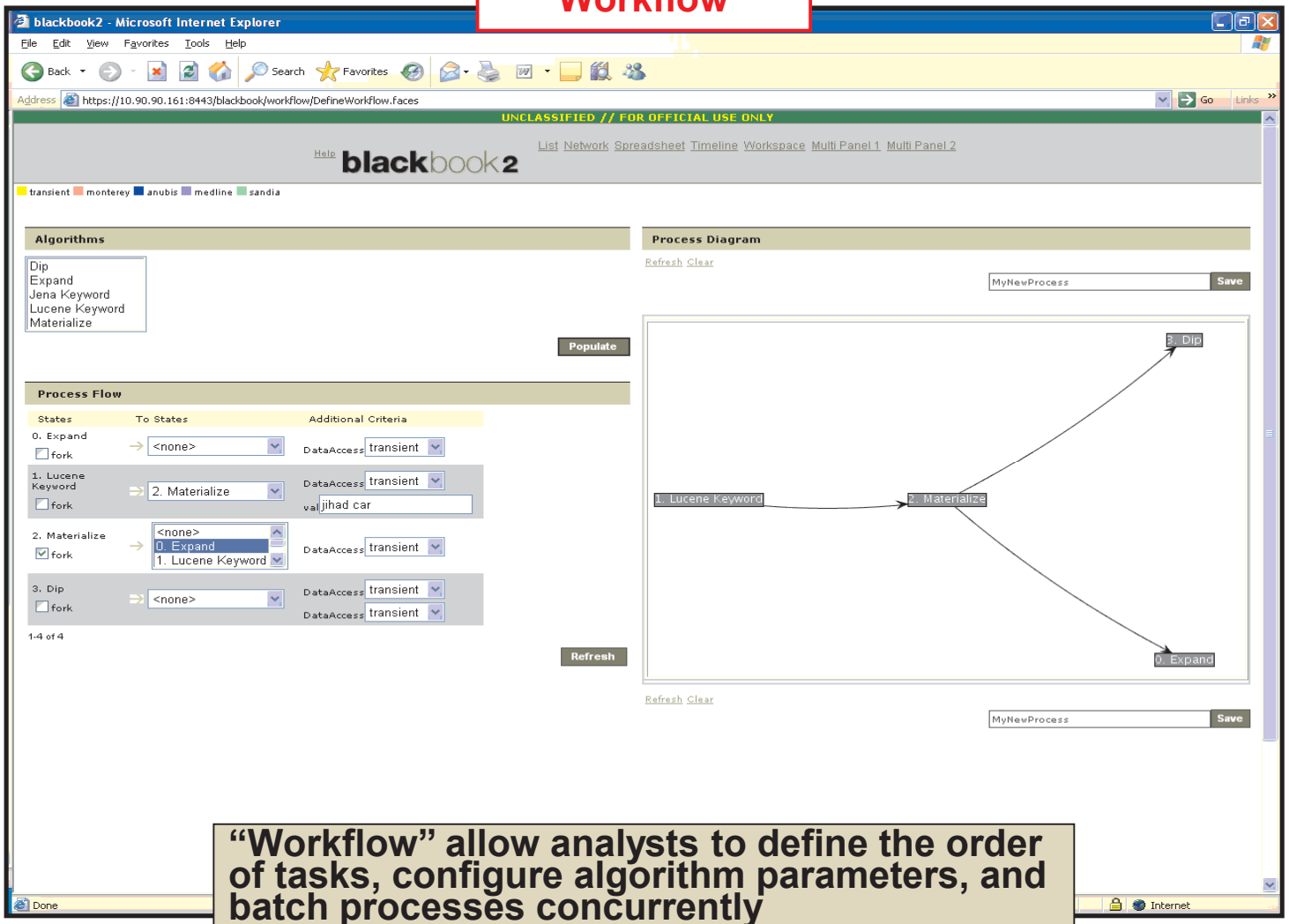


Allows analysts to upload their own ontology

fas

User Interface

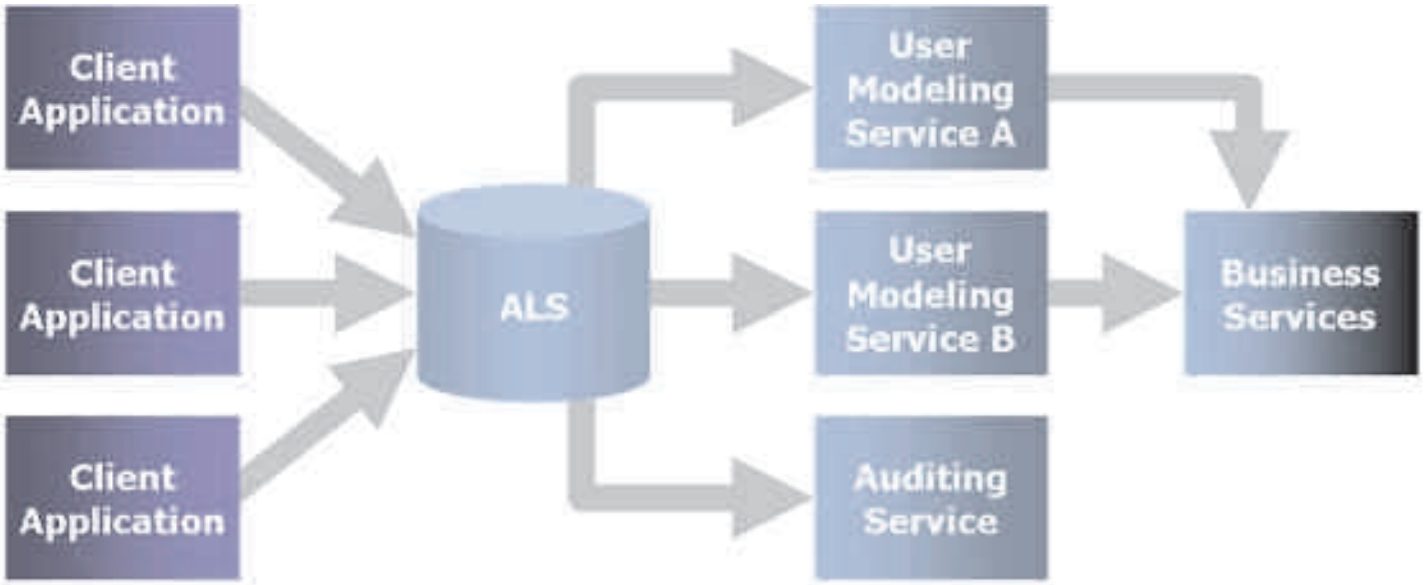
Workflow



fas

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

Analysis Log Service



Client Applications generate ALEs as users interact with the various applications.

