

From Semantic Search & Integration to Analytics

KMWorld/Intranets 2004

Session: Information Architecture, Track F, October 26, 2004

Amit Sheth

Semagix, Inc. and LSDIS Lab, University of Georgia

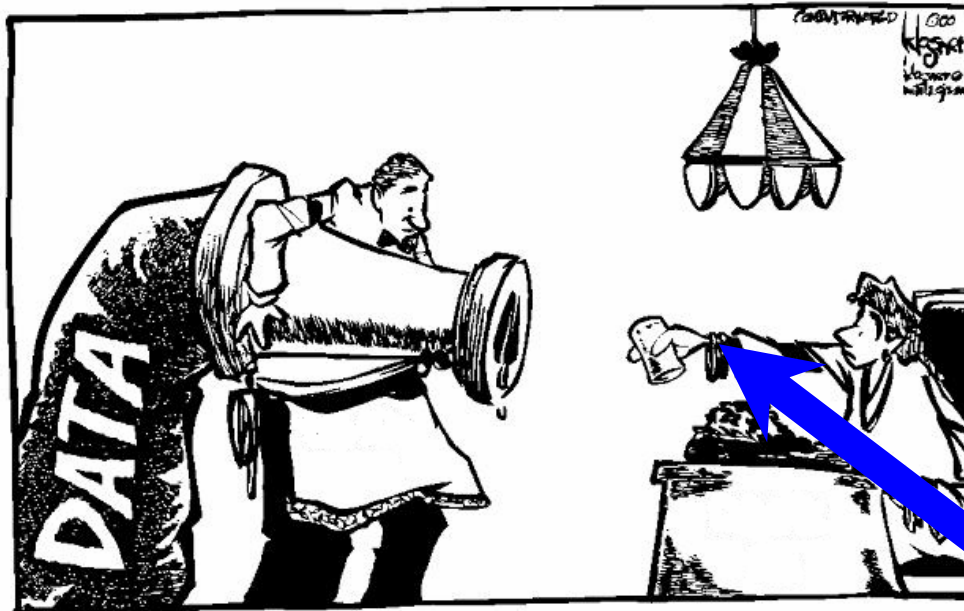
amit.sheth@semagix.com, amit@cs.uga.edu

© Semagix (when marked by Semagix logo); UGARF, Inc. and Amit Sheth (when marked by LSDIS logo)

SEMAGIX
POWER • THROUGH • RELEVANCE



What do you want to do with information?



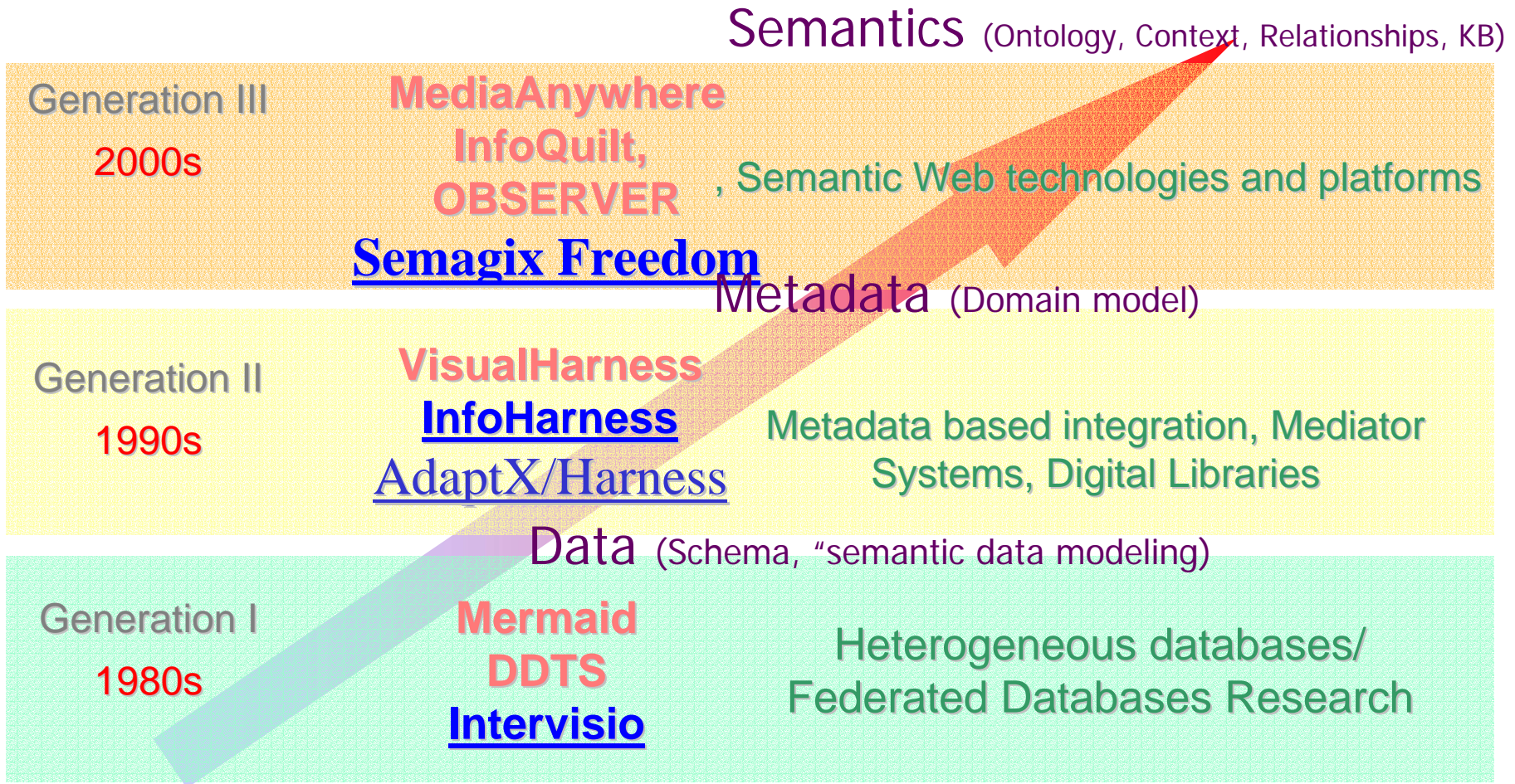
Increasing sophistication in applying semantics

- Relevant Information (Semantic Search & Browsing)
- Semantic Information Interoperability and Integration
- Semantic Correlation/Association, Analysis, Early Warning

**Not just relevant data,
but actionable information,
analysis and insight**



Three generation of Information Systems: Where we have come from, where we are going



Outline

- Observations
- Semantic Web and its key enablers
 - Ontology: What do real world ontologies look like, how are they created
 - Metadata: what enables automatic metadata extraction, how metadata enables integration of heterogeneous content
- What type of industry and scientific Semantic Applications are built
- Conclusions: Ontologies as a best-in-class approach; analytics as in most demanding applications

Emerging Trends, Changing Focus

- From syntax/structure to semantics
- From techniques that focus on either unstructured data (text) or structured content, to both types and semi-structured data
- From directly analyzing data (warehousing and mining) to ontology based processes of creating high quality metadata and analyzing metadata
- From search and browsing for delivering relevant documents; from locating entities within contents to discovering complex relationships and delivering actionable information with insights;
from semantic search to integration, mining, analytics

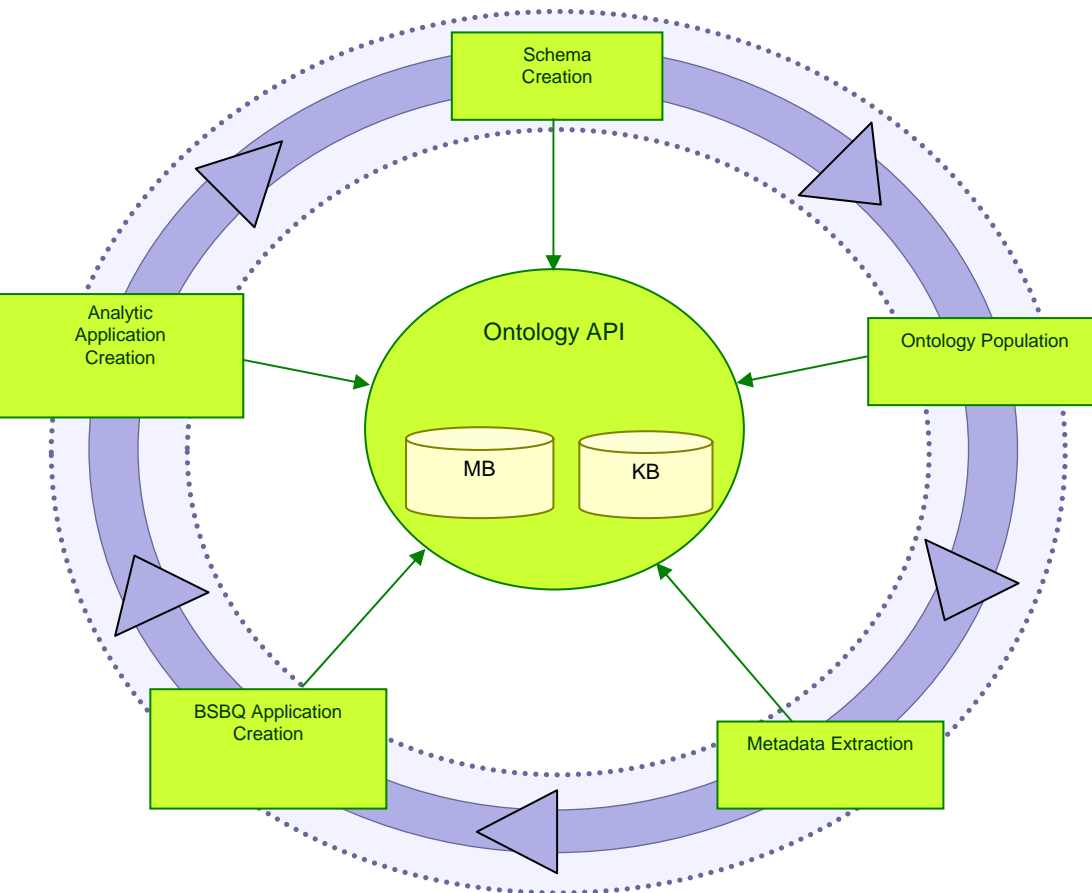
Empirical observations based on real-world efforts

- Comprehensive commercial products support development Ontology-based information systems and Semantic Applications
 - Several commercial companies with offerings that vary from tools to comprehensive enterprise software platforms
 - Deployed, business critical applications can be found
 - Applications validate the importance of ontology the current semantic approaches
 - Significant impact of academic research in early stage companies
- Empirical observations in this talk are based on development of several real-world ontologies and implemented/deployed semantic applications by Semagix and its partners/customers

Semantic Web in a nutshell

- Ontology as the centerpiece
- Metadata that associate meaning to content
- Computing (complex querying, inferencing, other reasoning) that support semantic applications

Ontology-driven Information System Lifecycle



Building a scalable and high performance system with support for:

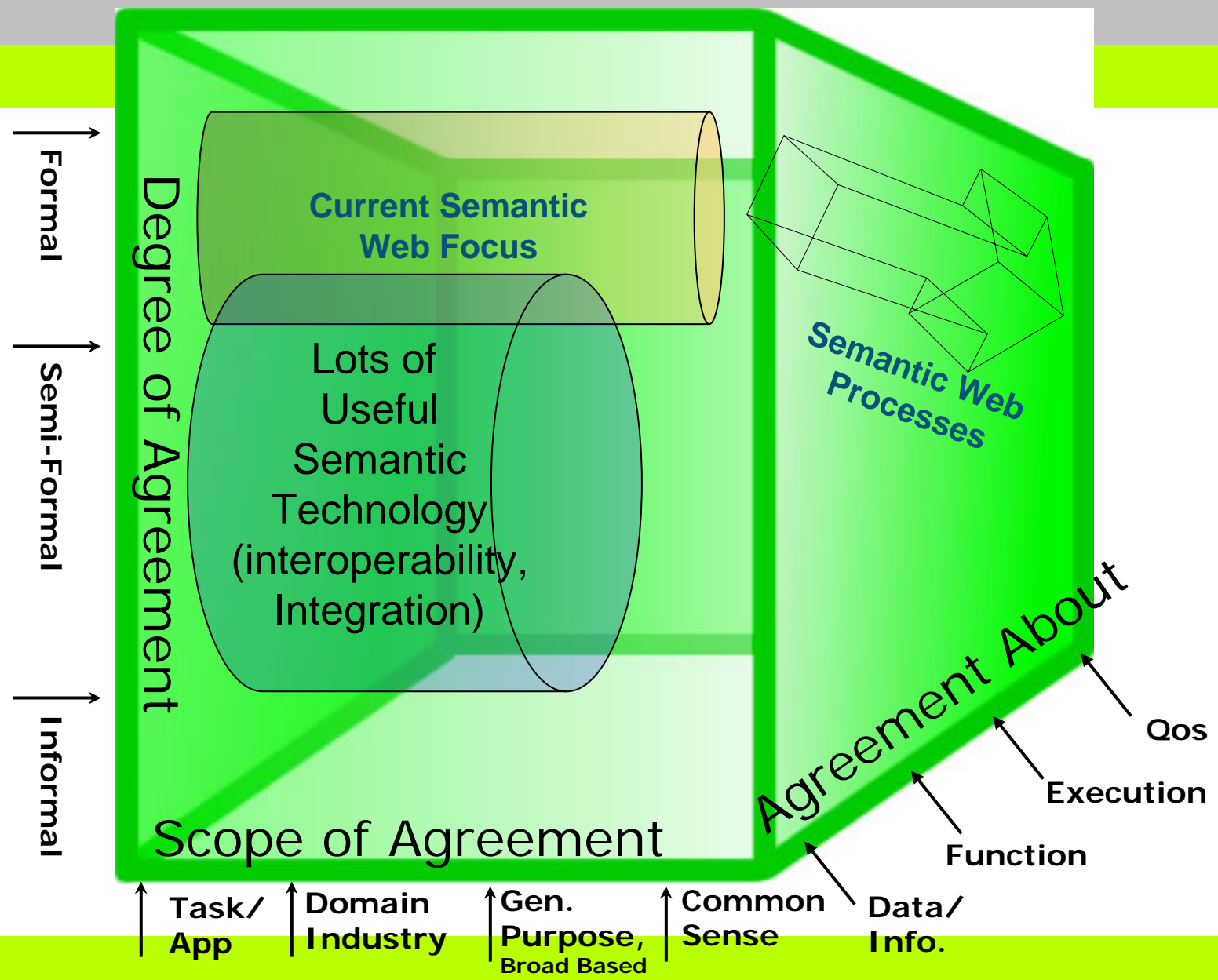
- Ontology creation and maintenance
- Knowledge-based (and other techniques) supporting Automatic Classification
- Ontology-driven Semantic Metadata Extraction/Annotation
- Utilizing semantic metadata and ontology
 - Semantic search/querying/browsing
 - Information and application integration - normalization
 - Analysis/Mining/Discovery – relationships

Central Role of Ontology

- Ontology represents agreement, represents common terminology/nomenclature
- Ontology is populated with extensive domain knowledge or known facts/assertions
- Key enabler of semantic metadata extraction from all forms of content:
 - unstructured text (and 150 file formats)
 - semi-structured (HTML, XML) and
 - structured data
- Ontology is in turn the center price that enables
 - resolution of semantic heterogeneity
 - semantic integration
 - semantically correlating/associating objects and documents



Broad Scope of Semantic (Web) Technology



Other dimensions:
how agreements are reached,
...



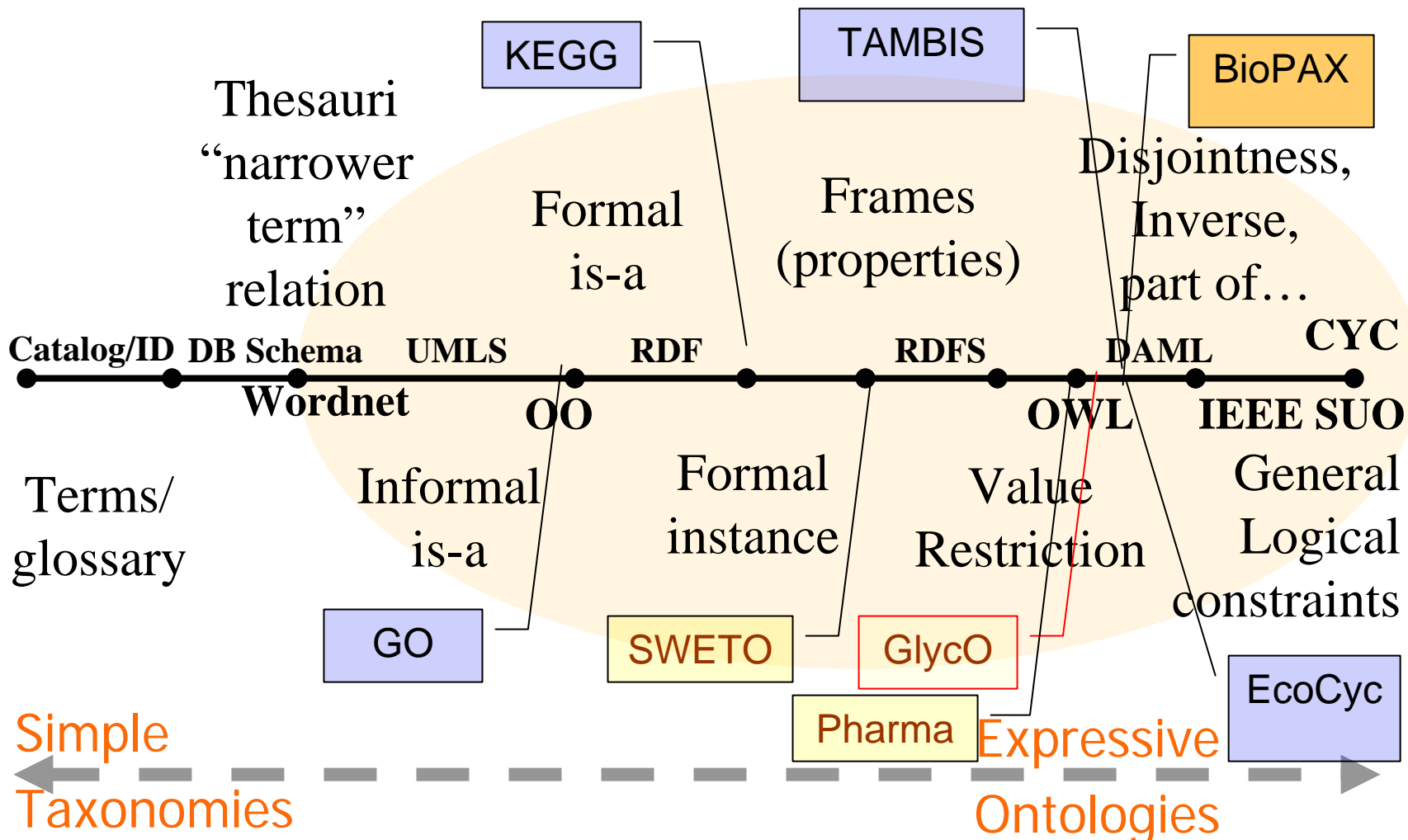
Types of Ontologies (or things close to ontology)

- Upper ontologies: modeling of time, space, process, etc
- Broad-based or general purpose ontology/nomenclatures: Cyc, CIRCA ontology (Applied Semantics), **SWETO**, *WordNet* ;
- Domain-specific or Industry specific ontologies
 - News: politics, sports, business, entertainment
 - Financial Market
 - Terrorism
 - Pharma
 - **GlycO**
 - (*GO (a nomenclature), UMLS inspired ontology, ...*)
- Application Specific and Task specific ontologies
 - Anti-money laundering
 - Equity Research
 - Repertoire Management

Fundamentally different approaches in developing ontologies
at the two end of the above spectrum



Expressiveness Range: Knowledge Representation and Ontologies





Building ontology

- Three broad approaches:
 - social process/manual: many years, committees
 - Can be based on metadata standard
 - automatic taxonomy generation (statistical clustering/NLP):
limitation/problems on quality, dependence on corpus, naming
 - Descriptive component (schema) designed by domain experts;
Description base (assertional component, extension) by automated processes

Option 2 is being investigated in several research projects;

Option 3 is currently supported by Semagix Freedom

Ontology can be very large

Semantic Web Ontology Evaluation Testbed – SWETO v1.4
is

- Populated with over 800,000 entities and over 1,500,000 explicit relationships among them
- Continue to populate the ontology with diverse sources thereby extending it in multiple domains, new larger release due soon
- Two other ontologies of Semagix customers have over 10 million instances, and requests for even larger ontologies exist

GlycO

- *is a focused ontology for the description of glycomics*
- *models the biosynthesis, metabolism, and biological relevance of complex glycans*
- models complex carbohydrates as sets of simpler structures that are connected with rich relationships

Ontology schema can be large and complex

GlycO statistics now (and growing)

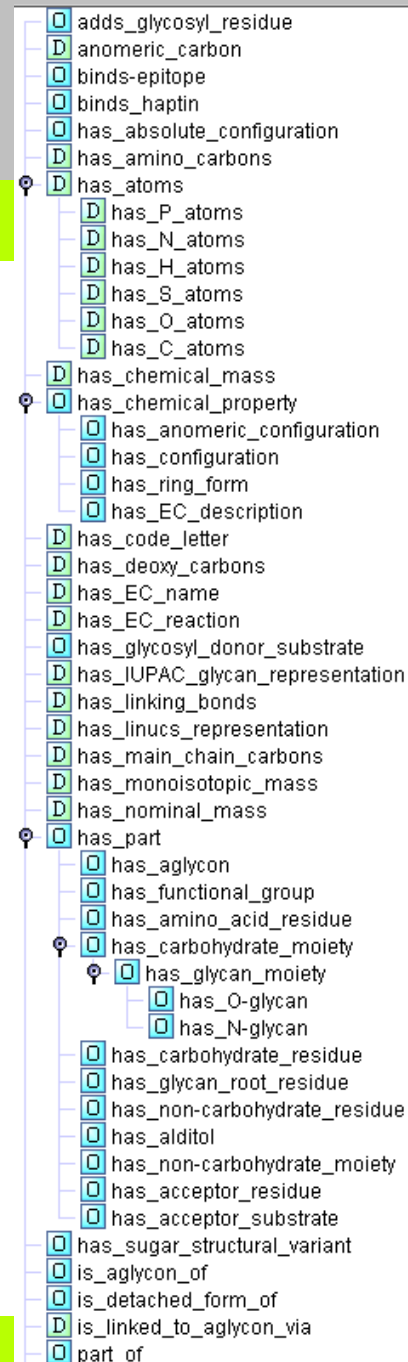
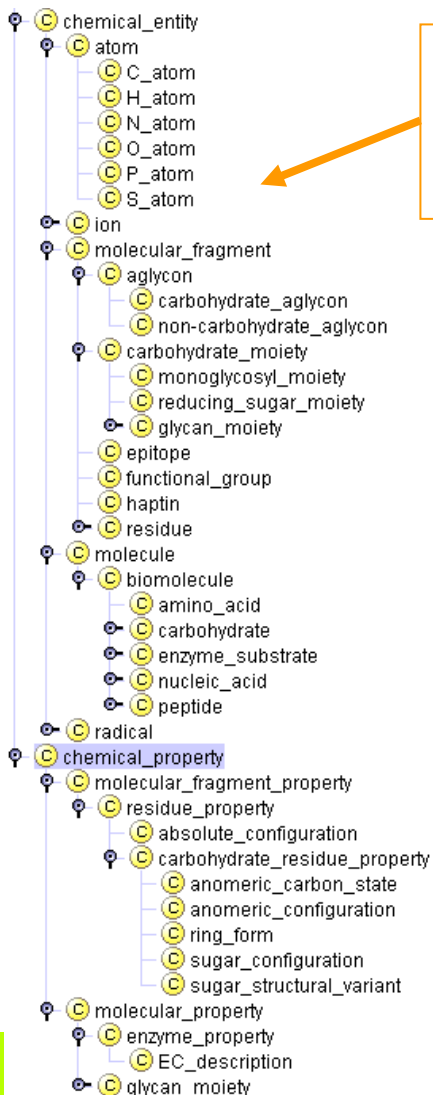
- 767 classes
- 142 slots
- Instances Extracted with Semagix Freedom:
 - 69,516 genes (From PharmGKB and KEGG)
 - 92,800 proteins (from SwissProt)
 - 18,343 publications (from CarbBank and MedLine)
 - 12,308 chemical compounds (from KEGG)
 - 3,193 enzymes (from KEGG)
 - 5,872 chemical reactions (from KEGG)
 - 2210 N-glycans (from KEGG)

