

blackbook2

3

Original slides created by:

Buster Fields Program Manager

Slides modified by:

Set Cruz and Lance Byrd

Contact Information

- Marathon Minds
 - Lance Byrd
 - lancebyrd@marathonminds.com
- Set Cruz
 - setcruz1@umbc.edu

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
-

Linked Data

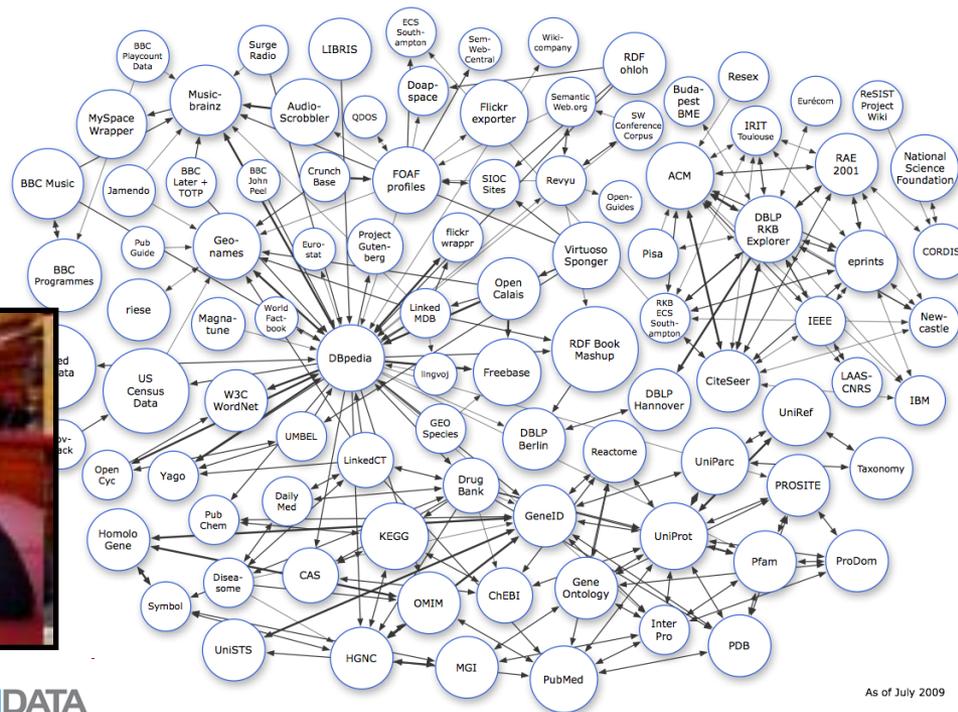
- The term Linked Data refers to a set of best practices for publishing and connecting structured data on the Web



- 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



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

VISUALS

DATA FUSION

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

Rapid search on single keywords, complex phrases, phonetic match

INDEXING

TRIPLE STORE

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

Apply filters, extractors, transformation algorithms

ALGORITHM

SECURITY

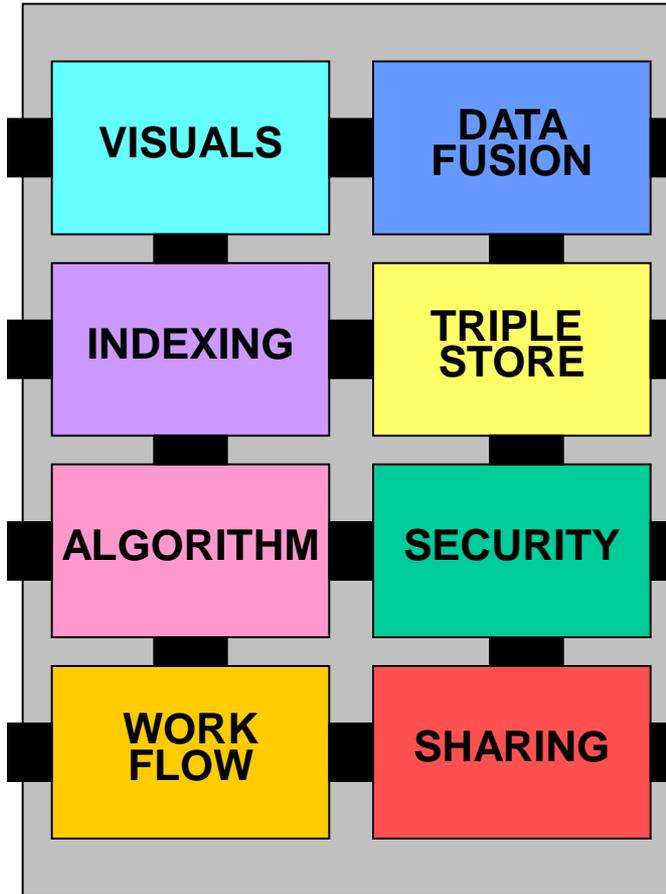
The auditing and adjudication of data as it is accessed and transformed