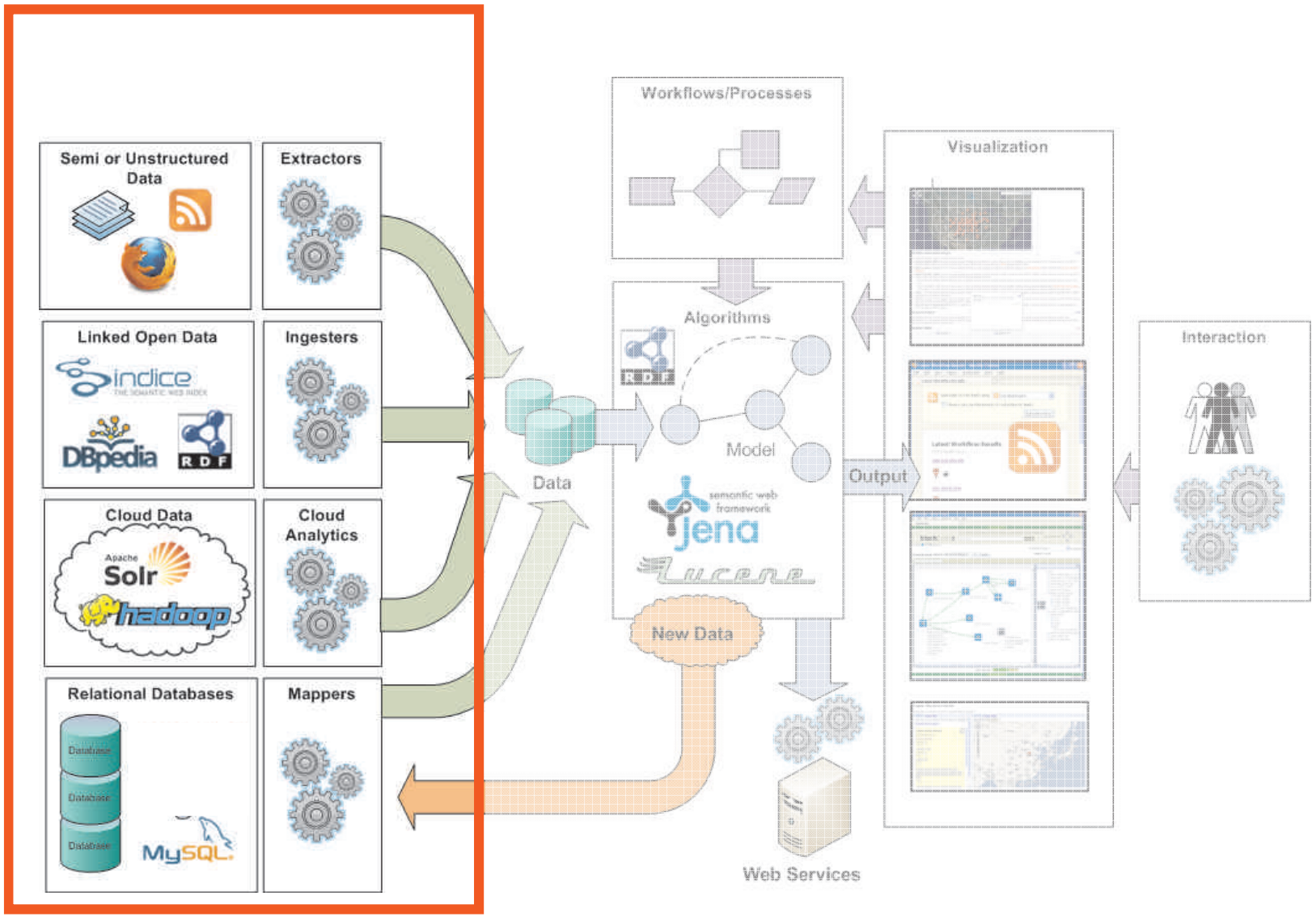
The ALEs are transmitted to the ALS.

The ALS stores the ALEs received from the client applications.