GlycO taxonomy

The first levels of the GlycO taxonomy

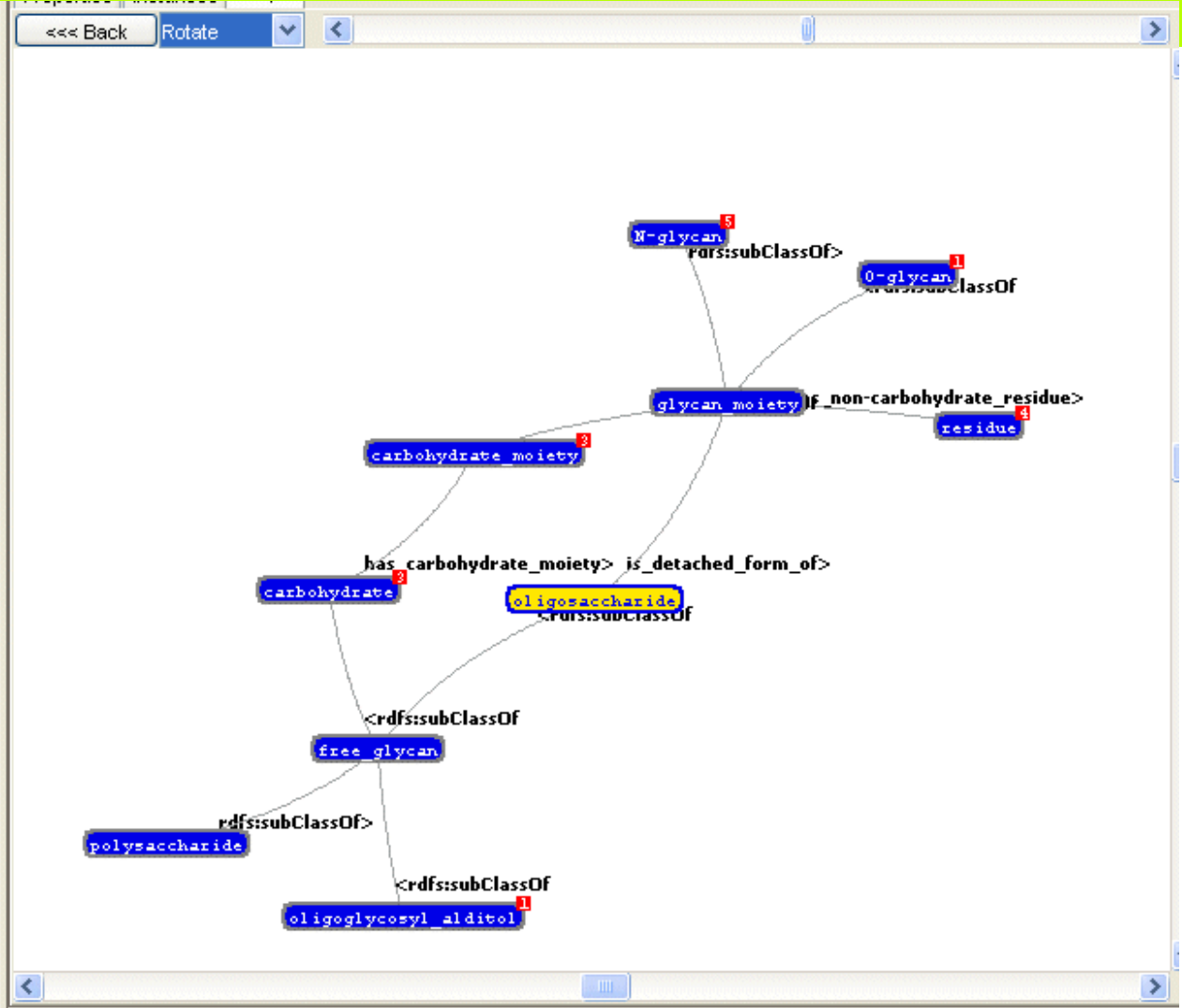
Most relationships and attributes in GlycO

GlycO exploits the expressiveness of OWL-DL. Cardinality constraints, value constraints, Existential and Universal restrictions on Range and Domain of properties allow the classification of unknown entities as well as the deduction of implicit relationships.



Query and visualization

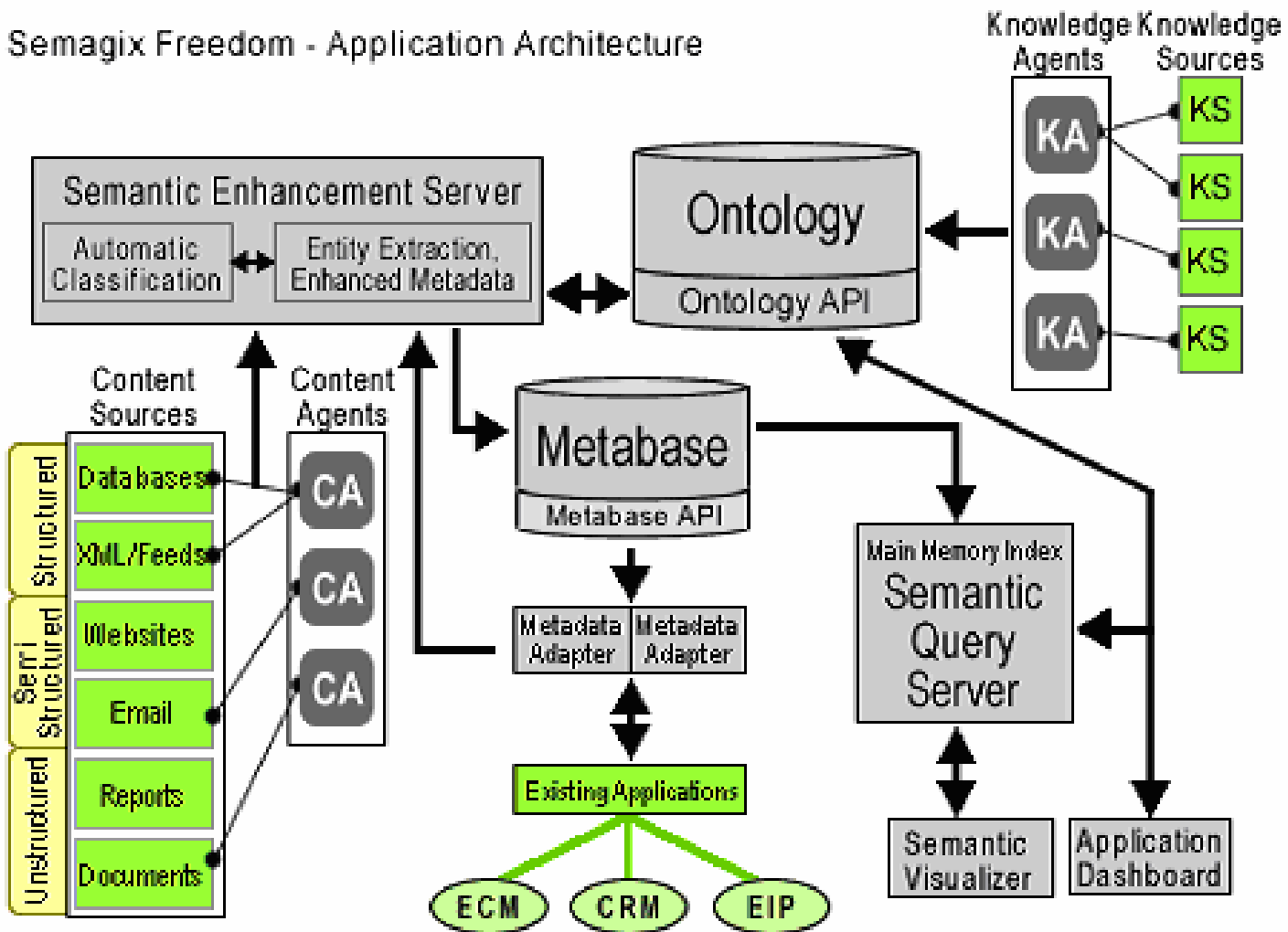
- owl:Thing
 - chemical_property
 - molecular_property
 - molecular_fragment_property
 - chemical_entity
 - radical
 - molecular_fragment
 - epitope
 - aglycon
 - haptin
 - functional_group
 - carbohydrate_moiety
 - reducing_sugar_moiety
 - glycan_moiety
 - monoglycosyl_moiety
 - residue
 - nucleotide_residue
 - carbohydrate_residue
 - amino_acid_residue
 - atom
 - molecule
 - ion



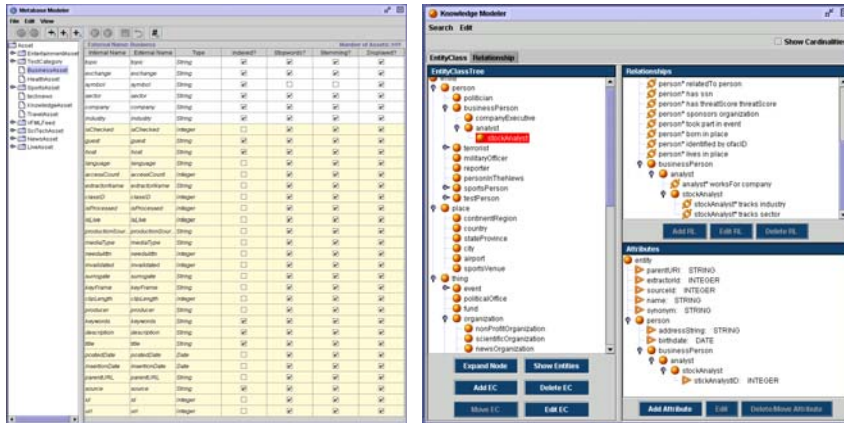
Description and Provenance Information for glycan_moiety : a glycan moiety is a glycan that is part of a glycoconjugate, which has a non-carbohydrate moiety

Semagix Freedom Architecture: for building ontology-driven information system

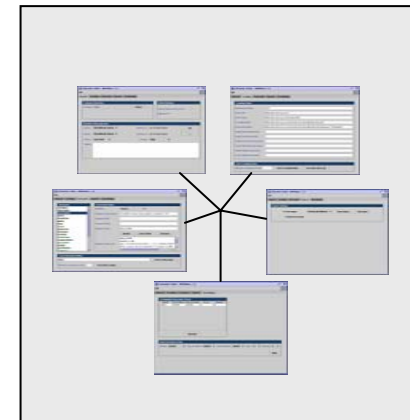
Semagix Freedom - Application Architecture



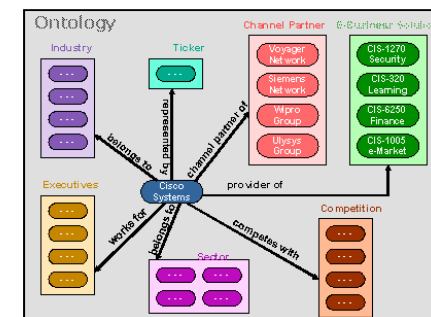
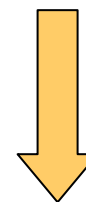
Ontology Creation and Maintenance Steps



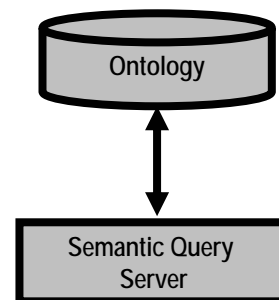
1. Ontology Model Creation (Description)



