Managing Knowledge with Semantics

The “If we only knew what we know” – Syndrome

Brooke Aker
CEO Expert System
baker@expertsystem.net
W3C Oil & Gas Meeting
December 2008
Eni is active in 70 countries with a staff of about 76,000 employees.

Oil & Natural Gas production: 1,74 million boe/day

2007 Market capitalization: € 91.6 billion

2007 RESULTS
- consolidated revenues: € 87.2 billion
- net profit: € 10 billion
Expert System

Semantic Processing Platform - Cogito

- 100+ customers including large corporations, government in;
  - business intelligence - enterprise search & data extensibility
  - market sentiment - customer care

- 100+ dedicated engineers focused on core semantic technology, applications, tools and services:
  - 200 man/years in the development of COGITO over the last 10 years.

- 15 years old, private & profitable
  - FY2007: $12M, 115+ employees, 40% growth each of last 2 years
  - Offices in Connecticut, California, Florida, UK, Italy, & Germany
Knowledge is the key for competitiveness

TO BE FAST, EFFECTIVE AND EFFICIENT

THE MAIN REQUIREMENT

TO SEARCH FOR INFORMATION

THE FIRST TASKS

TO PROVIDE KNOWLEDGE

THE ULTIMATE IMPERATIVE
Semantic Search: the “mosaic” metaphor

“What will be the nuclear strategy of China in the next 20 years?”

Keyword search

Semantic search
Semantic Search: the “mosaic” metaphor

- main clause/subordinate clause
- subject-action-object (SAO)
- keywords - headwords
- related concepts
- name of people
- name of geographical site
- name of companies
- name of technologies
- phone numbers
- plate numbers
- investment figures
- date figures
- ......
Knowledge is **connection**, not collection

- Standard Information retrieval
- Semantic engine
- Semantic search
- Grocery List
  - Ezekiel 9:2 Bread
  - Kuchen
  - Tomatoes
  - English muffins
  - Spinach
  - Salmon
  - Fava beans
  - Cabbage
  - Brown rice
  - Tortillas
  - Asda fruit
  - Crackers
- Connection
- Keldosophy AI
  - Teaching
  - All pages
  - Personal
  - About KBS
  - Bookmarks
<table>
<thead>
<tr>
<th>Technology - Semantic Methods</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>4 Methods</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphological Analysis</td>
<td>understand word forms</td>
<td>dog, dogs, and dog-catcher are closely related</td>
</tr>
<tr>
<td>Grammatical Analysis</td>
<td>understand the parts of speech</td>
<td>&quot;There are 40 rows in the table&quot; uses rows as a noun, vs. &quot;She rows 5 times a week&quot; uses rows as a verb</td>
</tr>
<tr>
<td>Logical Analysis</td>
<td>understand how words relate to other words</td>
<td>&quot;Jeffrey Skilling, represented by Attorney Daniel Petrocelli, is married to Rebecca Carter&quot;. Rebecca is married to Jeffrey not Daniel.</td>
</tr>
<tr>
<td>Semantic Analysis (disambiguation)</td>
<td>understand the context of key words</td>
<td>&quot;I used beef broth for my soup stock&quot; uses stock in the context of food, vs. &quot;The company keeps lots of stock on hand&quot; uses stock in the context of inventory.</td>
</tr>
</tbody>
</table>
Technology - Semantic Methods

Morphology / POS

Categories

Disambiguation

Logic (semantic triples)
Technology Stack & Precision

Development Studio

Linguistic Query Engine

Semantic Network

1. Morphology
2. Grammatical
3. Logic
4. Disambiguation

Develop & Add Custom Rules

80%

90%+

English
Arabic
Italian
German
French
Other Middle Eastern
Technology - Semantic Network

COGITO® Semantic Network:
- 350,000 words
- 2.8m relationships
Semantic Development Environment
Use Case One

DISTILLING KEY INFORMATION

PROCESSING AND TRACKING KNOWLEDGE
Use Case One

Question: “What will be the nuclear strategy of China in the next 20 years?”