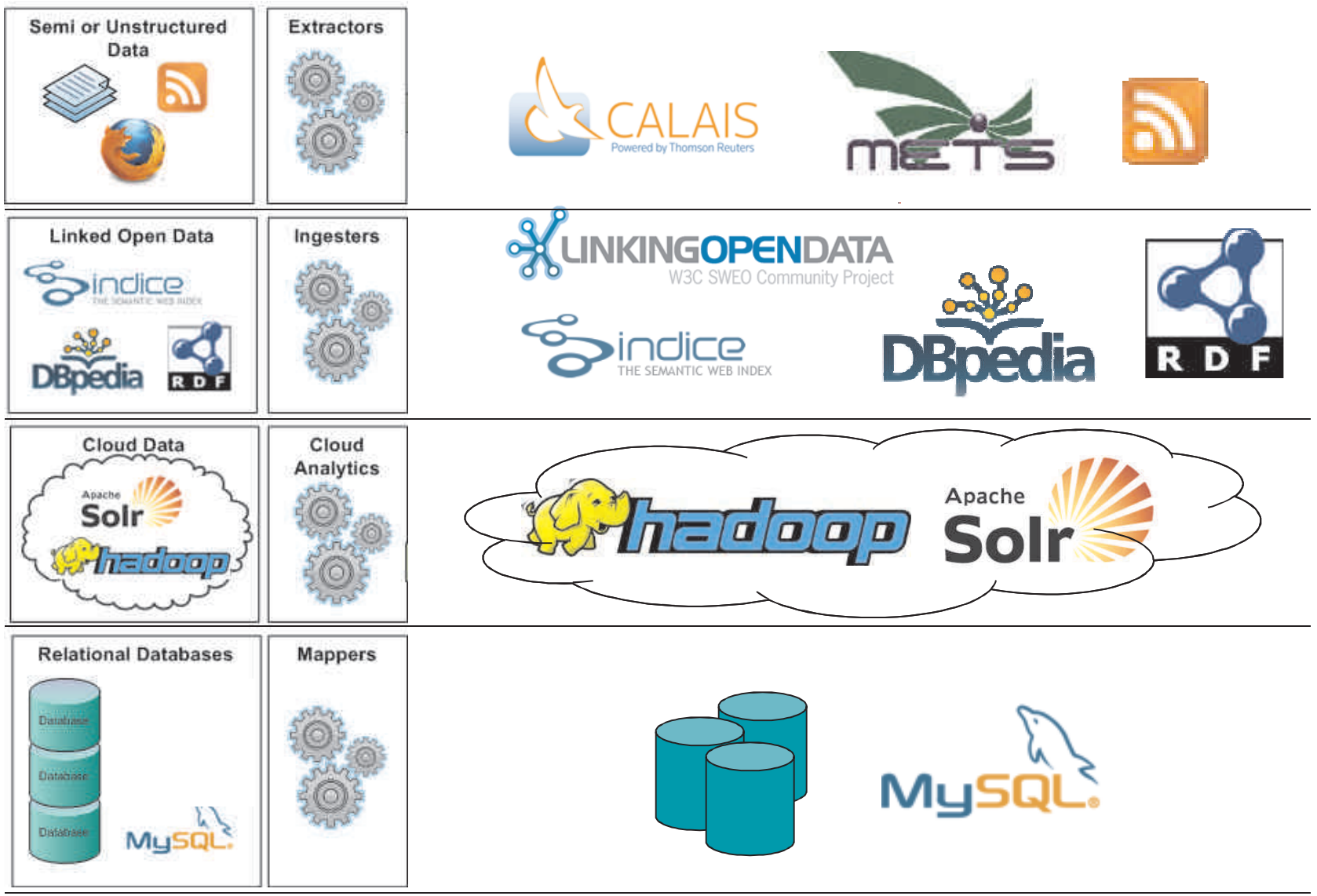
Services interested in using ALEs can query the ALS for ALEs.

Other services can consume the results of the user modeling services for their own purposes.

Data Tier


Semi or Unstructured Data
Extractors
Linked Open Data
indice
DBpedia
RDF
Ingesters
Cloud Data
Apache Solr
hadoop
Cloud Analytics
Relational Databases
Database
Database
Database
MySQL
Mappers
Data
Workflows/Processes
Algorithms
Model
semantic web framework
jena
lucene
New Data
Web Services
Visualization
Output
Interaction