2. Knowledge Agent Creation

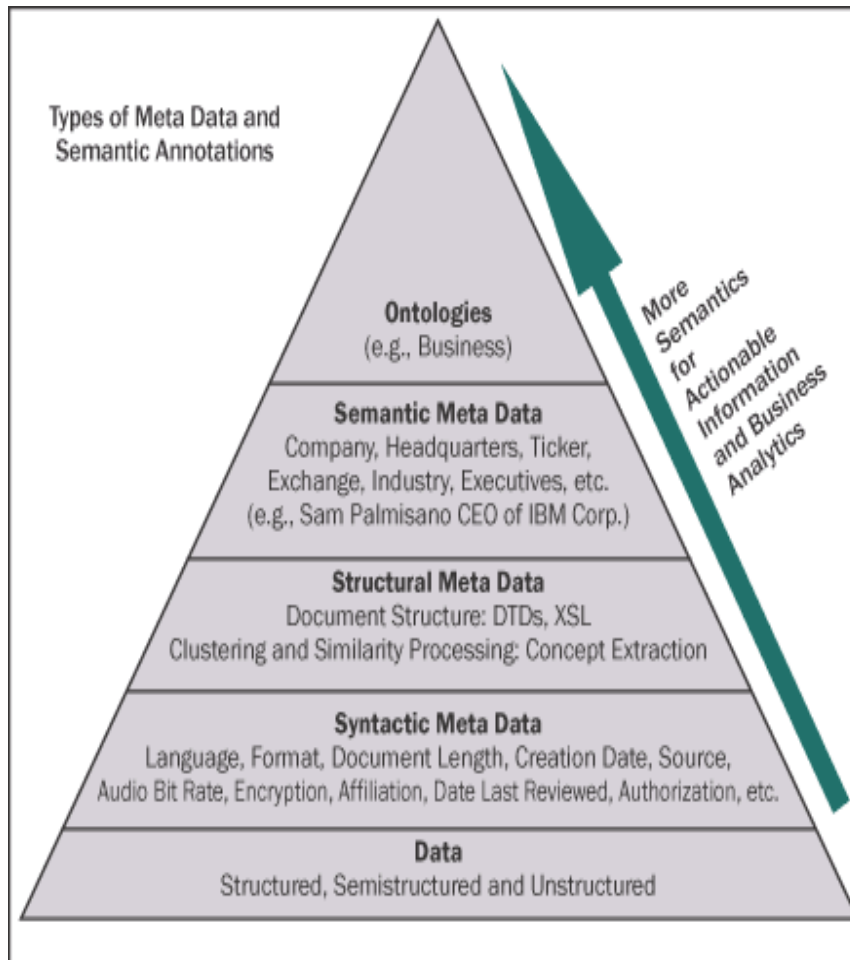


3. Automatic aggregation of Knowledge



4. Querying the Ontology

The Evolution of Meta Data

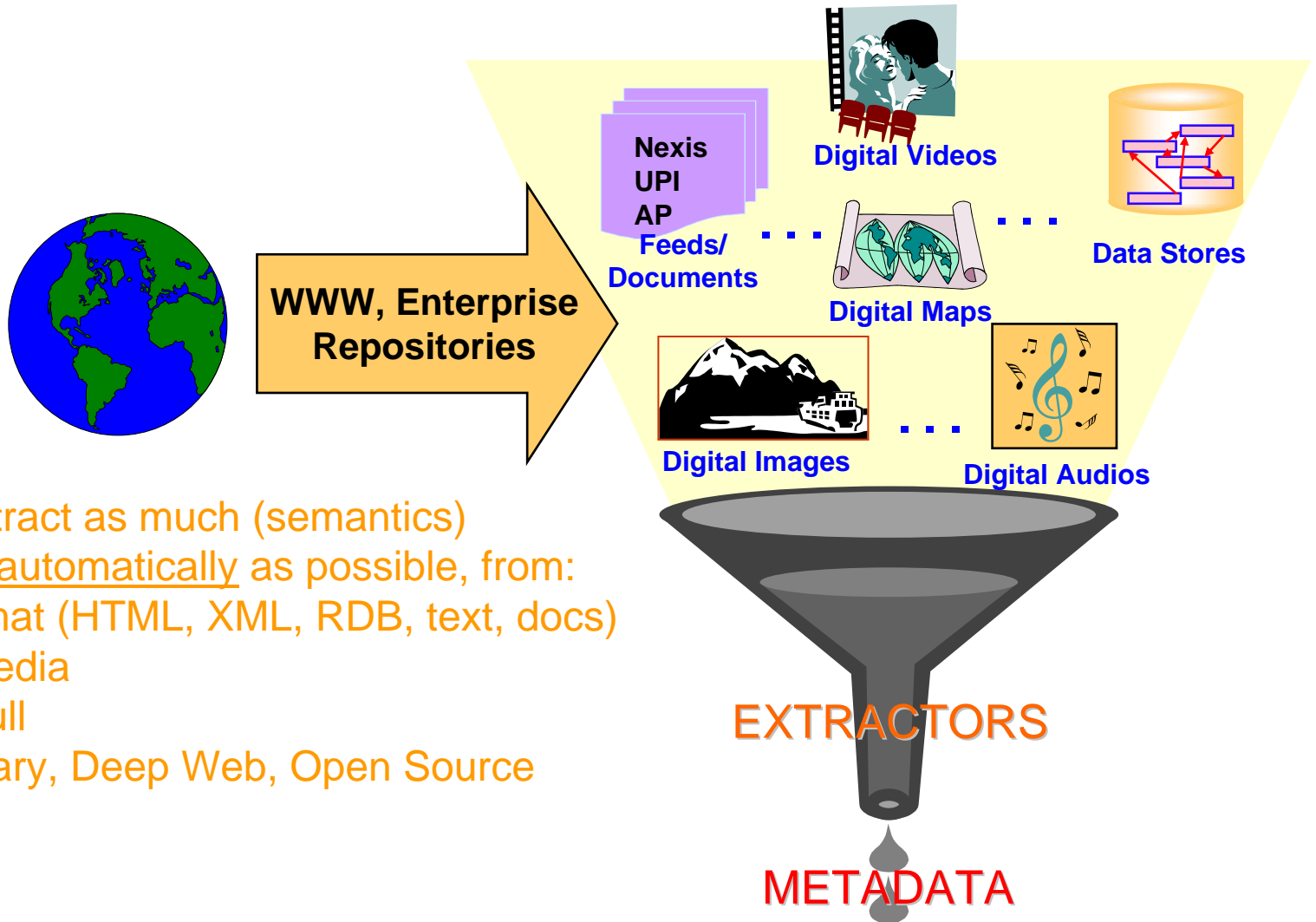


The more sophisticated technologies:

- Provide scalability and flexibility
- Handle all types of data (unstructured, semi-structured, structured)
- Accommodate SmartQuerying – flexible, intelligent querying
- Create SmartData – enhancing raw data with context and relationships
- Enable powerful enterprise decisionmaking



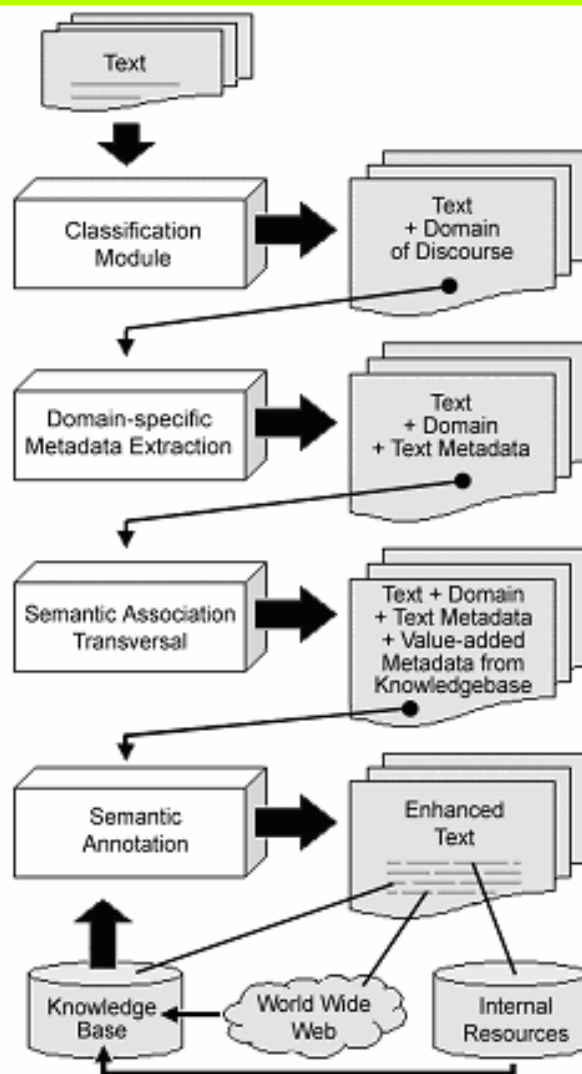
Metadata extraction from heterogeneous content/data



Create/extract as much (semantics) metadata automatically as possible, from:

- Any format (HTML, XML, RDB, text, docs)
- Many media
- Push, pull
- Proprietary, Deep Web, Open Source

Metadata Extraction and Semantic Enhancement



Semantic Annotation/ Metadata Extraction

+

Enhancement

[Bancroft, Hammond, Sheth]

Blue-chip bonanza continues

company company company
 Dow above 9,000 as [HP](#), [Home Depot](#) lead advance; [Microsoft](#) upgrade helps techs.
date time
 August 22, 2002: 11:44 AM EDT
phrase phrase
 By Alexandra Twin, CNN/Money Staff Writer
city company
[New York](#) (CNN/Money) - An upgrade of software leader [Microsoft](#) and strength in blue chips including
company company weekday
[Hewlett-Packard](#) and [Home Depot](#) were among the factors pushing stocks higher at midday [Thursday](#),
financial index
 with the [Dow Jones industrial average](#) spending time above the 9,000 level.
time financial index
 Around 11:40 a.m. ET, the [Dow Jones industrial average](#) gained 65.06 to 9,022.09, continuing a more
date stock exchange
 than 1,300-point resurgence since [July 23](#). The [Nasdaq](#) composite gained 9.12 to 1,418.37.
financial index
[The Standard & Poor's 500 index](#) rose 9.61 to 958.97.
company stockSym \$ \$
[Hewlett-Packard](#) ([HPQ](#): up \$0.33 to \$15.03, Research, Estimates) said a report shows its share of
 the printer market grew in the second quarter, although another report showed that its share of the
continent region continent
 computer server market declined in [Europe](#), the [Middle East](#) and [Africa](#).
company stockSym \$ \$
[Home Depot](#) ([HD](#): up \$1.07 to \$33.75, Research, Estimates) was up for the third straight day after
 topping fiscal second-quarter earnings estimates on Tuesday.
tech category company
 Tech stocks managed a turnaround. [Software](#) continued to rise after [Salomon Smith Barney](#) upgraded
company stockSym \$ \$
 No. 1 software maker [Microsoft](#) ([MSFT](#): up \$0.55 to \$52.83, Research, Estimates) to "outperform"
\$ \$ company
 from "neutral" and raised its price target to \$59 from \$56. Business software makers [Oracle](#)
stockSym \$ \$ company stockSym \$ \$
 ([ORCL](#): up \$0.18 to \$10.94, Research, Estimates), [PeopleSoft](#) ([PSFT](#): up \$1.17 to \$20.67,
company stockSym \$ \$
 Research, Estimates) and [BEA Systems](#) ([BEAS](#): up \$0.28 to \$7.12, Research, Estimates)
 all rose in tandem.