Enable automated and semi-automated control and composition

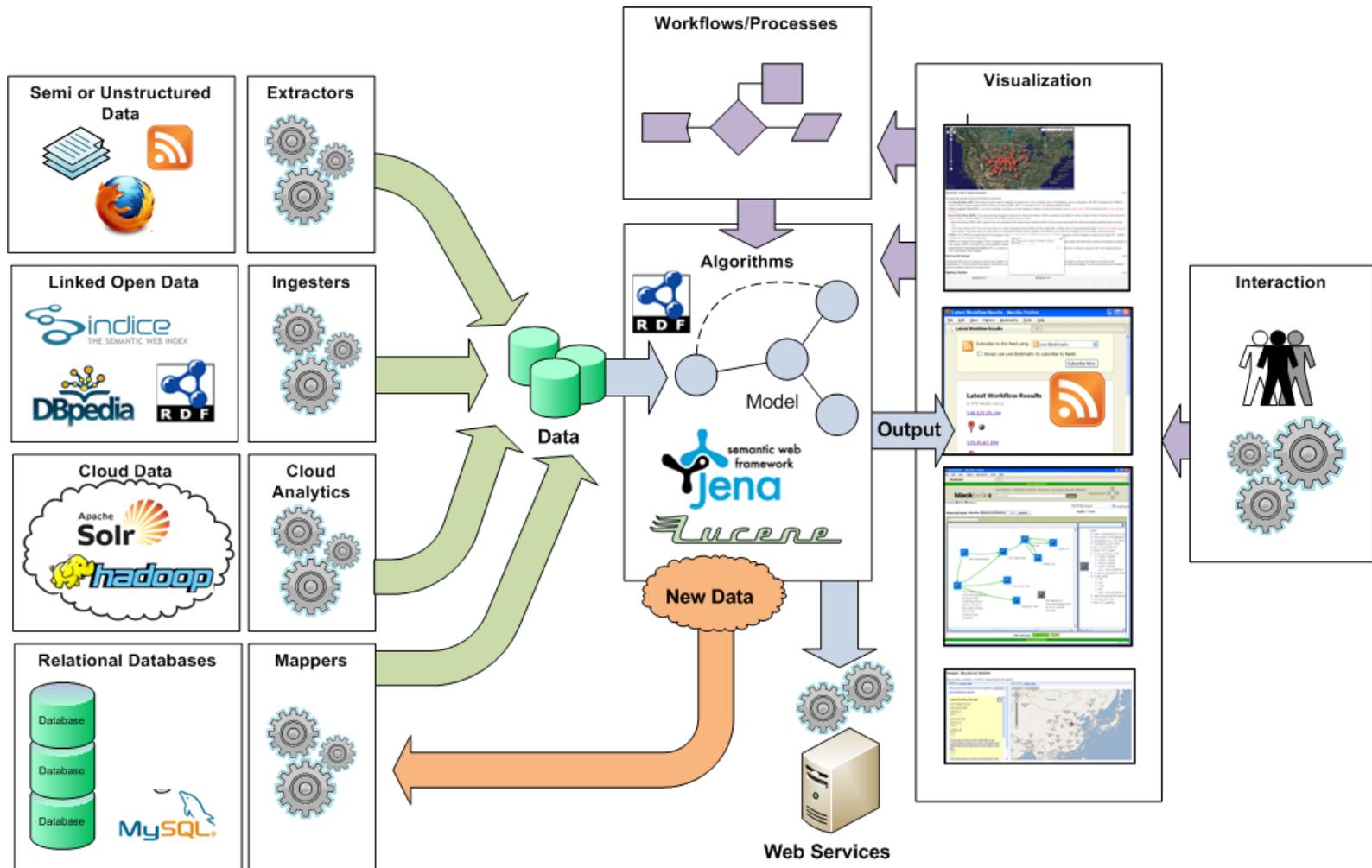
WORK FLOW

SHARING

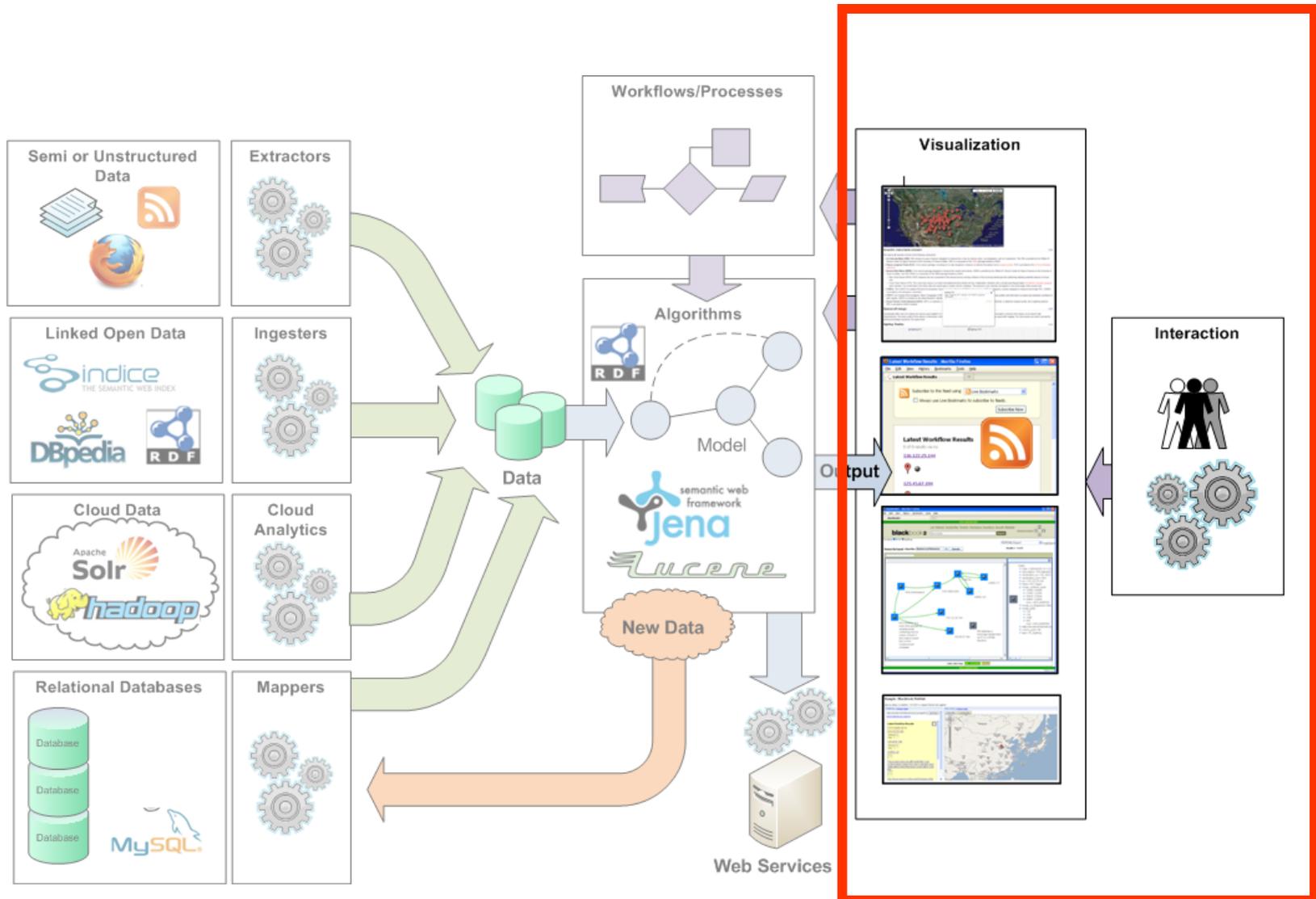
Discovery of web-services, and user workspaces