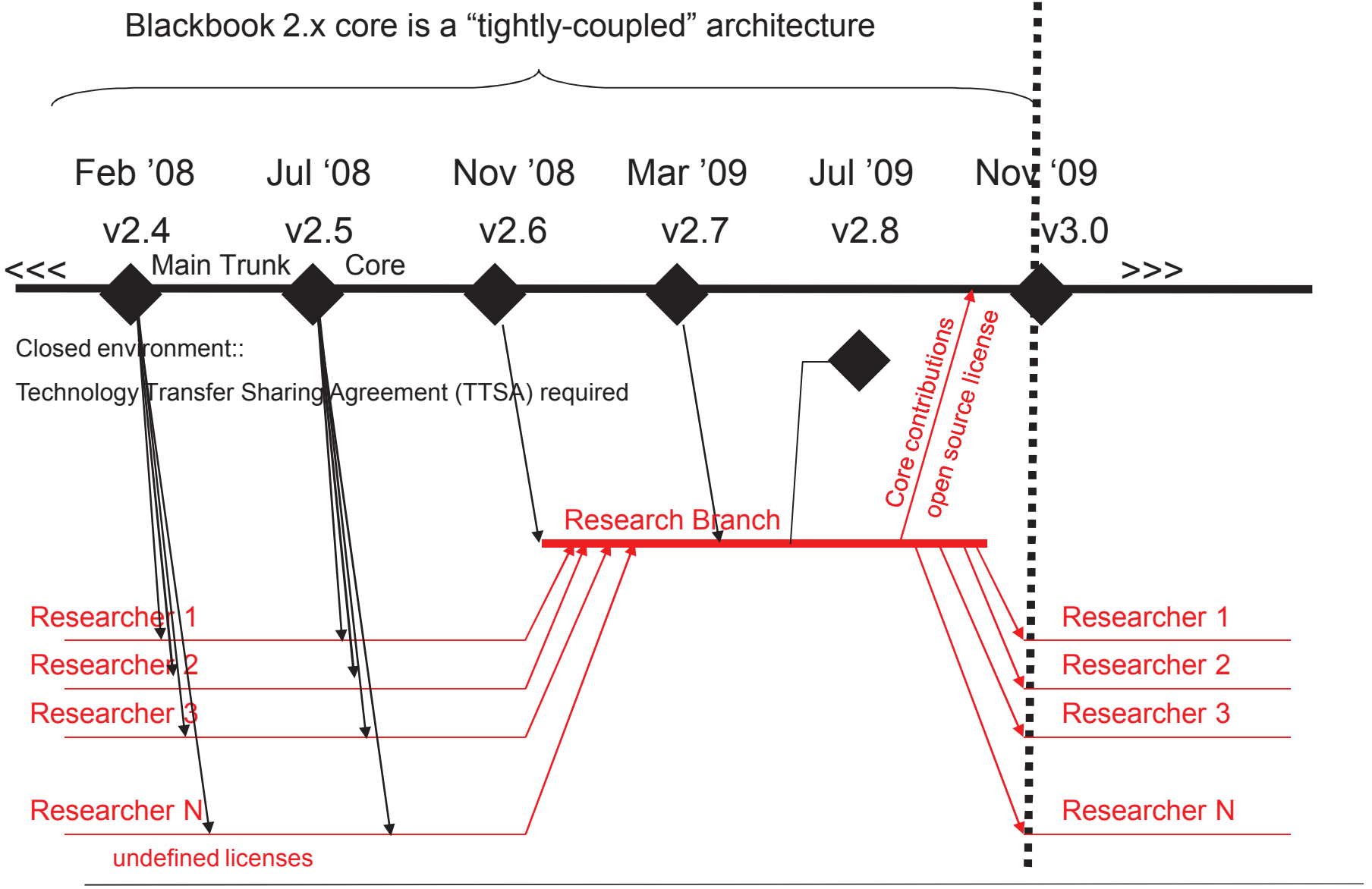
Data Integration Points



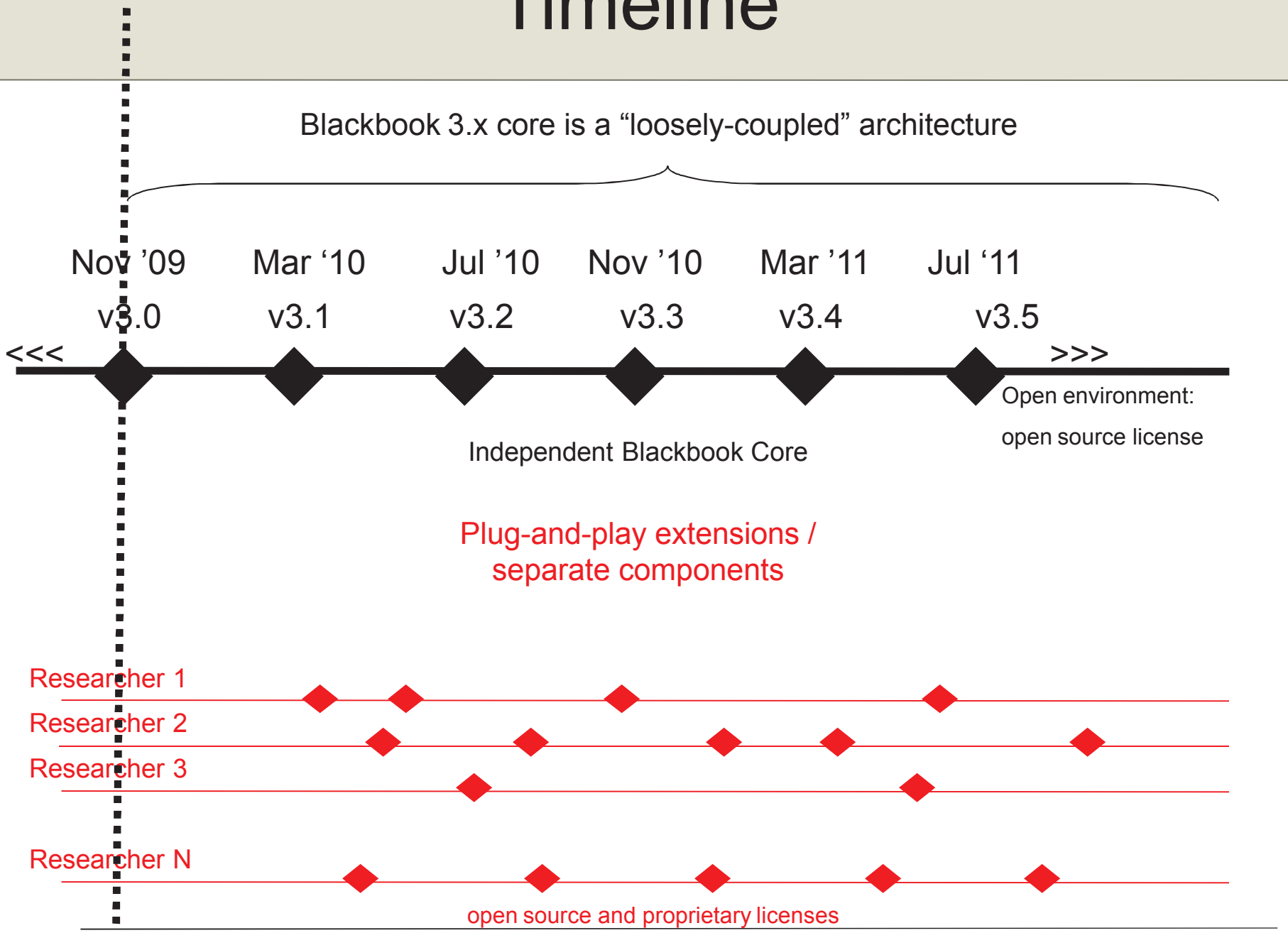
Future Capabilities

- Blackbook v3.0
  - Transition to a loosely-coupled architecture
  - Improve scalability allowing handling of large graphs
  - Implement secure SPARQL and Linked Data endpoints
  - Replace Java Applets views with AJAX-based WiGi and Simile
  - Interface to an entity extraction service (METS, Open Calais)
- Blackbook v3.1
  - Implement OSGI technology for algorithm “hot-deployment”
  - Demonstrate the mobile analytic concept
  - Improve visualization with rich interface
- Blackbook v3.2
  - Peer-to-Peer connectivity for Blackbook platforms