competes with

Automatic Semantic Annotation

```

<body>
<body>
<dist>
</body>
</body>
<p>P
is t
pla
<p>
Se
mc
del
els
Go
<p>S
loc
<p>M
Tin
coi
sai
<p>G
</body>
</html>
</Identification>
<NewsManagement>
<NewsItemType FormalName="News" />
<FirstCreated>20010901T142709</FirstCreated>
<ThisRevisionCreated>20010901T142709</ThisRevisionCreated>
<Status Vocabulary="urn:newsml:comtexnews.net:20010101:Com
Scheme="ComtexStatus" FormalName="Usable" />
<Urgency FormalName="5" />
</NewsManagement>
<NewsComponent Essential="yes" EquivalentsList="no">
<Role FormalName="Main" />
<NewsLines>
<HeadLine>The Debate is on the Future of Timor Sea LNG</H
<DateLine>Sep 01, 2001 (Gas-to-Liquids News/PBI Media via
<CopyrightLine>Copyright 2001 PBI Media, LLC. All rights res
<KeywordLine Origin="Comtex">australia</KeywordLine>
<KeywordLine Origin="Comtex">oil</KeywordLine>
<KeywordLine Origin="Comtex">transportation</KeywordLine>
</NewsLines>
<AdministrativeMetadata>
<Catalog>
<Resource>
<Urn>urn:newsml:comtexnews.net:20010201:Providers:1
<DefaultVocabularyFor Context="Provider/Party" />
</Resource>
<Resource>
<Urn>urn:newsml:comtexnews.net:20010201:Source:1</U
<DefaultVocabularyFor Context="Source/Party" />
</Resource>
<Resource>
<Urn>urn:newsml:comtexnews.net:20010201:SourceProp
<DefaultVocabularyFor Context="Property" />
</Resource>
</Catalog>
<FileName>22698565.xml</FileName>
<Provider>
<Party FormalName="Comtex" />
</Provider>
<Source>
<Party FormalName="Phillips Full" />
</Source>
<Property FormalName="SourceCode" Value="PHP" />
</AdministrativeMetadata>
<RightsMetadata>
<CopyrightDate>2001</CopyrightDate>
</RightsMetadata>
<DescriptiveMetadata>
<Language FormalName="en" />
<Property FormalName="PublicCompany"
Vocabulary="urn:newsml:comtexnews.net:20010201:Domest
<Property FormalName="CompanyName" Value="Phillips Petr
<Property FormalName="StockSymbol" Value="P" />
</Property>
</DescriptiveMetadata>
<ContentItem>
<MediaType FormalName="Text" />
<MimeType FormalName="text/vnd.IPTC.NITF" />
</DataContent>

```

```

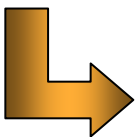
Value-added Semagix Semantic Tagging
<Language FormalName="en" />
<Property FormalName="PublicCompany" Vocabulary="urn:newsml:comtexnews.net:20010201:DomesticPublicCompanies:1">
<Property FormalName="CompanyName" Value="Phillips Petroleum Co." />
<Property FormalName="StockSymbol" Value="P" />
<Property FormalName="Competitor">
<Property FormalName="CompanyName" Value="BP p.l.c." />
</Property>
<Property FormalName="Competitor">
<Property FormalName="CompanyName" Value="Ultramar Diamond Shamrock Corp." />
</Property>
<Property FormalName="Competitor">
<Property FormalName="CompanyName" Value="Royal Dutch/Shell Group" />
</Property>
<Property FormalName="Competitor">
<Property FormalName="CompanyName" Value="Bartlesville, Oklahoma, United States of A" />
</Property>
<Property FormalName="Headquarters" Value="Bartlesville, Oklahoma, United States of A" />
<Property FormalName="StockExchange" Value="NYSE" />
<Property FormalName="Sector" Value="Energy" />
<Property FormalName="Industry" Value="Integrated Oil and Gas" />
+ <Property FormalName="CompanyExecutive" Value="Augustine, Norman R." />
+ <Property FormalName="CompanyExecutive" Value="Boren, David L." />
+ <Property FormalName="CompanyExecutive" Value="Chappell, Jr., Robert E." />
+ <Property FormalName="CompanyExecutive" Value="Devlin, Robert" />
+ <Property FormalName="CompanyExecutive" Value="Horner, Larry D." />
+ <Property FormalName="CompanyExecutive" Value="Roy, J. Stapleton" />
+ <Property FormalName="CompanyExecutive" Value="Tobias, Randall L." />
+ <Property FormalName="CompanyExecutive" Value="Tschinkel, Victoria J." />
<Property FormalName="CompanyPosition" Value="Director" />
</Property>
<Property FormalName="CompanyExecutive" Value="Turner, Kathryn C." />
<Property FormalName="CompanyPosition" Value="Director" />
</Property>
<Property FormalName="CompanyExecutive" Value="Meyers, Ph.D., Kevin" />
<Property FormalName="CompanyPosition" Value="Executive Vice President, Alaska Operations" />
</Property>
<Property FormalName="CompanyExecutive" Value="Lowe, John" />
<Property FormalName="CompanyPosition" Value="Senior Vice President, Planning and Strategic Transactions" />
</Property>
<Property FormalName="CompanyExecutive" Value="Mulva, J. J." />
<Property FormalName="CompanyPosition" Value="Chairman of the Board" />
<Property FormalName="CompanyPosition" Value="Chief Executive Officer" />
</Property>
<Property FormalName="CompanyExecutive" Value="Batchelder, E. L." />
<Property FormalName="CompanyPosition" Value="Vice President" />
<Property FormalName="CompanyPosition" Value="Chief Information Officer" />
</Property>
<Property FormalName="CompanyExecutive" Value="Whitworth, J. Bryan" />
<Property FormalName="CompanyPosition" Value="Chief Administrative Officer" />
<Property FormalName="CompanyPosition" Value="Executive Vice President" />
<Property FormalName="CompanyPosition" Value="General Counsel" />
</Property>
<Property FormalName="CompanyExecutive" Value="Carrig, John" />
<Property FormalName="CompanyPosition" Value="Chief Financial Officer" />
<Property FormalName="CompanyPosition" Value="Senior Vice President" />
<Property FormalName="CompanyPosition" Value="Treasurer" />
</Property>
</Property>
<Property FormalName="PrivateCompany" Value="Shell Oil Co." />
<Property FormalName="Competitor">
<Property FormalName="CompanyName" Value="BP p.l.c." />
</Property>
<Property FormalName="Competitor">
<Property FormalName="CompanyName" Value="Chevron Corp." />
</Property>
<Property FormalName="Competitor">
<Property FormalName="CompanyName" Value="Exxon Mobil Corp." />
</Property>
<Property FormalName="Headquarters" Value="Houston, Texas, United States of Ameri" />
<Property FormalName="Sector" Value="Energy" />
<Property FormalName="Industry" Value="Integrated Oil and Gas" />
</Property>

```

Content 'Enhancement' Rich Semantic Metatagging

Value-added relevant metatags added by Semagix to existing COMTEX tags:

- Private companies
- Type of company
- Industry affiliation
- Sector
- Exchange
- Company Execs
- Competitors



Customer Needs Driving Innovation

Horizontal Needs

Industry Needs

<p>Content Exploitation</p> <ul style="list-style-type: none"> • Understand and leverage siloed data • Increase worker productivity • Better KM across enterprises 	<p>Knowledge Discovery</p> <ul style="list-style-type: none"> • Access/leverage universe of data • More accurate competitive/threat assessment 	<p>Competitive Advantage</p> <ul style="list-style-type: none"> • Outmaneuver competitors • Improve enterprise decision making • Less damage control
<p>AML</p> <ul style="list-style-type: none"> • Comply with current/future regulations • Ensure broker/trade compliance • Reduce risks and costs • Enhance CRM 	<p>Homeland Security</p> <ul style="list-style-type: none"> • Improve intelligence gathering/analysis • Enable information sharing/preserve security • Create effective first responder programs 	<p>Pharmaceuticals</p> <ul style="list-style-type: none"> • Represent/update known data • Expedite drug discovery process • Enhance speed-to-market • Reduce redundancy

- Pharmaceuticals: intelligent literature search/mining, drug discovery
- Government and Intelligence
- Glycomics: semantic annotation of scientific (e.g., mass spectroscopy) data, complex processes with bioinformatics web services

VideoAnywhere and Taalee Semantic Search Engine (2000)

BLENDING BROWSING & QUERYING INTERFACE

Targeted e-shopping/e-commerce

ATTRIBUTE & KEYWORD QUERYING

- All Assets
- Business
- Movie Assets
 - Film Festivals
 - Interviews
 - Movies
 - Reviews
- News
- Travel

SEMANTIC BROWSING

uniform view of worldwide distributed assets of similar type

Details

Title	Alien-Resurrection
Director	
Actors	Sigourney Weaver Winona Ryder
Contents	The experiments begin. An unholy combination of human and alien genetics, made possible by an uneasy alliance between a renegade band of smugglers and a zealous cadre of scientists and officials. One subject is familiar -- a woman horrifically linked to the alien species that now elicits so much scrutiny. Ripley is back and all is not what it seems... Alien Resurrection stars Sigourney Weaver as Ripley, the role she originated in Ridley Scott's 1979 classic, Alien, and reprised in James Cameron's Aliens (for which she received an Academy Award nomination)

Search Results

Title	Details	View	Buy It
Alien-Resurrection			
Alien Resurrection			
Alien Resurrection			
Alien Resurrection			
Alien Resurrection			


assets access

Blended Semantic Browsing and Querying (Intelligence Analyst Workbench): 2002

BBC News | Middle East | Egypt sentences militants to death - Micro...

File Edit View Favorites Tools Help

Most of the defendants were said to belong to the armed Islamic movement **Al Jihad**, and **some of them are** known to be closely associated with Osama Bin Laden, the Saudi dissident accused by the United States of masterminding attacks on American embassies in east Africa last year.



All nine of the condemned men are living as fugitives outside Egypt. They include the Jihad leader, Ayman al **Zawahri**, who is thought to be in Afghanistan and his brother, **Mohammed**.

Eleven others, convicted of conspiring to overthrow the government, received life sentences with hard labour.

Some of the men have known links with Osama Bin Laden. They included several militants extradited from other countries.

Ahmed Ibrahim al Naggar, one of 12 handed over to Egypt from Albania last summer, wore the red clothing of a condemned man because he had been sentenced to death in an earlier trial for plotting an attack on Cairo's Khan al-Khalili market - a major tourist attraction

Internet

Knowledgebase Browser: Back

bin Laden, Osama

Classifications:
personInTheNews

Aliases:
 Bin Laden
 Osama

bin Laden, Osama memberOf organization:
 Al Qaeda

Search Terms: bin Laden Osama

Page: 1 2 3 4 5 6 7 8 9 10 [Next](#) [Next 10 pages](#)

Precise Matches:

- Blair predicts Bin Laden killing**
The likely outcome of US-led military action in Afghanistan is the death of 11 September suspect Osama Bin Laden, the UK... 01:33 10/25/2001 BBC
- Carmen bin Laden**
In an exclusive interview Carmen bin Laden, the wife of an older brother of Osama bin Laden, speaks frankly about her... 01:42 10/25/2001 ABC News
- Rumsfeld: 'We'll get Bin Laden'**
US Defence Secretary Rumsfeld says his comments to a newspaper that Bin Laden may never be found have been overstated. 01:41 10/25/2001 BBC
- Anthrax detected at facility handling White House mails**
Anthrax has been detected at the mail facility that handles mail for the White House. 02:47 10/24/2001 Channel News Asia
- Kuwait fears Bin Laden funding link**
The BBC's Frank Gardner reports on alleged links between some Kuwaiti charities and Osama Bin Laden's al-Qaeda network. 01:14 10/23/2001 BBC
- US steps up attacks on Taliban, as winter nears**
With winter barely three weeks away, the US is escalating its air attacks on Afghanistan's ruling Taliban, hoping to topple... 01:49 10/23/2001 Channel News Asia
- British special forces to help US find Osama: report**
Britain has agreed to a US request to deploy all its Special Air Service (SAS) forces to try and help smoke out terrorist... 02:02 10/21/2001 Channel News Asia
- Eliminate Bin Laden, CIA told**
President Bush has given the CIA sweeping powers to remove the threat from the al-Qaeda terrorist network, it is revealed. 01:08 10/21/2001 BBC
- Eliminate Bin Laden, CIA told**
President Bush has given the CIA sweeping powers to remove the threat from the al-Qaeda terrorist network, it is revealed. 01:41 10/21/2001 BBC
- Bin Laden 'received UN cash'**
A BBC investigation reveals the UN once funded a charity suspected of being a front for Osama Bin Laden. 03:10 10/20/2001 BBC

Page: 1 2 3 4 5 6 7 8 9 10 [Next](#) [Next 10 pages](#)

SEMAGIX

Person:

Location:

Organization:

Other:

Item Details: Back

Kuwait fears Bin Laden funding link
The BBC's Frank Gardner reports on alleged links between some Kuwaiti charities and Osama Bin Laden's al-Qaeda network.

