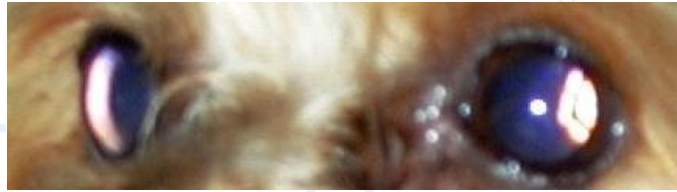


Semantic Technologies:

Providing Answers
Today



Vision of Semantic Search



Vision of semantic search is the availability of ***search tools*** that would improve retrieval

Internet: Answering questions using natural language. *e.g. What was wrong with Curt Schilling's shoulder?*

Enterprise: Content preparation and "intelligent indexing" for more efficient retrieval in vertical domains - life sciences, financial services or telecom *e.g. Categorizing (tagging), normalizing and de-duplicating all content related to adverse reactions for a particular drug formulation*

Reality of Semantic Search

- Problem – linguistic variations in concept expression
 - **Technology: natural language processing (NLP)**
- Problem – huge numbers of documents that are the same or versions of the same
 - **Technologies : text mining, text analytics, normalizing & de-duping**
- Problem – amount of content exceeds amount of human expertise to analyze & categorize
 - **Technologies : entity extraction, contextual analysis, auto-categorization**
- Problem – understanding trends and relative values expressed in content
 - **Technology : sentiment analysis**
- Problem – retrieving & federating contextually related and relevant content
 - **Technologies – All of the above**

Partial List of Semantic Companies at Work In the Enterprise and Out

- [Attensity](#)
- [Basis Technology](#)
- [Cambridge Semantics](#)
- [**Expert System**](#)
- [Cognition](#)
- [Concept Searching](#)
- [**Connotate**](#)
- [EasyAsk](#)
- [Information Extraction Systems](#)
- [**Lexalytics**](#)
- [**Linguamatics**](#)
- [Metatomix](#)
- [MuseGlobal](#)
- [Nstein](#)
- [Patterns & Predictions](#)
- [SAP:Inxight Federal Systems](#)
- [SAS: Teragram Linguistic Technologies](#)
- [**Sinequa**](#)
- [Smartlogic](#)
- [**TEMIS**](#)
- [TopQuadrant](#)

Case 1: I2E from Linguamatics

Used by: **Biogen Idec**

Need: Sifting through massive biomedical literature corpora to deliver **FACTS** in context to scientists

Use: Applying agile natural language processing (NLP) to **mine Medline**, other citations and full-text PDFs to **create a database of biomarkers; maintain currency of database content**

Benefits:

- **Minimize human curatorial effort**
- **Reduces time to retrieve** facts by 10X to 1000X depending on the project
- Added benefit – **entity tagging** of the database against their proprietary thesaurus (proteins, diseases, adverse events), enhancing its value

Case 2: Luxid from

Used by: **Thomson Reuters**

Need: Eliminate **or reduce manual tagging** for healthcare & scientific content by **providing intelligent contextual content analysis**

Use: To provide customer facing **topical navigation** to access published content

Benefits:

- **Tagging Efficiency** after rules and infrastructure are established
- Highly **efficient feedback loop** to help subject matter experts make adjustments to assigned controlled vocabulary
- Entity tagging providing opportunities for leveraging content **-sentiment & semantic analysis**

Caveat (s): Will always need editorial oversight for new topical areas of research

Case 3:



Used by: **Cormine Intelligent Data** for **WorldTech International**

Need: **Faster throughput for monitoring** and mining content from more than 2,000 Web sources

Use: **Data mining** embedded in Cormine's support of WorldTech customer facing Newton portal

Benefits:

- **Connotate "intelligent agents" detect** new & targeted content
- **Data mining** for relevant content
- **Collecting and archiving results** for delivery to Cormine's Zoogma intelligence platform for entity extraction, auto-categorization, duplication detection and normalization

Case 4: Cogito

from  **EXPERT SYSTEM**
SEMANTIC INTELLIGENCE

Used by: **Eni** (integrated energy company)

Need: **Speed up competitive intelligence filtering** to better understand oil & gas business risk & opportunity

Use: Cogito platform to **capture weak signals** of marketplace discontinuity and technology evolution

Benefits:

- Improve internal **information flow**
- Better **organization** of company's **knowledge assets** and better **insight into content** of information resources
- Providing targeted responses by **extracting contextually relevant meaning** from text documents



Case 5: ***Salience 4*** from **Lexa**LYTICS

Used by: **FT Search, Inc.**

Need: To provide **entity extraction** for **text analysis** to enhance search results for news feed customers

Use: **Newssift.com**

Benefits:

- **Extracting sentiment** and tone from news feed content
- **Revealing positive, neutral or negative sentiment** to users
- **Providing visual cues using graphs** to readers to engage them more actively in content analysis

Case 6: SINEQUA

Used by: **Atos Origin**

Need: To replace existing multiple embedded search applications with **federated search to improve relevancy, ranking, & performance**

Use: **Single-point intranet search** across all applications

Benefits:

- **Rapid deployment, improved results** with no tuning
- **Ease of adding new connectors** for new applications
- Very **fast** indexing and retrieval
- Impressive, intuitive **contextual navigation**

Thank you for listening !

**PLEASE CONTACT ME AT LYNDA@GILBANE.COM
FOR CLARIFICATION OR CONTACTS FOR PRODUCTS DESCRIBED**

<http://gilbane.com>



THE GILBANE GROUP

Readings: Semantic Technologies

Chickowki, Ericka. ***Understanding Semantic Web Technologies.*** Baseline, 08/05/2008, 4p.

Gilbane Group. Enterprise Search Blog on Semantic Search and Semantic Technologies

Miller, Ron. ***Semantic Search takes root in the enterprise.*** Medford, NJ: InfoToday: IN Enterprise search sourcebook 2008, 09/2008, pp. 42-45

Making data more usable through semantic search. DigitalNow Blog, 08/25/2008, 1p. Moulton comments on taxonomy and ontology.

The Meaning of Semantics Depends...On who you ask and why you are asking, Lynda Moulton, at Boston KM Forum, Symposium at Bentley University on "Semantics – the Next Frontier for Leveraging Knowledge", October 7, 2008