Timeline

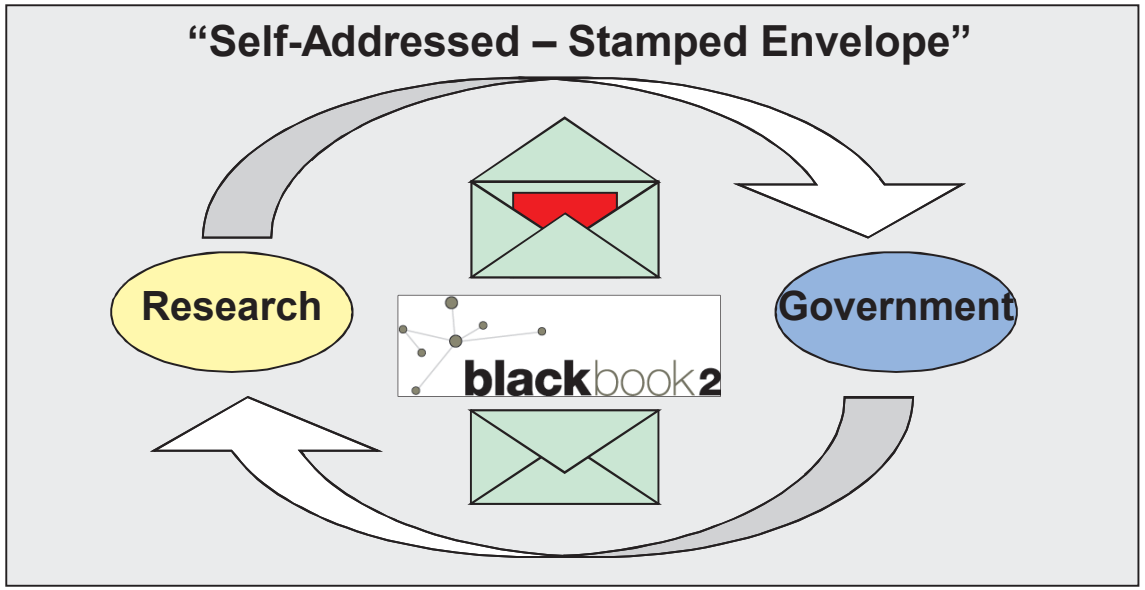

Blackbook 2.x core is a “tightly-coupled” architecture
Feb ’08
v2.4
Jul ‘08
v2.5
Nov ’08
v2.6
Mar ’09
v2.7
Jul ’09
v2.8
Nov ‘09
v3.0
<<<
Main Trunk
Core
>>>
Closed environment::
Technology Transfer Sharing Agreement (TTSA) required
Core contributions
open source license
Research Branch
Researcher 1
Researcher 2
Researcher 3
Researcher N
undefined licenses
Researcher 1
Researcher 2
Researcher 3
Researcher N

Timeline


Blackbook 3.x core is a “loosely-coupled” architecture
Nov ’09
v3.0
Mar ‘10
v3.1
Jul ’10
v3.2
Nov ’10
v3.3
Mar ’11
v3.4
Jul ‘11
v3.5
<<<
>>>
Open environment:
open source license
Independent Blackbook Core
Plug-and-play extensions /
separate components
Researcher 1
Researcher 2
Researcher 3
Researcher N
open source and proprietary licenses