Current Capabilities

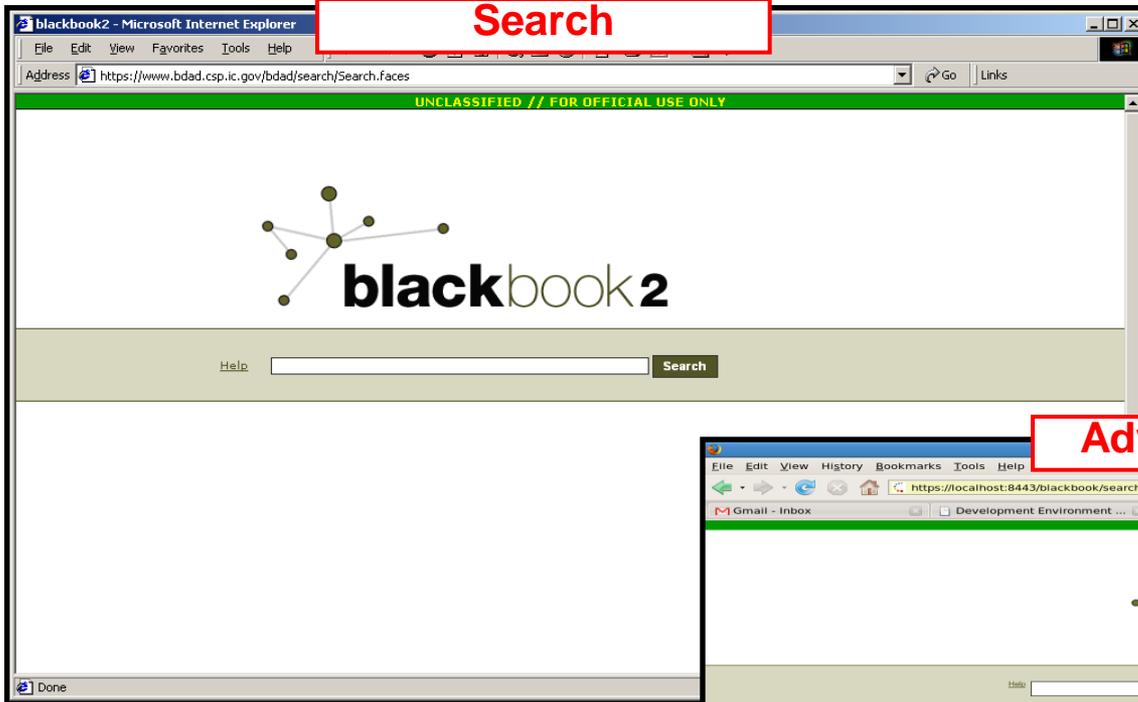


Presentation Tier

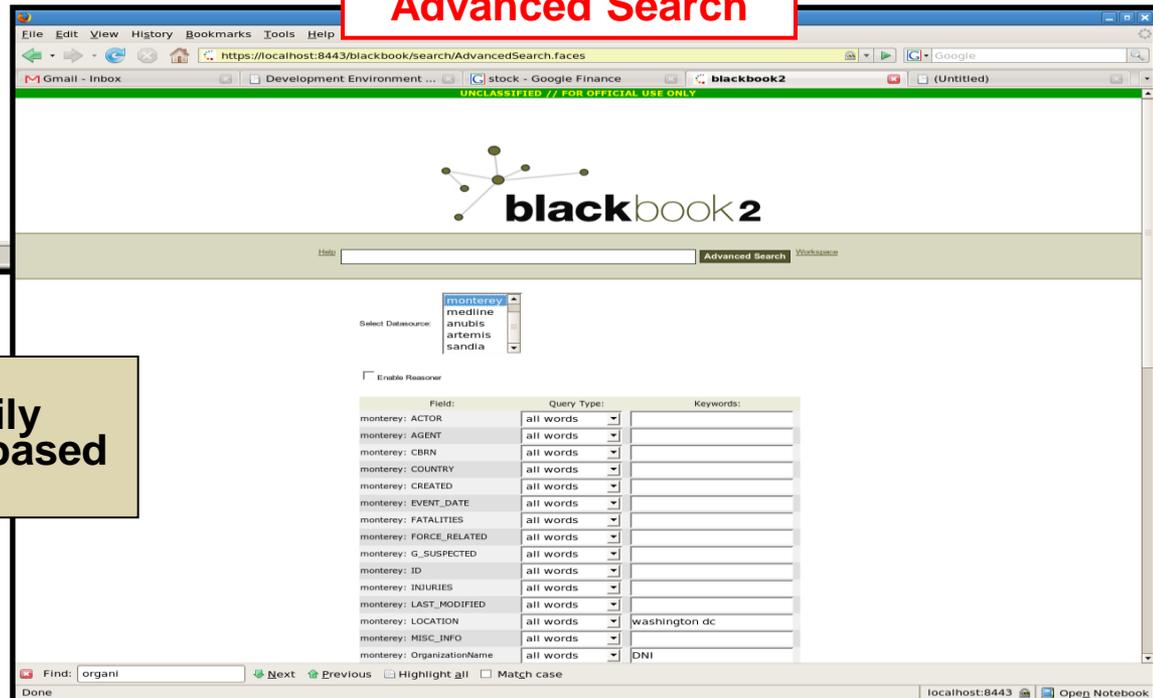


User Interface

Search



Advanced Search



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

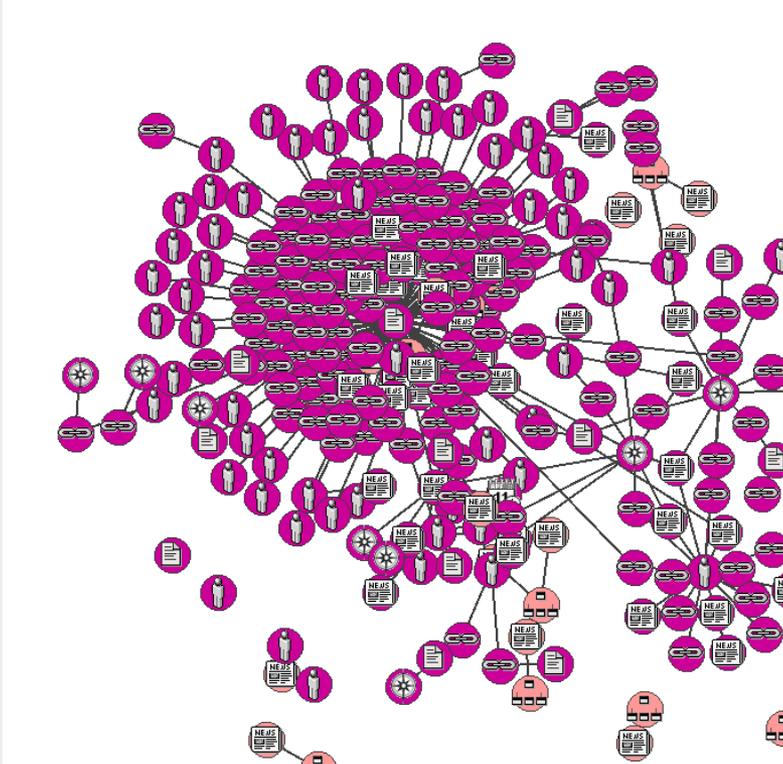
User Interface

Network (AJAX)

Help **blackbook2** List Network Four Eyes Viewer Google Map Timeline Workspace Assertions Manager Advanced Search

monterey medline sandia the911report anubis RDF Export Export Results

Four Eyes Viewer
256 nodes and 253 edges.



Appearance Settings:

Width:	100	1200	680
Height:	100	1200	680
Node Size:	1	20	18
Maximum number of nodes to render client side:	0	2000	118

Show node icons.
 Show node labels.
 Show only materialized data.

Server Side Layout Settings:

Maximum Time Allowed (seconds):
1 30 5

Interaction Settings:

Distance to farthest affected node:
1 100 100

Other Settings

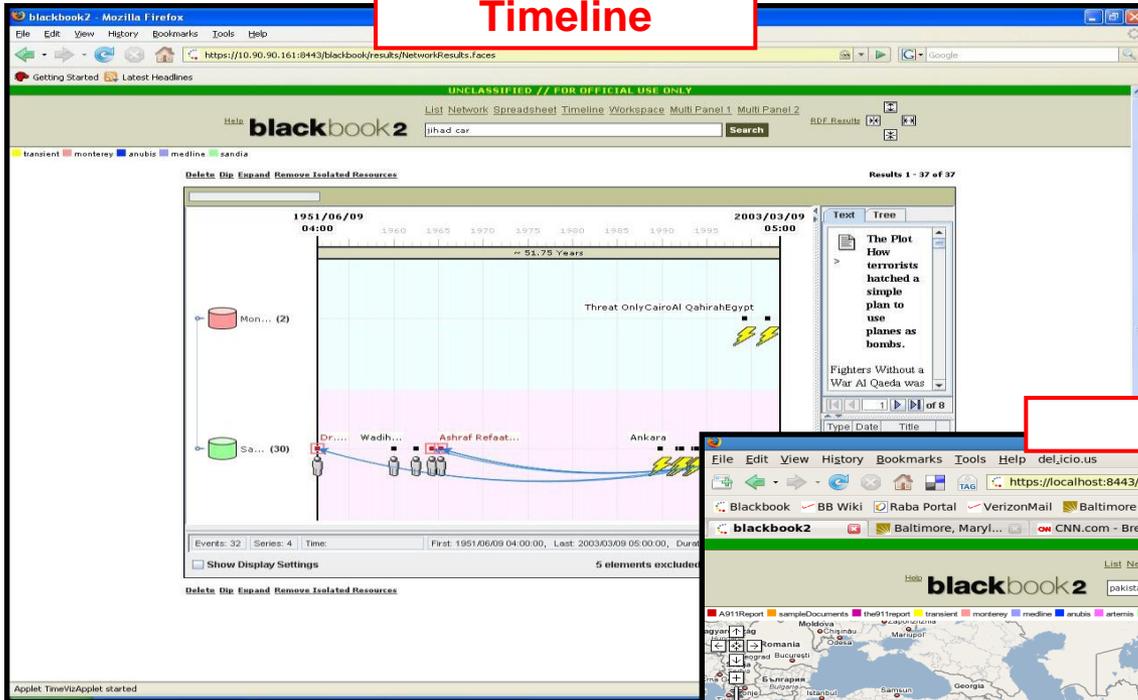
Use Blackbook Data.
Select a graph to load:
_bb_bush.dnv

Level of Detail:
0 0.0 0

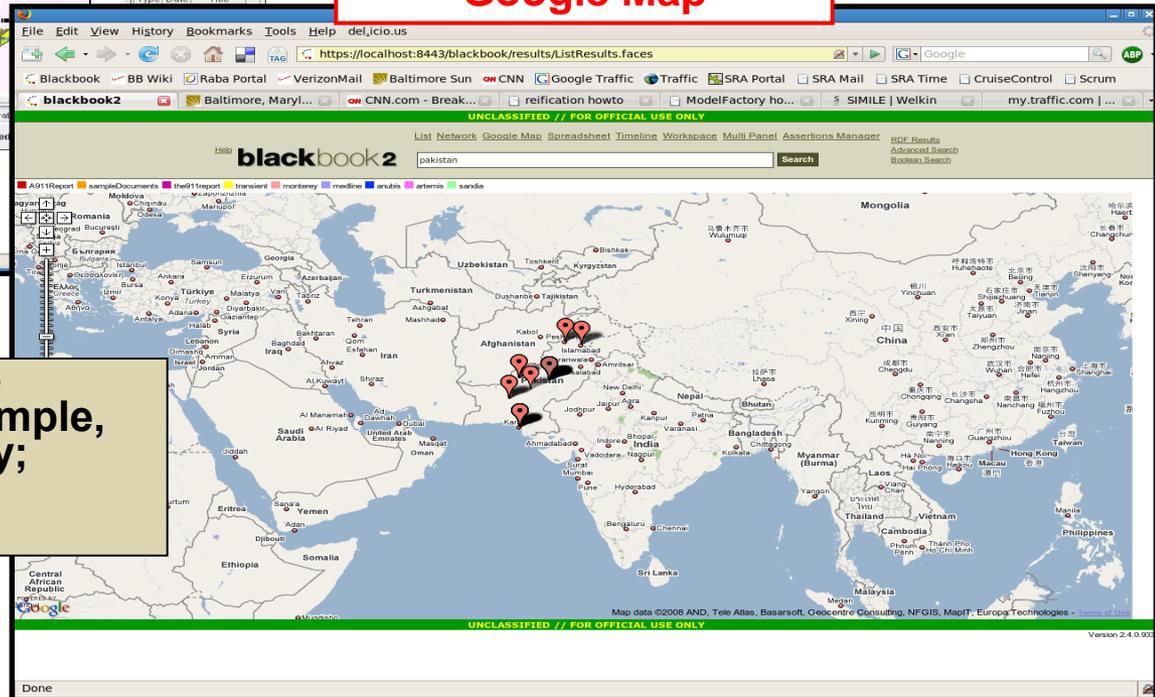
An AJAX-based network visualization, called "WiGi", optimizes client-server processing for large graphs. Planned to be released as early as Blackbook v3.0 (Nov 2009)

User Interface

Timeline



Google Map



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

User Interface

Ozone: Blackbook Widget

The screenshot displays the iGoogle Developer sandbox interface. At the top, there is a navigation bar with links for Home, Weather, My Entities, Multiple SetPref..., and My Gadgets. A search bar is prominently featured in the center, with the text "Welcome to the iGoogle Developer sandbox" below it. The main content area is divided into several sections:

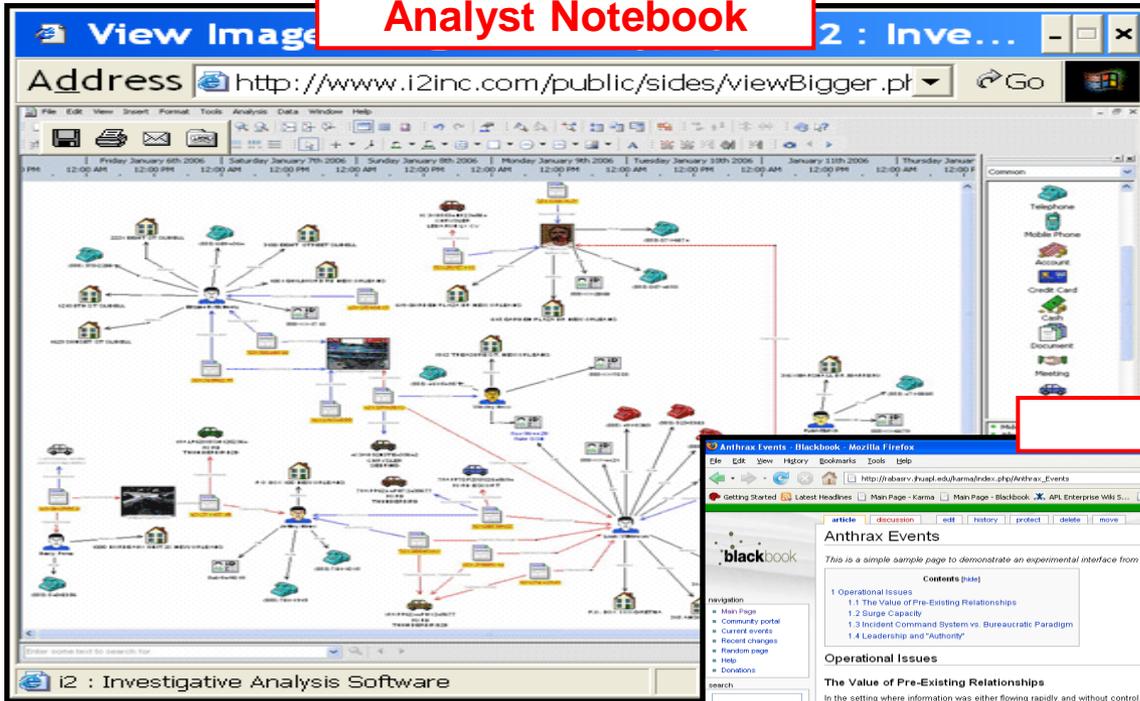
- My Entities:** A form with fields for `rss_url` (http://blackbook2/rss/), `entity_list` (vessel1,vessel2,vessel3), `process_def` (1), and `base_wiki_url` (http://blackbook2/wiki). There are "Save" and "Cancel" buttons.
- Weather:** Two weather widgets. The first is for "Halethorpe, MD" showing a current temperature of 36°F, cloudy conditions, and a 4-day forecast. The second is for "Kill Devil Hills, NC" showing a current temperature of 37°F, clear conditions, and a 4-day forecast.
- My Gadgets:** A section for managing gadgets, including a "New docs" link, a list of gadgets (myAttention.xml, developer.xml, myEntities.xml) with "Inlined" and "Cached" checkboxes, and an "Add a gadget" input field.
- Multiple SetPref - Iframe:** A section with a text area containing the instruction: "Each page load should increment the value of each usepref. Reload page and make sure each usepref is incremented."