[View Complete Story](#)

Click to play:
 112K

People: [bin Laden, Osama](#)

Organizations: [Al Qaeda](#)

Location: [Kuwait, Middle East](#)

Media: video

Produced by: BBC

Length: 01:14

Visualizer Content: BSBQ Application

The screenshot displays a web application interface with a search bar containing 'procrit' and a 'Find Entity' button. The search results list several articles, with the selected article being 'Epoetin alfa: current and future indications and nursing implications'. A 'Content Details' window is open over this article, showing the following information:

- Title:** Epoetin alfa: current and future indications and nursing implications.
- Text:** Cancer-related anemia commonly is associated with fatigue and
- Authors:** Buchsel, Patricia C; Murphy, Barbara J
- Side Effects:** fatigue
- Drug Class:** recombinant hormone
- Drugs:** Epoetin Alfa; Procrit
- Companies:** Ortho Biotech Products, L.P.
- Hormones:** erythropoietin
- Symptoms:** fatigue
- producer:** PubMed

The background shows a search results list with the following items:

- FDA notifications. Watch out for counterfeit Procrit, 2 lot (no description available)
- Darbepoetin alfa administered
- Erythropoietic agents as neu
- Epoetin alfa: current and futu
- Pure Red-Cell Aplasia and Re
- Role of oral versus IV iron su
- Erythropoietin (Procrit; Epog
- Role of iron in optimizing res

At the bottom left, there is a logo for SEMAGIX with the tagline 'POWER • THROUGH • RELEVANCE'. On the right side, there is a partial view of a network diagram with nodes and labels such as 'therapy related)', '(chronic disease)', 'fat', 'of breath', 'male appe', and 'blood cell count in'.

Semantic Information Integration in *Portals*

The screenshot shows the KnowledgePlex website interface. At the top, there's a search bar and navigation links. The main header features the KnowledgePlex logo and the tagline "The professional resource for affordable housing and community development".

your focus sidebar (right):

- Welcome, **Yash Warke**
- Topic: **Workforce Development**
- Role: **Public**
- Other topics from your profile: **Best Practices & Models, Public Housing, Health Care / Child Care**
- Locales: **GA**
- Roles: **Public**

Workforce Development main section:

This page summarizes KnowledgePlex content about **Workforce Development**.

Workforce Development is a general term to describe various efforts to improve the work and work-readiness skills of all citizens. Workforce development efforts are often referred to generically as "Job Training" or "Jobs" programs, and often administered by single-purpose employment organizations or by multi-purpose community-improvement organizations as a part of larger community economic development programs.

documents section:

Document Title	Author	Date
HUD's Daily Focus	www.hud.gov	Sep 15 2003
Some Tips on Collecting Data	www.nhi.org	Sep 14 2003
Fundraising: Multiple Mail Appeals	www.nhi.org	Sep 14 2003
Profile of a Healthy Fundraising Program, by Kim Klein	www.nhi.org	Sep 14 2003
Trends in Naturalization	Fix,Passel,Suher	Sep 17 2003

news section:

Current news on **Workforce Development**.

News Title	Author	Date
Trends in Naturalization	Fix,Passel,Suher	Sep 17 2003
Families that Receive Government Benefits after Leaving Welfare Are Less Likely to Return	Urban Institute	Sep 11 2003
Half of Welfare Recipients without Barriers are Working	Urban Institute	Sep 11 2003
Half of Former Welfare Recipients Are Working or Worked Recently	Urban Institute	Sep 11 2003

discussion section:

Discussion Title	Date
New Sources Of Development Finance? Affordable Housing Development & Finance	Sep 11 20:29
Success In Financing Commercial Deals Economic Revitalization	Sep 11 20:30
Immigration And Fair Housing Complaints Fair Housing	Sep 11 20:32
Foster Children And Support Housing Homelessness	Sep 11 20:34
Low Interest Rates And Mortgage Lending Abuses Homeownership / Mortgage Markets	Sep 11 20:36
Smart Growth Partnerships Land Use & Housing Planning	Sep 11 20:39
Accountability For Non-profits Organizational Development	Sep 11 20:40
Ida Program Growth And Success Personal Finance & Asset Creation	Sep 11 20:42
Future Of Hope Vi Public Housing	Sep 11 20:44
Arts & Cultural Program Examples Social / Comprehensive Development	Sep 11 20:45

calendar section:

Calendar Item	Date
Sidebar to Cover Story	Sep 15 2003
Pacific Hotel (Seattle, WA)	Sep 15 2003
Anderson Gardens	Sep 15 2003
Insight	Sep 15 2003
USDA Hopes to Fund Pilot Agriculture Innovation Centers in 2003	Sep 15 2003

Left Sidebar (topics):

- All Topics**
- Affordable Housing Development & Finance**
 - Housing Preservation / Expiring Use
 - Multifamily Housing
 - Property Management
 - Senior / Special Needs Housing
 - Single Family Housing
- Economic Revitalization**
 - Business Recruitment & Retention
 - Commercial Development & Finance
 - Historic Preservation
 - Microenterprise
 - Workforce Development**
- Fair Housing**
- Homelessness**
 - Best Practices & Models
 - Development & Finance Programs
- Homeownership / Mortgage Markets**
 - Homebuyer Assistance
 - Mortgage Lending / Predatory Lending
- Land Use & Housing Planning**
 - Environmental Issues
 - Gentrification & Preserving Affordability
 - Housing Market Trends & Studies
 - Landbanking
 - Planning & Zoning
 - Smart Growth
 - Transportation & Parking
- Organizational Development**
 - CDC Capacity
 - Leadership / Skills Training
 - Measuring Success / Impact
 - Organizational Effectiveness & Performance
- Personal Finance & Asset Creation**
 - Financial Literacy
 - Financial Strategies & Tools

User profile as a context for semantic integration of diverse yet relevant content

Semantic integration and presentation of various types of personalized content items in one place

Sample content item that is explicitly or implicitly associated semantically to facets in user profile

Equity Research Dashboard with Blended Semantic Querying and Browsing

Automatic 3rd party content integration

Motorola, Inc. 2:36:22 PM EDT

Symbol	Change	Price	Volume
MOT	+0.61	15.53	8,935,500

Equity Indices

DOW	8,570.15
NASDAQ	1,493.15
S&P 500	1,002.75

Company News

WebLink Wireless Reveals Text 2 Voice...	09/24/2001	COMTEX
Insignia's Jeode PDA Edition To Be In...	09/24/2001	COMTEX
Altera Teams Up with Virginia Tech Re...	09/19/2001	COMTEX

Analysis News

CSFB sees quality in Qualcomm	09/10/2001	CBS Marketwatch
Merrill Downgrades Motorola to 'Near...	09/06/2001	BusinessWeek Online
Motorola	09/06/2001	ON24

Earnings News

EXPANSION: OGLVYINTERACTIVE ESPANA F...	09/18/2001	COMTEX
LES ECHOS: STMICROELECTRONICS LOOKS F...	09/10/2001	COMTEX
Motorola Reduces Third-Quarter Sales ...	09/06/2001	Bloomberg

Industry and Competition News

Techs keep falling	09/05/2001	CNNFN
Hot Stocks: Federated, May Department...	07/05/2001	CNNFN
Landis acquires QUAYONE	07/02/2001	COMTEX

Market Commentary News

Volatility buffets telecom sector	09/21/2001	CBS Marketwatch
PCs, chips plunge; storage pushes up	09/17/2001	CBS Marketwatch
Workers Return; Not Business as Usual	09/12/2001	CNBC

Mergers & Acquisitions

Platinum Equity Acquires Multiservice...	09/04/2001	COMTEX
Motorola Sees Job Cuts as Chip Lines Cut	08/15/2001	CNBC
A Novo Broadband Signs Binding Agree...	08/14/2001	COMTEX

Resources for Motorola, Inc.

- TheStreet.com: Key Stats | Income | Cash Flow | Broker Rating
- WIZARD: SEC Financials
- Market Guide: Snapshot | Analyst Corner | Company Profile
- sage: Message Boards
- Zacks: Analyst Recommendations
- MarketGauge: Price and volume action

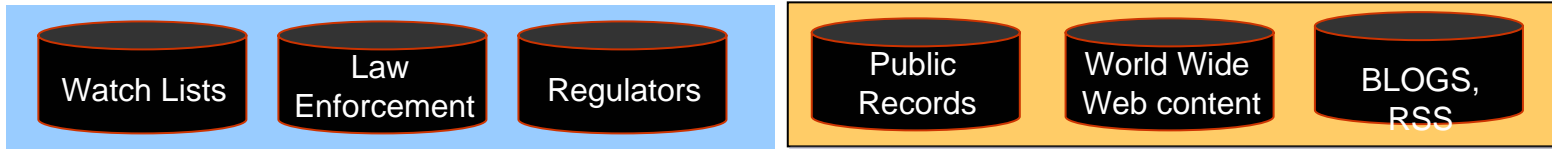
Focused relevant content organized by topic (*semantic categorization*)

Related relevant content not explicitly asked for (semantic associations)

Automatic Content Aggregation from multiple content providers and feeds

Competitive research inferred automatically

Semantic Integration of Heterogeneous Data for AML application



Structure Government Data

Un-structure text, Semi-structured Data

Establishing New Account

A screenshot of a web form titled 'Establishing New Account'. It includes fields for Name, Address, Date of Birth, and other personal information. There are also checkboxes for 'I am Human' and 'I am Bot'.

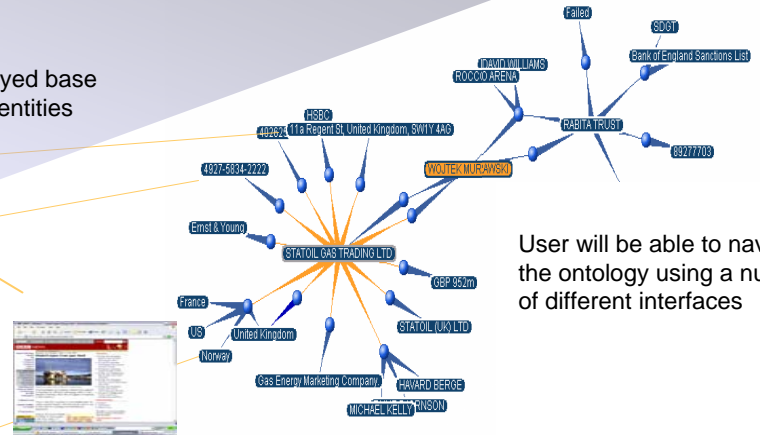
Scores the entity based on the content and entity relationships

A screenshot of a 'Customer Identification and Risk Assessment' report for Mohammed Hussein Al-Amoudi. The report includes a summary table with columns for 'Entity', 'Result', and 'Score'.

Entity	Result	Score
© Verification (ChoicePoint)	na	0
Office of Foreign Assets Control (OFAC)	Watch (Indirect)	35
Australian Department of Foreign Affairs & Trade (DFAT)	No Match	0
Bank of England (BoE)	Watch (Indirect)	35
De Nederlandsche Bank (DNB)	No Match	0
European Union (EU)	No Match	0
State Secretariat for Economic Affairs (SECO)	No Match	0
SWIFT Financial Control	Watched	20
WatchCheck Content	No Match	0
Media Search (Lexis/Nexis or Factiva)	No Match	0
Sanction SDN List	No Match	0
ESB	Watch (Indirect)	35
FAIR	Watch (Indirect)	35

Snippets is displayed base on best match to entities

Two screenshots showing search results and snippets for Mohammed Hussein Al-Amoudi. The top screenshot shows a list of search results with columns for 'Entity', 'Result', and 'Score'. The bottom screenshot shows a detailed view of a search result with a snippet of text.



User will be able to navigate the ontology using a number of different interfaces

CIRAS KYC Check

Individual Check - Microsoft Internet Explorer

File Edit View Favorites Tools Help

CIRAS

Customer Identification and Risk Assessment

User: Larry Parker | [Sign Off](#)

[Check Individual](#) **Check Organization** [Status & Results](#) [Reporting](#)

LMZ Shipping, LLC ENTITY Score: 65 | [Print \(PDF\)](#)
Brief explanation of score...