We will semantically explore this “corpus” of 2,220,000 documents.
<table>
<thead>
<tr>
<th>Keyword</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>nuclear</td>
<td>32,579</td>
</tr>
</tbody>
</table>

**Keyword: nuclear**

**Frequency: 32,579**
concept: “nuclear energy”
frequency: 3,332

concept: “nuclear power”
frequency: 3,384

- nuclear weapon frequency: 2.796
- nuclear arsenal frequency: 197
- nuclear attack frequency: 112
- nuclear medicine frequency: 13
<table>
<thead>
<tr>
<th>Place</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>26,161</td>
</tr>
</tbody>
</table>
How many times, in the same sentence with place: China frequency: 985
How many times, in the same sentence with

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Headword</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>nuclear</td>
<td>nuclear power</td>
<td>1286</td>
</tr>
<tr>
<td>china</td>
<td>china</td>
<td>1032</td>
</tr>
<tr>
<td>power</td>
<td>power</td>
<td>799</td>
</tr>
<tr>
<td>energy</td>
<td>energy</td>
<td>635</td>
</tr>
<tr>
<td>development</td>
<td>development</td>
<td>272</td>
</tr>
<tr>
<td>technology</td>
<td>technology</td>
<td>129</td>
</tr>
<tr>
<td>plants</td>
<td>plants</td>
<td>126</td>
</tr>
<tr>
<td>capacity</td>
<td>capacity</td>
<td>120</td>
</tr>
<tr>
<td>cooper</td>
<td>cooper</td>
<td>108</td>
</tr>
<tr>
<td>agreement</td>
<td>agreement</td>
<td>104</td>
</tr>
<tr>
<td>construction</td>
<td>construction</td>
<td>88</td>
</tr>
<tr>
<td>International</td>
<td>International</td>
<td>72</td>
</tr>
<tr>
<td>peaceful</td>
<td>peaceful</td>
<td>65</td>
</tr>
<tr>
<td>new</td>
<td>new</td>
<td>65</td>
</tr>
<tr>
<td>percent</td>
<td>percent</td>
<td>65</td>
</tr>
<tr>
<td>usa</td>
<td>usa</td>
<td>66</td>
</tr>
<tr>
<td>2013</td>
<td>2013</td>
<td>65</td>
</tr>
<tr>
<td>industry</td>
<td>industry</td>
<td>65</td>
</tr>
<tr>
<td>india</td>
<td>india</td>
<td>65</td>
</tr>
<tr>
<td>plan</td>
<td>plan</td>
<td>61</td>
</tr>
<tr>
<td>signed</td>
<td>signed</td>
<td>61</td>
</tr>
<tr>
<td>two</td>
<td>two</td>
<td>61</td>
</tr>
<tr>
<td>one</td>
<td>one</td>
<td>61</td>
</tr>
<tr>
<td>country</td>
<td>country</td>
<td>58</td>
</tr>
<tr>
<td>national</td>
<td>national</td>
<td>56</td>
</tr>
<tr>
<td>have</td>
<td>have</td>
<td>55</td>
</tr>
<tr>
<td>most</td>
<td>most</td>
<td>52</td>
</tr>
<tr>
<td>plans</td>
<td>plans</td>
<td>51</td>
</tr>
<tr>
<td>research</td>
<td>research</td>
<td>50</td>
</tr>
<tr>
<td>world</td>
<td>world</td>
<td>49</td>
</tr>
<tr>
<td>pakistan</td>
<td>pakistan</td>
<td>50</td>
</tr>
</tbody>
</table>

strategy

- Concept: scheme - Noun. A carefully devised plan ...
  - Freq.: 33

strategy, scheme - Noun. A carefully devised plan ...
- strength, force, specialty - Noun. An asset of spec...
  - Freq.: 5
- strength, potency, effectiveness - Noun. Capacity...
  - Freq.: 2
- strengthen, fortify - Verb. To make something ...
  - Freq.: 12
- strengthen, fortify - Verb. To make something ...
  - Freq.: 5
- strike - Verb. To enter someone's mind or occup...
  - Freq.: 1
- strong, strong - Adj. A strength or in ...
  - Freq.: 4
- structure - Noun. The composition of ...
  - Freq.: 2
- student, pupil, education - Noun. A learner who ...
  - Freq.: 4
- study, report - Noun. An amount described ...
  - Freq.: 8
- study, work - Noun. An amount intended to learn ...
  - Freq.: 4
- subject, matter, task, issue - Noun. Some situat...
  - Freq.: 4
Electricity consumption is growing fast
Environment protection is serious
Nuclear power is safe, economical and clean
China government has made an active strategy for nuclear power development to meet the situation
During 2006-2020,
- China will construct 30 new NPPs,
- power 40,000MWe,
- percentage 4%,
- capacity of fuel cycle facilities will also be expanded.
The virtuous funnel: from collection to connection