At the bottom of the interface, there are links for "Add a theme" and "Mobile - Advertising Programs - Business Solutions - Privacy Policy - Help - About Google".

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

User Interface

Analyst Notebook



Mediawiki

The screenshot shows a Mediawiki page titled "Anthrax Events" within the Blackbook application. The page content includes a table of contents, a section on "Operational Issues" with sub-points, and a section on "The Value of Pre-Existing Relationships" with a paragraph of text. Below the text is a diagram showing relationships between names and locations. The diagram has nodes for "capitol", "anthrax", "Derring", "Fredericks", "Smith", "Evans", "Atta", "Jones", "Simpson", and "Washington". Arrows indicate connections between these nodes. A red oval highlights the diagram and the text above it.

Operational Issues

- 1.1 The Value of Pre-Existing Relationships
- 1.2 Surge Capacity
- 1.3 Incident Command System vs. Bureaucratic Paradigm
- 1.4 Leadership and "Authority"

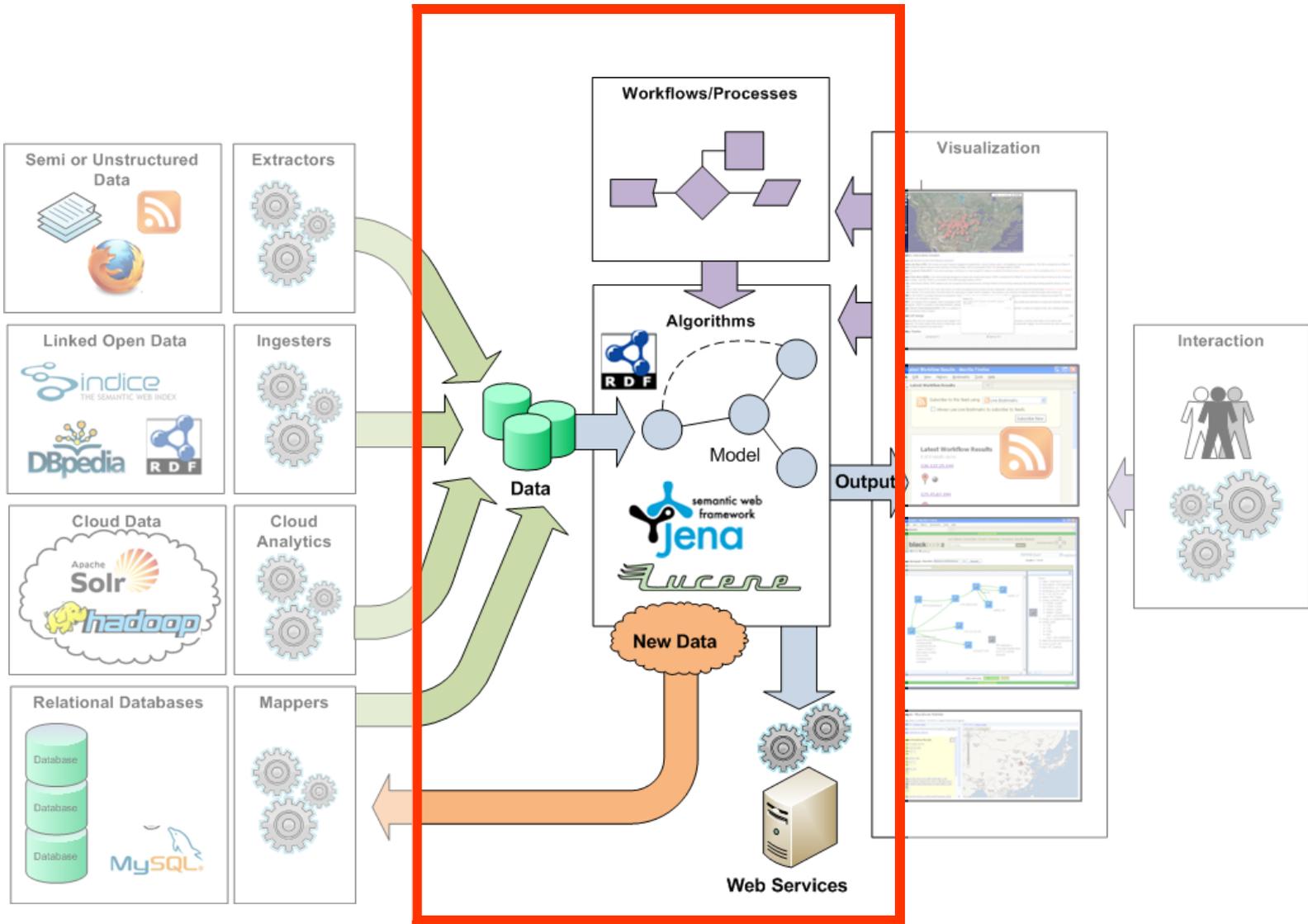
The Value of Pre-Existing Relationships

In the setting where information was either flowing rapidly and without control, or was not flowing at all, where command and control structures were changing rapidly as more and more agencies responded, and where there was a void of leadership structure, the vital need for credible and reliable information was essential. The value of pre-existing and often informal relationships and communication structures cannot be stressed enough. Information was exchanged frequently through midlevel communication chains that had been developed over years of planning and exercising together. These pre-existing channels served to bypass the obstacles of the time-consuming bureaucratic review processes and to bypass the need for administrative and credibility checks that go into the introductory phases of meetings and press conferences. We spoke to people we knew as trusted colleagues in channels that were well rehearsed to get the information and get it done.