Check ID: 14092 Status: Pending — [Cancel or Refer this Record](#)

Summary | [Match Details](#) | [Media Content](#) | [Attributes](#) | [Associations](#) | [WebFountain](#)

Check	Result	Score
ID Verification (ChoicePoint)	Verified	0
Office of Foreign Asset Control (OFAC)	▲ Match	20
Australian Department of Foreign Affairs & Trade (DFAT)	No Matches	0
Bank of England (BOE)	▲ Match	0
De Nederlandsche Bank (DNB)	No Matches	0
European Union (EU)	No Matches	0
State Secretariat for Economic Affairs (SECO)	No Matches	0
CPB Global Name Database ("Swiss Check")	No Matches	0
WorldCheck Content	▲ Match	--
Media Search (Lexis/Nexis or Factiva)	No Matches	--
Section 326 List	No Matches	0

[Perform a New Check](#)

Visualizer Content: BSBQ Application

The screenshot displays the Semantic Visualizer interface within a Microsoft Internet Explorer browser window. The main search area shows 'ENRON' entered in the search box, with 'Find Entity' and 'Instances' buttons. A network graph on the right side of the page shows nodes for 'Belden, Timothy', 'J.P. Morgan Chase & Co.', 'Citigroup Inc.', and 'Securities Exchange Commission'. A 'Content Details' window is open, showing information for 'ANSCHUTZ, Philip F.'.

SEMAGIX
POWER • THROUGH • RELEVANCE

Semantic Visualizer

Help

CITIGROUP INC.

Synonyms

CITIGROUP INC.

Classifications

Organization

Attributes

name	value
Countries	USA
Location	New York
WC Category	CORPORATE
Display Name	CITIGROUP INC.

Related Content

- ANSCHUTZ, Philip F.
- CITIGROUP INC.
- REED, John Shepard
- WELL, Sanford I
- CITIGROUP GLOBAL MARKETS INC.

Content Details

Title: ANSCHUTZ, Philip F.

Description: Sept - 2002 charged by NY Attorney General with illegally profiting from the now banned practice of spinning in 57 IPOs arranged by Citigroup. May - 2003 agreed to pay \$4.4 million to settle charges without admitting or denying guilt. Californian Supreme Court has rejected his bid to stop an action brought by the California State Teachers Retirement System, which claims to have lost \$200m in stock and bonds invested in Qwest, an Anschutz firm. He is now likely to face trial and could be held personally liable in securities fraud case. He also

WC Category: INDIVIDUAL

Entity: WC33640

URLs: <http://albany.bizjournals.com/albany/stories/2003/05/12/daily2> Go

Java Applet Window

Copyright © 2004 Semagix Ltd.

CIRAS KYC Check Associations

Individual Check - Microsoft Internet Explorer

File Edit View Favorites Tools Help

CIRAS
Customer Identification and Risk Assessment

User: Larry Parker | [Sign Off](#)

[Check Individual](#) **Check Organization** [Status & Results](#) [Reporting](#)

LMZ Shipping, LLC ENTITY Score: 65 | [Print \(PDF\)](#)
Brief explanation of score...

Check ID: 14092 Status: Pending — [Cancel or Refer this Record](#)

[Summary](#) | [Match Details](#) | [WebFountain](#) | [Media Content](#) | [Attributes](#) | **Associations**

[Launch Associations Visualizer](#)

Is Related to	Wojtech Moroski Wojtech Moroski is related to Rabbita Trust Rabbita Trust appears on FBI Watch List
Undertakes	12 Thompson Av, Athens, Georgia
Is related to	Wojtech Moroski
Active in	Washington
Active in	US

[Perform a New Check](#)

SEMAG!X
POWER • THROUGH • RELEVANCE

[Terms, conditions, caveats and small print](#)
© Copyright 2004. Semagix. All Rights Reserved.

View Risk Scores for a specific company or customer

Transaction and Customer List - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media

CIRAS 10:43:56 8/1/2003 View Ontology

Client Information

Company: STATOIL GAS TRADING LTD
Individual:
Location:
Other:

Risk Score

STATOIL GAS TRADING LTD 65
Aggregate 65

Accept Reject

Company Knowledge

STATOIL GAS TRADING LTD [Compa
Synonyms:
Statoil Gas Trading
Relationships:
HAVARD BERGE works for
RUNE BJORNSON works for
MICHAEL KELLY works for
WOJTEK MURAWSKI is a shareholder in STATOIL GAS TRADING LTD
STATOIL GAS TRADING LTD is audited by Ernst & Young

Score Breakdown - Details

Score Component	Score	Reason
shareholder check	65	has a shareholder WOJTEK MURAWSKI who works for RABITA TRUST which appears on Bank of England Sanctions List
shareholder check	65	has a shareholder WOJTEK MURAWSKI who works for RABITA TRUST which appears on Bank of England Sanctions List

Aggregate Score: 65

Know Your Customer Check
Retrospective Check
Application Date: 29/01/2003
Request Outcome: Failed

Retrospective Client Check All Transactions Investigate Client Investigate Transaction New

Done Internet

start > RE: WF conten... CD Drive (D:) Camtasia Internet Expl... Adobe Acrobat - ... 100% 10:43 AM Friday 8/1/2003



European Law Enforcement Agency – Case Study

To identify and target prolific offenders by creating a holistic view of crime incident data and related entities

Needs / Business Requirements

- Merge and link case data from multiple sources using effective identification and disambiguation of entities associated with cases
- Link cases to a taxonomy of modus operandi based on analysis of unstructured textual information such as witness statements and crime reports
- Ability to use pre-defined or investigation-specific case profiles for search and match, with matching determined by configurable scoring of case attributes and associated entities
- Positive and negative searching of cases based on presence /absence of key characteristics or behaviours
- Ability to explore and browse case data starting from any specific associated entity via link analysis and intelligent entity identification and annotation in supporting unstructured textual documents

System Requirements

- Several hundred users
- Integrate with enterprise single sign-on systems

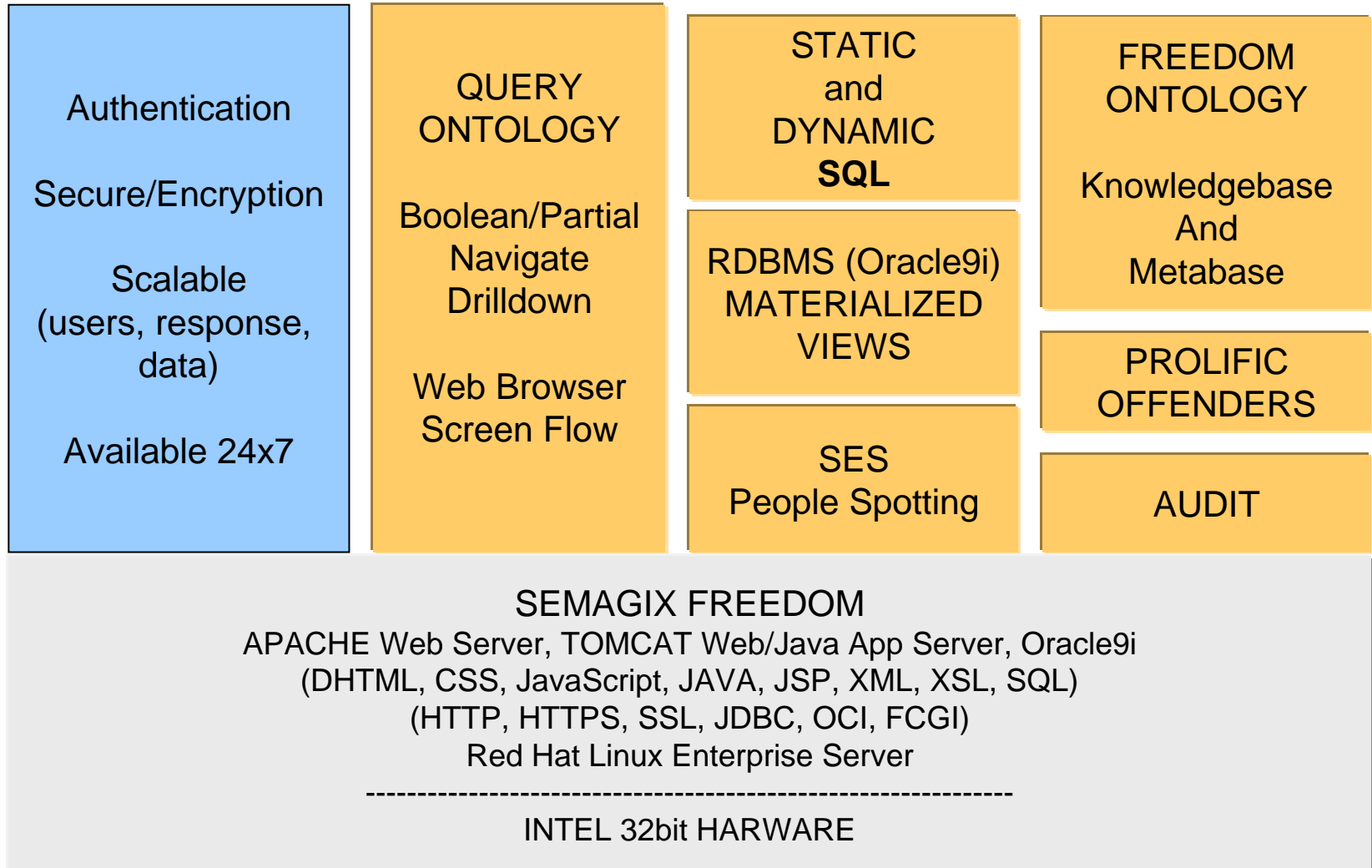
Technical Challenges

- Handle extremely large scale datasets – in excess of 10 million incidents – with a very high rate of daily incremental update
- Enforce visibility and privacy rules appropriate to the user and the available data sets
- Identify case behaviours / modus operandi in unstructured text

Results

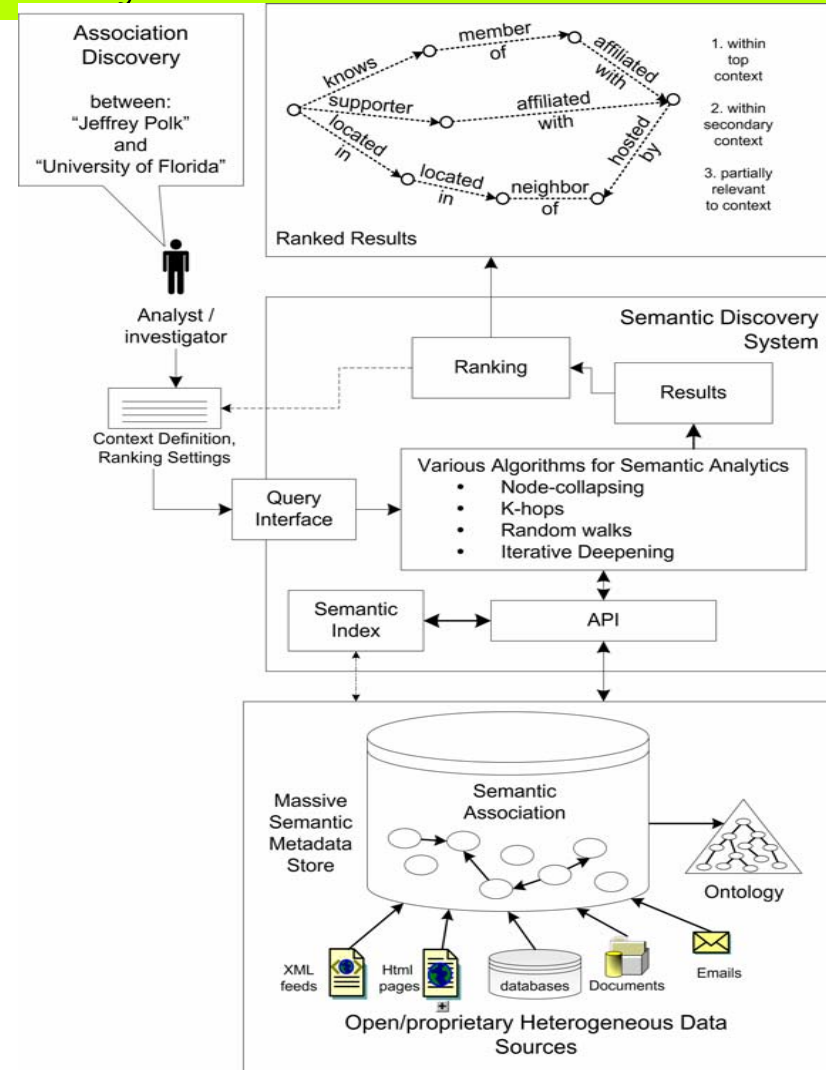
- Superior and more timely identification of prolific officers
- Better prioritization of cases
- Greater investigator productivity and effectiveness

European Law Enforcement Agency – Architecture Overview



SemDIS Prototype: Relationship Discovery

- SemDIS project
 - Discover relationships between entities in semantic knowledge bases
 - For prototype, use a subset of SWETO as a dataset






Turn Ranking On

Semantic Association Query Engine - Microsoft Internet Explorer

File Edit View Favorites Tools Help


Back Forward Stop Home Search Favorites Media Print Mail News RSS

Address <http://vader.cs.uga.edu:8080/semdis/matches> Go Google Links



LSDIS

Large Scale Distributed Information Systems



University of Georgia
Computer Science Department

Select Entities of Interest

More than one class found.
Select best match from list below:

Entity one:

Chee-Keng Yap
Yap

One class found.