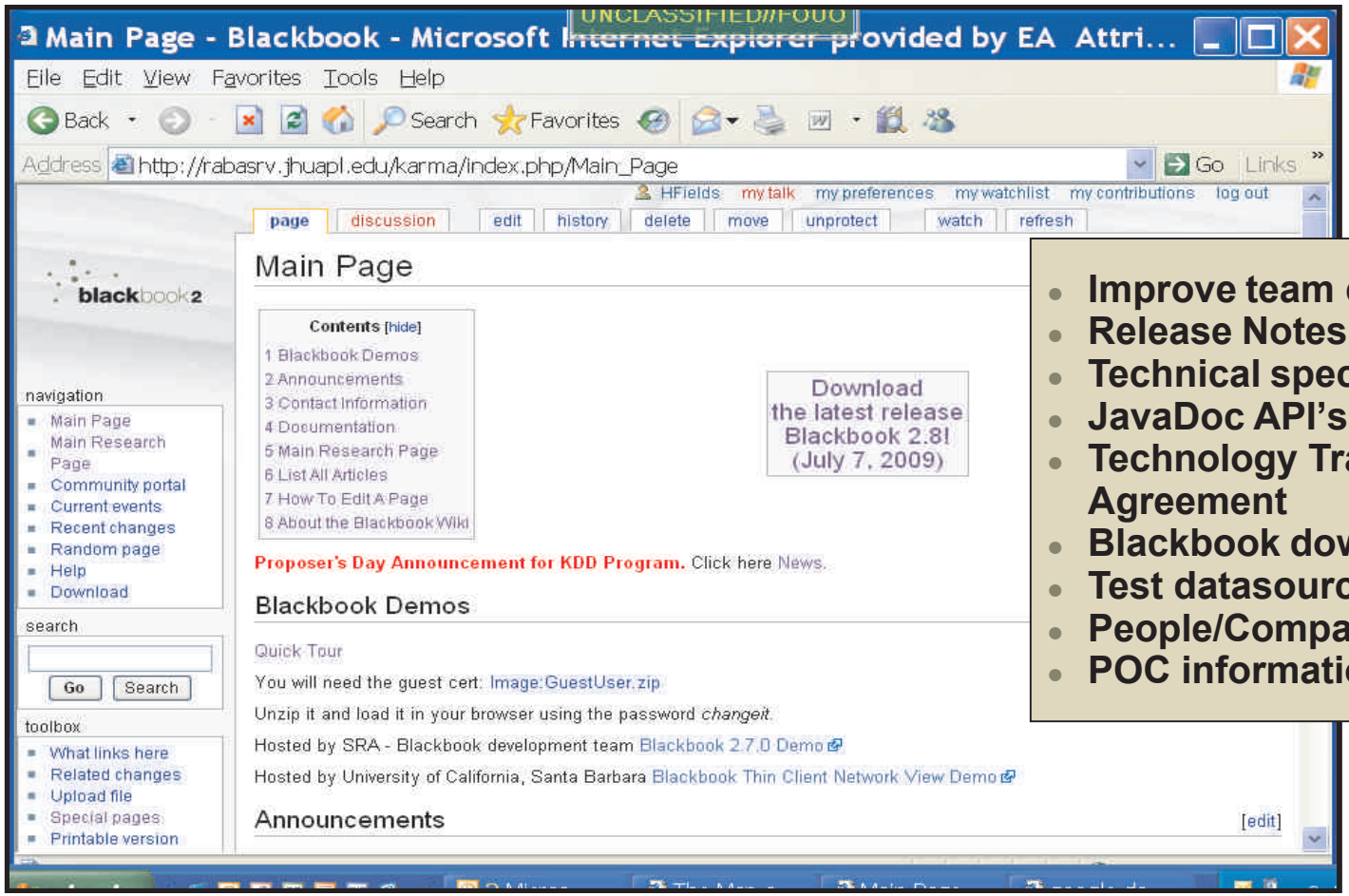
Technology Transfer

- Knowledge Discovery and Dissemination (KDD) program
  - Led by Dr Art Becker
- Blackbook provides a common integration framework for technology transfer



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".

Blackbook Wiki



- **Improve team collaboration**
- **Release Notes**
- **Technical specs, documentation**
- **JavaDoc API's**
- **Technology Transfer Sharing Agreement**
- **Blackbook download access**
- **Test datasources**
- **People/Company list**
- **POC information**

Blackbook wiki can be accessed from the internet:
http://blackbook.jhuapl.edu

Process: Blackbook wiki account

Step 1:

Requester sends an email to the KDD Program Management Office (PMO), with the following information:

- First Name
- Last Name
- Affiliation (Company Name, Academic Institution, Government Agency)
- Work Phone
- Unclassified email address

- KDD PMO email: dni-iarpa-baa-09-10@ugov.gov

Process: Blackbook wiki account

Step 2:

KDD PMO will verify that a valid Technology Transfer Sharing Agreement (TTSA) form is on file for ALL companies and academic institutions. A TTSA is not required for government agencies.

- Blackbook software is not open source licensed – yet!
- A TTSA protects government's intellectual property

If a TTSA is not on file, the KDD PMO will email a TTSA to the requester

If a TTSA is on file, then Step 5

Process: Blackbook wiki account

Step 3:

Requester has a company representative sign the TTSA

- The TTSA is an agreement between the Government and the requester's company or academic institution
- The TTSA is NOT an agreement between the Government and the requester as an individual

Requester emails a signed TTSA to the KDD PMO

Process: Blackbook wiki account

Step 4:

KDD PMO will sign the TTSA and will archive

KDD PMO will email a signed copy of the TTSA to the requester

Process: Blackbook wiki account

Step 5:

KDD PMO will create a Blackbook wiki account for the requestor, as an individual

He/she may download the Blackbook software

Thank You