Object	Type	Properties
Smith	Person	firstname: Joe lastname: Smith location: 16th Street NW date: 2001-09-16
Jones	Person	firstname: Sam lastname: Smith

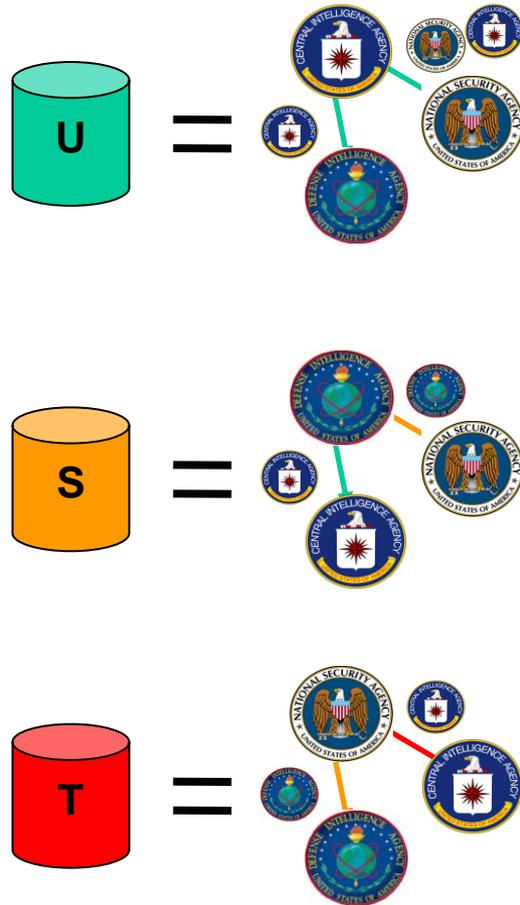
Blackbook is developing a framework called "Aqueduct", allowing interoperability between ozone widgets and wikis.

Middle Tier

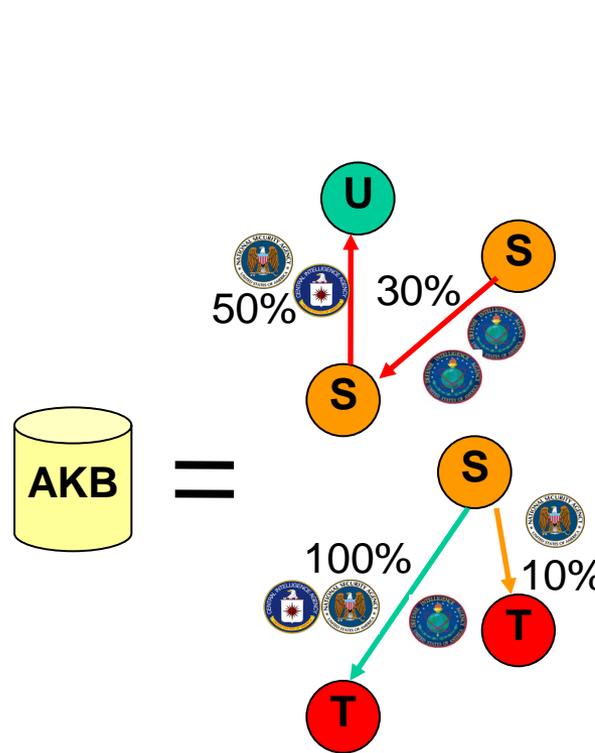


Security, Confidence, Affiliation

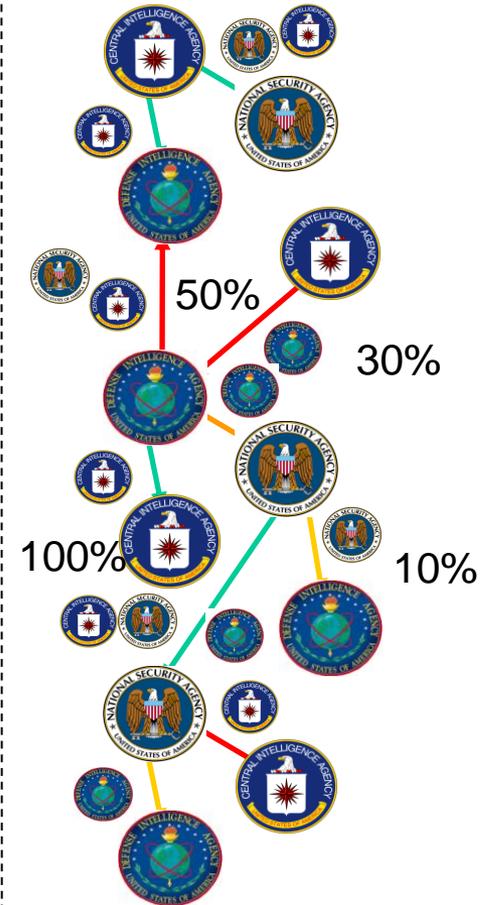
Original Datasource



Analyst Knowledge Base



Composite Knowledge



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 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

User Interface

Workflow

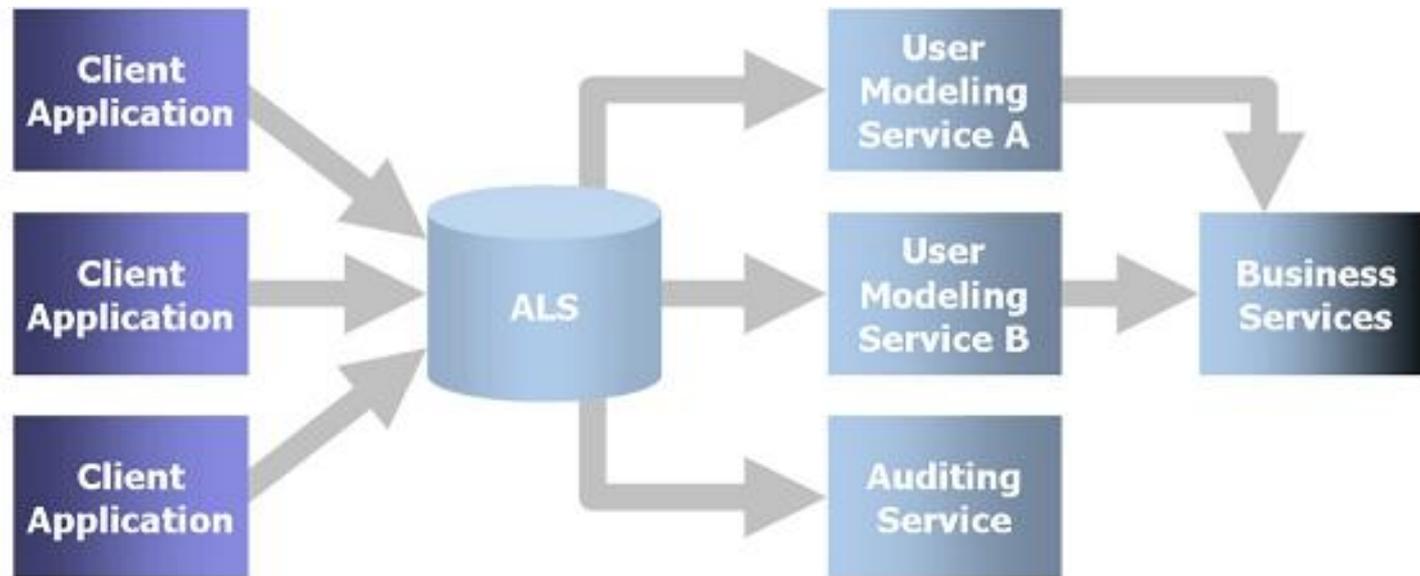
States	To States	Additional Criteria
0. Expand <input type="checkbox"/> fork	<none>	DataAccess: transient
1. Lucene Keyword <input type="checkbox"/> fork	2. Materialize	DataAccess: transient val: jihad car
2. Materialize <input checked="" type="checkbox"/> fork	<none> 0. Expand 1. Lucene Keyword	DataAccess: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

1-4 of 4

The Process Diagram shows a flow from '1. Lucene Keyword' to '2. Materialize', which then branches to '3. Dip' and '0. Expand'. The interface also includes buttons for 'Populate', 'Refresh', 'Refresh Clear', 'MyNewProcess', and 'Save'.