“nuclear energy” + China
Corpus of 2,220,000 documents

nuclear

6716 occurrences

32,579 “nuclear”

26,161 “China”

3384 nuclear power

3332 nuclear energy

33 sentences

33 sentences
Serendipity, maybe

unexpected concept: tar sand occurrences: 66
location: Canada, Alberta
occurrences: 41
announced C$ 230 million (US$ 200 million) investment over four years for researching clean
energy technologies including next-generation nuclear power. It also flagged the possibility of
using nuclear power for extracting oil from Alberta’s oil sands, reducing the substantial demand for
natural gas and cutting CO2 emissions from this.

Globe & Mail 18/1/07, CBC 18/1/07.

AUSTRALIA

Australian parliamentary report published
After 20 months working, the House of Representatives Standing Committee on Industry & Resources
has published a 732-page report Australia’s Uranium: Greenhouse Friendly Fuel for an Energy
Hungry World. Its Chairman said that the Committee had “achieved a unanimous and bipartisan
position on the need to remove all impediments to the further development of Australia’s uranium
resources. All members are agreed that present restrictions on uranium exploration and mining are
illogical, inconsistent and anti-competitive”. "Australia is uniquely placed to make a significant
contribution to emissions reductions through increased production and supply of uranium." He said
that "Australia should throw the world a climate lifeline through the expanded production and export
of this greenhouse-friendly fuel" since renewables and energy efficiency "alone have no prospect
whatevsoever of meeting rapidly-growing demands for energy and advancing greenhouse-gas emissions
to the degree required."

In respect to uranium exports to developing countries such as China: "As a matter of energy justice,
Australia should not deny countries who wish to use nuclear power in a responsible manner the
benefits from doing so." Parliament has now approved the bilateral treaty with China.
The knowledge funnel

"nuclear energy" + China
Corpus of 2,220,000 documents

32,579 "nuclear" occurrences
6716 occurrences
3384 sentences

nuclear power
nuclear energy
3332 occurrences

nuclear power
nuclear energy
Tar sand

nuclear power
nuclear energy
Alberta
Canada

34 sentences
3-Step Walk With Semantic Search

Oil as keyword = 399 results
3-Step Walk With Semantic Search

Oil as concept = 279 results
3-Step Walk With Semantic Search

Oil as a semantic concept = 322 results
3-Step Walk With Semantic Search

Oil as a semantic concept & objects = 38 results
Use Case Two

CAPTURING WEAK SIGNALS & UNDERSTANDING TECHNOLOGY EVOLUTION
Playing with a very popular “trio”

Subject
Action
Object
SAO structure identification can boost technical problem solving.

Subject: sponge
Action: cleans
Object: CAR

Solution

Problem
SAO structure identification can boost technical problem solving

**Problem**
- steam reforming
- bacteria
- catalyst reactor

**Solution**
- produce
  - electrolysis
  - HYDROGEN
Extracting knowledge through semantic queries on SAO’s

Find all the solutions

Set a problem

Find the connected problems
Capturing weak signals of "breakthrough"

- **Inject** criogenic liquids
- Ultrasonic stimulation
- Magnetic stimulation

- **Inject gas**
- Inject polymers

**Problem**

**Solution**

- Increase pressure
- **Enhance** OIL RECOVERY

- Radical Innovation
- Disruptive Technology
- Breakthrough
Conclusions
Main achievements from semantics implementation

- avoid high cost

- of wasting effort on poor content
  ✓ early selection of high quality information without “reading” text

- of not finding information
  ✓ management of unstructured information

- of unnecessary re-creation of content
  ✓ because you can “know what you know”
Thank you