Entity two:

Ravi Ramamoorthi

Enter a k value for the search:

Ranking on:


DoneInternet




Configure Ranking Schema with Context in 'Colleges' and 'Academic Depts.'

Semantic Association Query Engine - Microsoft Internet Explorer

Address: http://vader.cs.uga.edu:8080/semis/search?psel1=http%3A%2F%2Fisd.cs.uga.edu%2Fproj%2Fsemis%2Ftestbed%2F%23SWEET_2150038psel2=http%3A%2F%2Fisd.cs.uga.edu%2Fproj%2Fser



Large Scale Distributed Information Systems



University of Georgia
Computer Science Department

Configure Ranking Criteria

Context Specification

Finish! Locality

rdf:subClassOf

faculty member at>

Academic Department

University

Computer Science Researcher

Researcher

listed author in>

Expand Node

Hide Node

Select Node

Add to Region 'College Depts.'

k:

Subsumption Adjustment

Trust Adjustment

Rarity Adjustment

Popularity Adjustment

Association Length Adjustment

Favor Rare Associations:

Favor Popular Associations:

Favor Long Associations:

Submit Query

k:

k:

k:

k:

k:

Applet com.touchgraph.linkbrowser.LinkBrowserApplet started


Internet




Returns Ranked Results

Semantic Association Query Engine - Microsoft Internet Explorer

Address: <http://vader.cs.uga.edu:8080/semdis/franker>

**LSDIS**
Large Scale Distributed Information Systems


University of Georgia
Computer Science Department

Associations Found

Results 1 - 10 of 47. Search took: 4.389 seconds

Association	Ranking Score	Context	Association Length	Subsumption	Trust	Rarity	Popularity
1. Chee-Keng Yap <small>†faculty_member_at: New York University Department of Computer Science †has_academic_department: New York University †located_in: New York †located_in: Columbia University †has_academic_department: Columbia University Department of Computer Science †faculty_member_at: Ravi Ramamoorthi</small>	0.3756270205836345	█	█		█	█	█
2. Chee-Keng Yap <small>†listed_author_in: Refinement Methods for Geometric Bounds in Constructive Solid Geometry. †published_in: ACM Trans. Graph. †published_in: Frequency space environment map rendering. †listed_author_in: Ravi Ramamoorthi</small>	0.2538365896668301		█		█	█	█
3. Chee-Keng Yap <small>†listed_author_in: Refinement Methods for Geometric Bounds in Constructive Solid Geometry. †published_in: ACM Trans. Graph. †published_in: Chromium: a stream-processing framework for interactive rendering on clusters. †listed_author_in: Ren Ng †listed_author_in: All-frequency shadows using non-linear wavelet lighting approximation. †listed_author_in: Ravi Ramamoorthi</small>	0.25343627662676194		█		█	█	█
4. Chee-Keng Yap <small>†listed_author_in: On k-Hulls and Related Problems. †published_in: SIAM J. Comput. †published_in: Ranking Algorithms: The Symmetries and Colorations of the n-Cube. †listed_author_in: Jay P. Fillmore †listed_author_in: Spherical averages and applications to spherical splines and interpolation. †published_in: ACM Trans. Graph. †published_in: Frequency space environment map rendering. †listed_author_in: Ravi Ramamoorthi</small>	0.2533669312668104		█		█	█	█
5. Chee-Keng Yap <small>†listed_author_in: On k-Hulls and Related Problems. †published_in: SIAM J. Comput. †published_in: On Backtracking: A Combinatorial Description of the Algorithm. †listed_author_in: Jay P. Fillmore †listed_author_in: Spherical averages and applications to spherical splines and interpolation. †published_in: ACM Trans. Graph. †published_in: Frequency space environment map rendering. †listed_author_in: Ravi Ramamoorthi</small>	0.2533669312668104		█		█	█	█
6. Chee-Keng Yap <small>†listed_author_in: Reversal Complexity. †published_in: SIAM J. Comput. †published_in: Ranking Algorithms: The Symmetries and Colorations of the n-Cube. †listed_author_in: Jay P. Fillmore †listed_author_in: Spherical averages and applications to spherical splines and interpolation. †published_in: ACM Trans. Graph. †published_in: Frequency space environment map rendering. †listed_author_in: Ravi Ramamoorthi</small>	0.2533669312668104		█		█	█	█
7. Chee-Keng Yap <small>†listed_author_in: Reversal Complexity. †published_in: SIAM J. Comput. †published_in: On Backtracking: A Combinatorial Description of the Algorithm. †listed_author_in: Jay P. Fillmore †listed_author_in: Spherical averages and applications to spherical splines and interpolation. †published_in: ACM Trans. Graph. †published_in: Frequency space environment map rendering. †listed_author_in: Ravi Ramamoorthi</small>	0.2533669312668104		█		█	█	█
8. Chee-Keng Yap <small>†listed_author_in: Precision-Sensitive Euclidean Shortest Path in 3-Space. †published_in: SIAM J. Comput. †published_in: Ranking Algorithms: The Symmetries and Colorations of the n-Cube. †listed_author_in: Jay P. Fillmore †listed_author_in: Spherical averages and applications to spherical splines and interpolation. †published_in: ACM Trans. Graph. †published_in: Frequency space environment map rendering. †listed_author_in: Ravi Ramamoorthi</small>	0.2533669312668104		█		█	█	█
9. Chee-Keng Yan <small>†listed_author_in: Precision-Sensitive Euclidean Shortest Path in 3-Space.</small>							

Done Internet

Ontology Quality

- Many real-world ontologies may be described as semi-formal ontologies
 - populated with partial or incomplete knowledge
 - may contain occasional inconsistencies, or occasionally violate constraints (e.g. all schema level constraints may not be observed in the knowledgebase that instantiates the ontology schema)
 - often ontology is populated by many persons or by extracting and integrating knowledge from multiple sources
 - analogy is “dirty data” which is usually a fact of life in most enterprise databases.

Ontology Representation Expressiveness

- Applications vary in terms of expressiveness of representation needed.
- Trade-off between expressive power and computational complexity applies both to knowledge creation/maintenance and to inference mechanisms for such languages. It is often very difficult to capture the knowledge that instantiates the more expressive constructs/constraints.
- Many business applications end up using models/languages that lie closer to less expressive languages.
- On the other hand, we have seen a *few* applications, especially in scientific domains such as biology, where more expressive languages are needed, and OWL-Full or FOL is not adequate.

Ontology Size / Population / Freshness

- Ontology population is critical. Among the ontologies developed by Semagix or using its technology, a median size of ontology is over 1 million instances/facts and relationship instances each (at least two have exceeded 10 million instances). This level of knowledge makes the system very powerful (as it is applied). Furthermore, in many cases, it is necessary to keep these ontologies current or updated with facts and knowledge on a daily or more frequent basis. Both the scale and freshness requirements dictate that populating ontologies with instance data needs to be automated.

Metadata Extraction

Large scale metadata extraction and semantic annotation is possible. IBM WebFountain [Dill et al 2003] demonstrates the ability to annotate on a Web scale (i.e., over 2.5 billion pages), while Semagix Freedom related technology [Hammond et al 2002] demonstrates capabilities that work for a few million documents per day per server. However, the general trade-off of depth versus scale applies. Storage and manipulation of metadata for millions to hundreds of millions of content items requires database techniques with the challenge of improving performance and scale in presence of more complex structures

Semantic Technology Building Blocks

- A vast majority of the Semantic (Web) Technology Applications that have been developed or envisioned rely on three crucial capabilities: **ontology creation, semantic annotation (metadata extraction) and querying/inferencing**. Enterprise-scale applications share many requirements in these three respects with pan Web applications. All these capabilities must scale to many millions of documents and concepts (rather than hundreds to thousands) for current applications, and applications requiring billions of documents and concepts have also been discussed (esp. in intelligence and government space) but not yet deployed.

Primary Technical Capabilities/Key Research Challenges

- Two of the most basic “semantic” techniques are “named entity identification”, and “semantic ambiguity resolution”. [It would be nice to have relationship extraction too.] A tool for annotation is of little value if it does not support ambiguity resolution. Both require highly multidisciplinary approaches, borrowing for NLP/lexical analysis, statistical and IR techniques and possibly machine learning techniques. A high degree of automation is possible in meeting many real-world semantic disambiguation requirements, although pathological cases will always exist and complete automation is unlikely.

Content Heterogeneity

- Support for heterogeneous content is key – it is too hard to deploy separate products within a single enterprise to deal with structured, semi-structured and unstructured data/content management. New applications involve extensive types of heterogeneity in format, media and access/delivery mechanisms (e.g., news feed in RSS, NewsML news, Web posted article in HTML or served up dynamically through database query and XSLT transformation, analyst report in PDF or WORD, subscription service with API-based access to Lexis/Nexis, enterprise's own relational databases and content management systems such as Documentum or Notes, e-mails, etc). Semi-structured data (XML-based data and RDF based metadata) is growing at an explosive rate.

Processing

- Semantic query processing with the ability to query both ontology and metadata to retrieve heterogeneous content is highly valuable. Consider “Give me all articles on the competitors of Intel”, where ontology gives information on competitors, supports semantics (with the understanding that “Palm” is a company and that “Palm” and “Palm, Inc.” are the same in this case), and metadata identifies the company to which an article refers, regardless of format of the article.
- Analytical applications could require sub-second response time for tens of concurrent complex queries over a large metadata base and ontology, and can benefit from further database research. High performance and highly scalable query processing techniques that deal with more complex representations compared to database schemas and with more explicit roles of relationships, is important. Have not found great use of DL reasoning.



Conclusion

- Great progress from work in semantic information interoperability/integration of early 90s until now, re-energized by the vision of Semantic Web, related standards and technological advances
 - Standards defined by W3C are very timely and are bringing some level of interoperability
- No longer an exotic technology
 - beyond proof of concept and now facing main stream engineering challenges
 - some industries very open to ontologies
 - in other industries applications hide semantic technologies



Conclusion

- Great progress from work in semantic information interoperability/integration of early 90s until now, re-energized by the vision of Semantic Web, related standards and technological advances
- Technology beyond proof of concept
- But difficult research and engineering challenges ahead
- Researchers should be mindful of state of the art commercial technologies and real applications

For Further Information

- Article in Data Engineering special issue on Making the Semantic Web Real (Dec. 2003)
<http://wwwt.semagix.com/documents/SemanticWebTechinAction.pdf>
- Commercial Technology: <http://www.semagix.com/download.html>
- Research: Semantic Association and Semantic Discovery Projects:
<http://lsdis.cs.uga.edu/proj/proj.html>
- Publications and Presentations: <http://lsdis.cs.uga.edu/lib/lib.html>