“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.

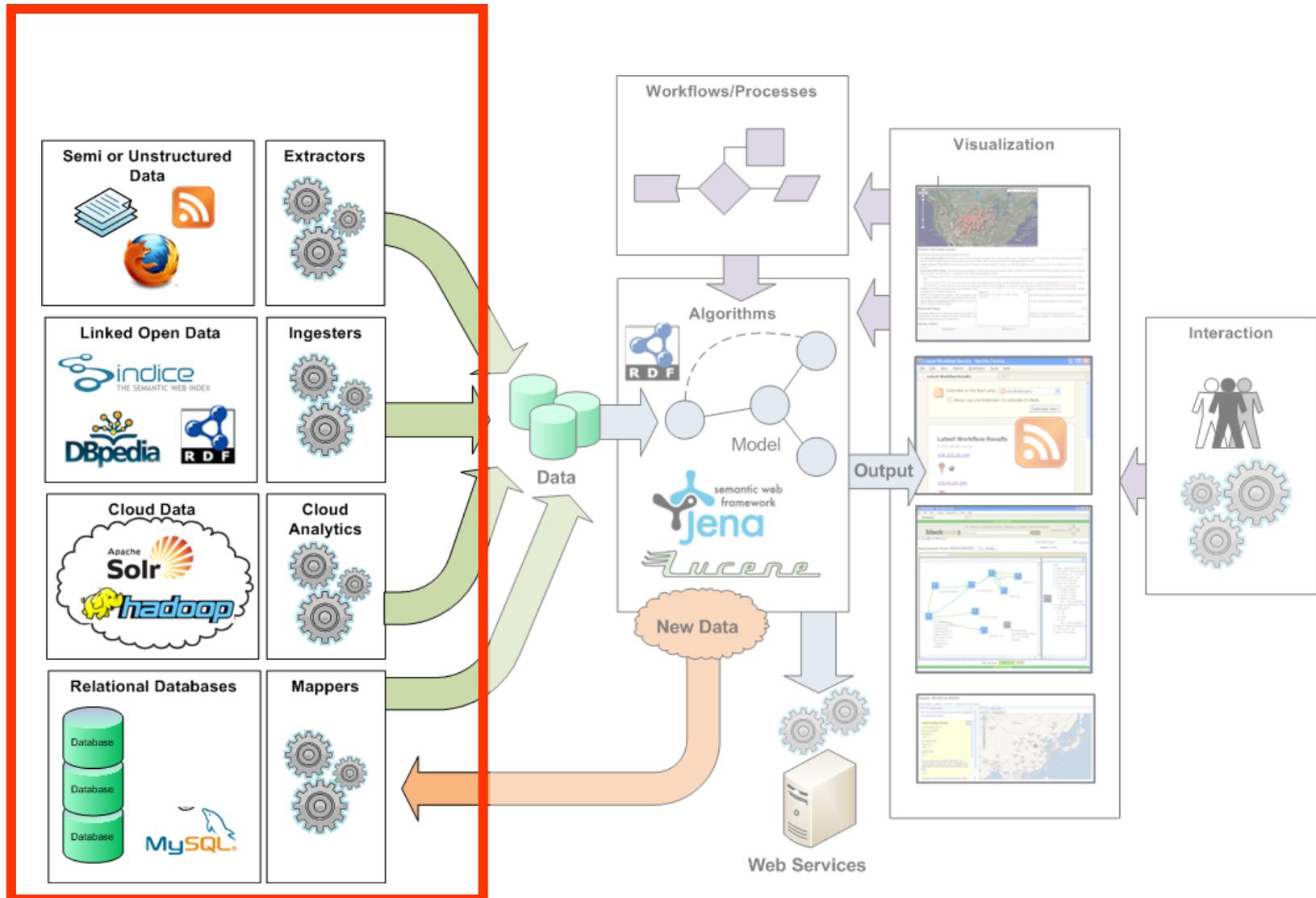
The ALEs are transmitted to the ALS.

The ALS stores the ALEs received from the client applications.

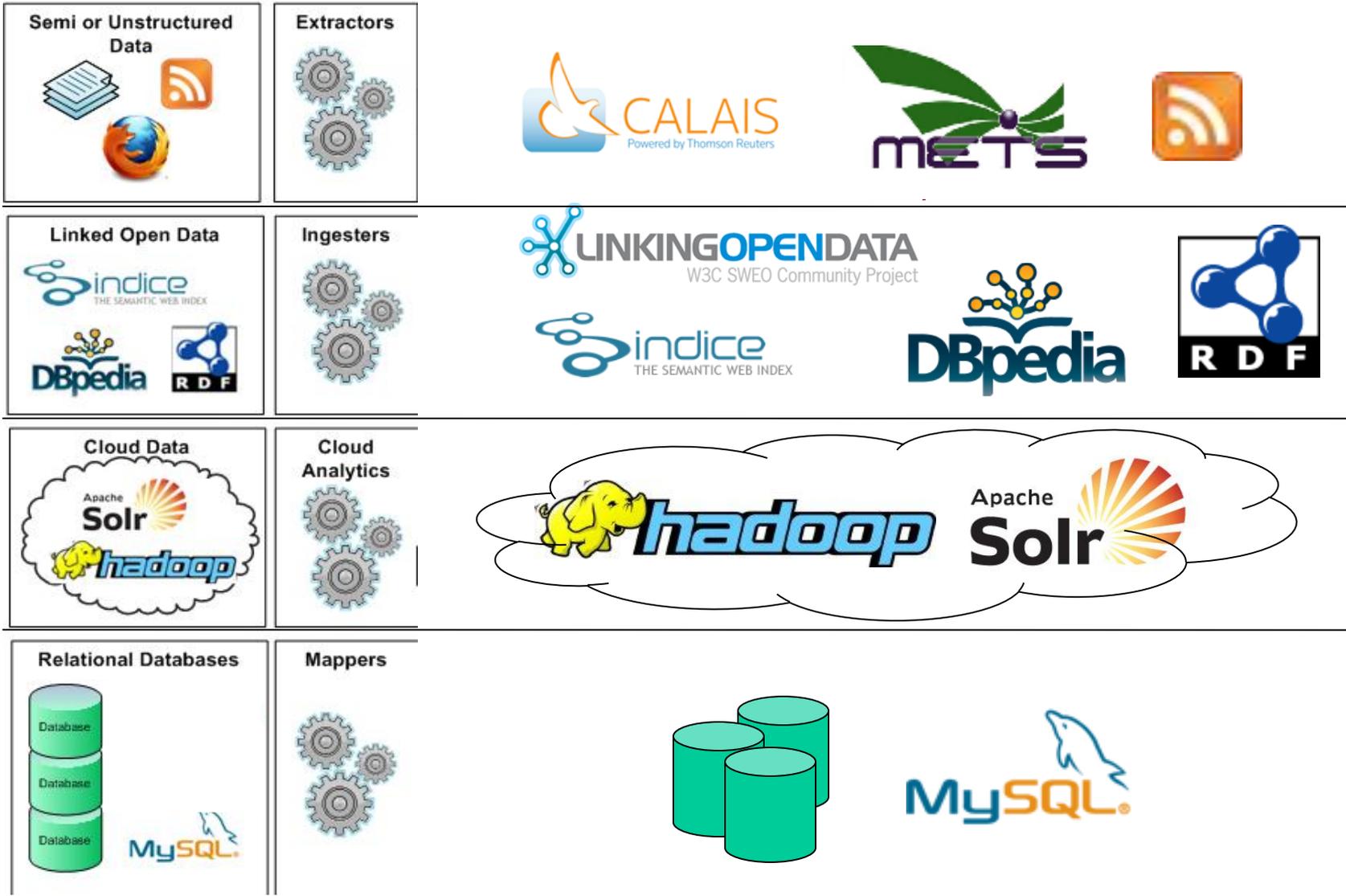
Services interested in using ALEs can query the ALS for ALEs.

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

Data Tier



Data Integration Points

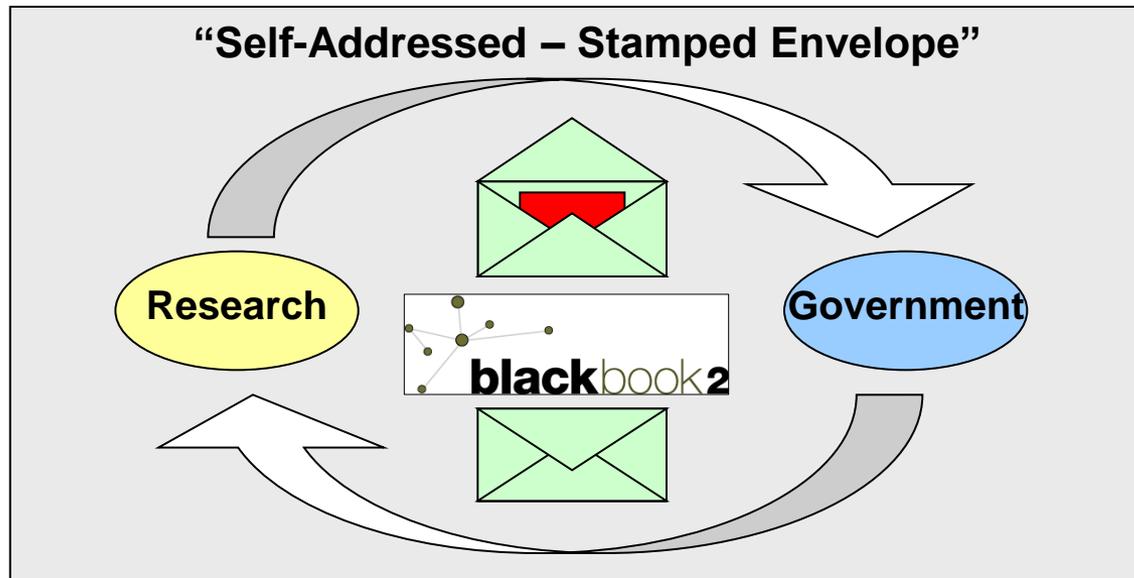


Future Capabilities

- **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)
-

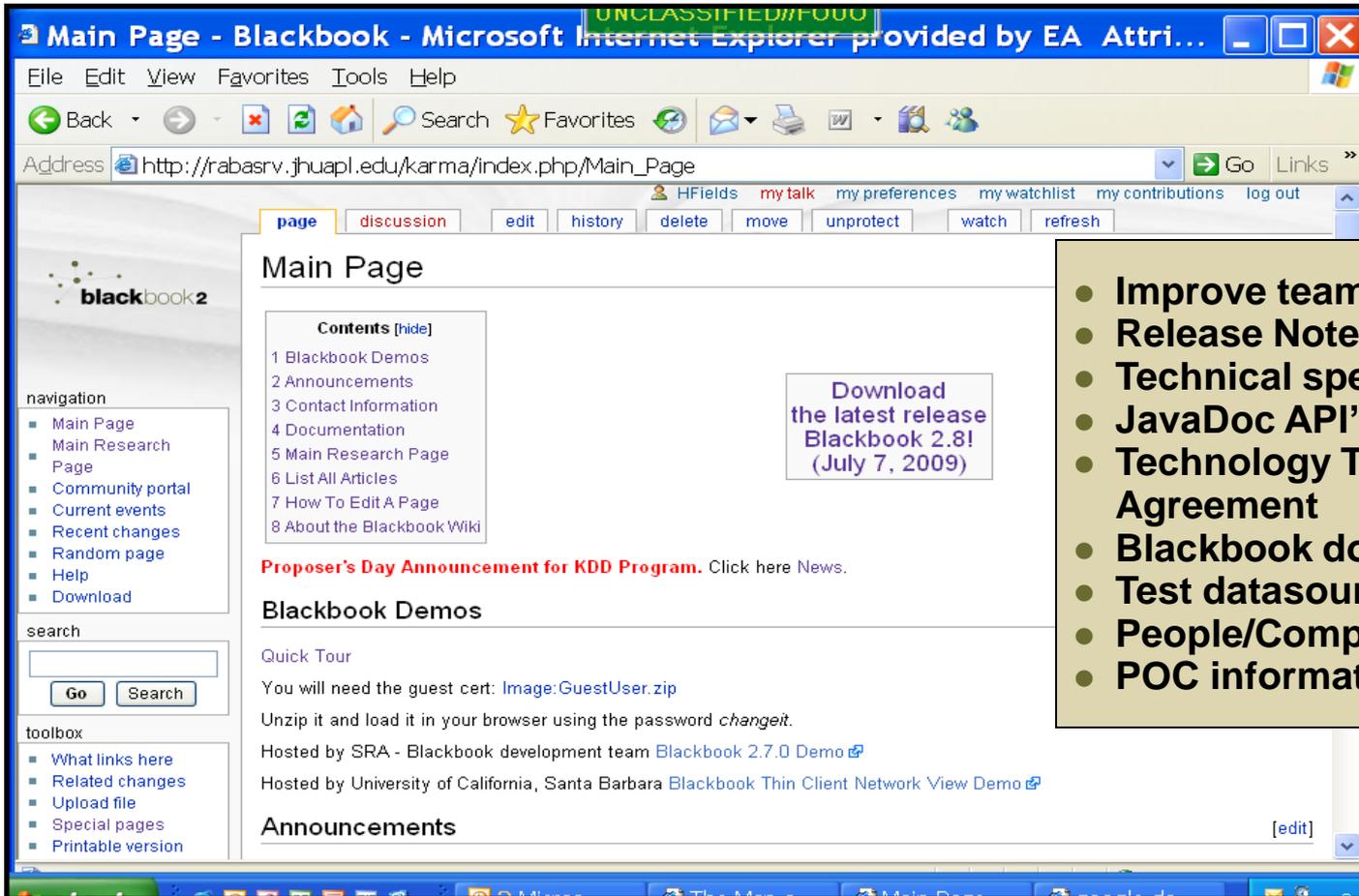
Technology Transfer

- 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
