EXPERT SYSTEM

OSINT made easy:

A Live Demo on How to Tackle Unstructured Data

Gian Piero Oggero
Director Strategic Accounts,
Intelligence Division
National Police Academy Prague 27-11-12

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

A FLOOD OF UNSTRUCTURED DATA & INFORMATION

More than 80% of the knowledge on which our daily jobs are based is unstructured (emails, documents, web pages, articles, information from social media, etc.).

- Over 294 Billion emails sent daily;
- Over 6.1 Trillion text messages sent in 2010;
- And what about phone, calls, faxes, chat sessions, etc.?

Sources: Radicati Group, ITU.

EVERY 60 SECONDS ON THE INTERNET

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

SCENARIOS OF INTEREST

Every organization must constantly monitor events, signals and situations that influence decision making.

THE LIMITS OF TRADITIONAL APPROACHES

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

Keyword Technology or Statistics

Breaks text into single words without considering the context, like reading a language that we don't understand:

Az IBM szokásosan nagy hangsúlyt helyez a továbbképzésre, így munkatársai évente számos szakmai tanfolyamon vesznek részt.

Shallow Linguistic Technology

Recognizes words and identifies their most basic forms (lemmas), but cannot distinguish between different meanings.

Sell -> Selling -> Sold

Neither understands the *meaning* of words.

KEYWORD TECHNOLOGY

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

No understanding of the real content within documents; for example, this sentence:

A new advanced search capability called Utopia will enable Web users to access information on the Internet in a whole new way -- providing them with the most credible search results possible through any existing technology

and this sentence:

Internet users A the new credible results enable capability Web to information through technology called on advanced will in a new access way -- providing them with Utopia most search the search possible existing any whole

are the same!

SHALLOW LINGUISTIC TECHNOLOGY

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

No understanding of the relation between elements and of the meaning; for example, this text:

The Dow fell 46.58, or 0.42 percent, to 11,002.14. The Standard & Poor's 500 index fell 1.44, or 0.11 percent, to 1,263.85, and the Nasdaq composite gained 6.84, or 0.32 percent, to 2,162.78.

and this text:

The Dow gained 46.58, or 0.42 percent, to 11,002.14. The Standard & Poor's 500 index fell 1.44, or 0.11 percent, to 1,263.85, and the Nasdaq composite fell 6.84, or 0.32 percent, to 2,162.78.

are the same!

WHY WE ARE DIFFERENT

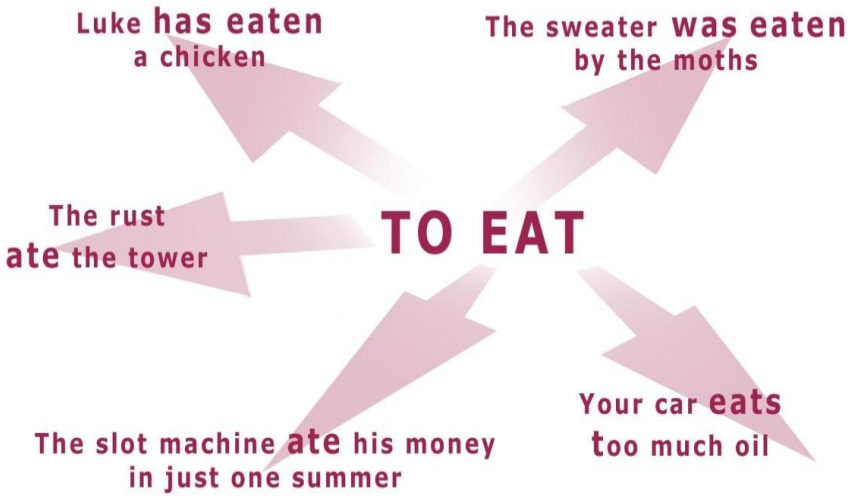
Semantic technology understands the meaning of words in the same way you learned to read.

- **It understands the relationships between words.**

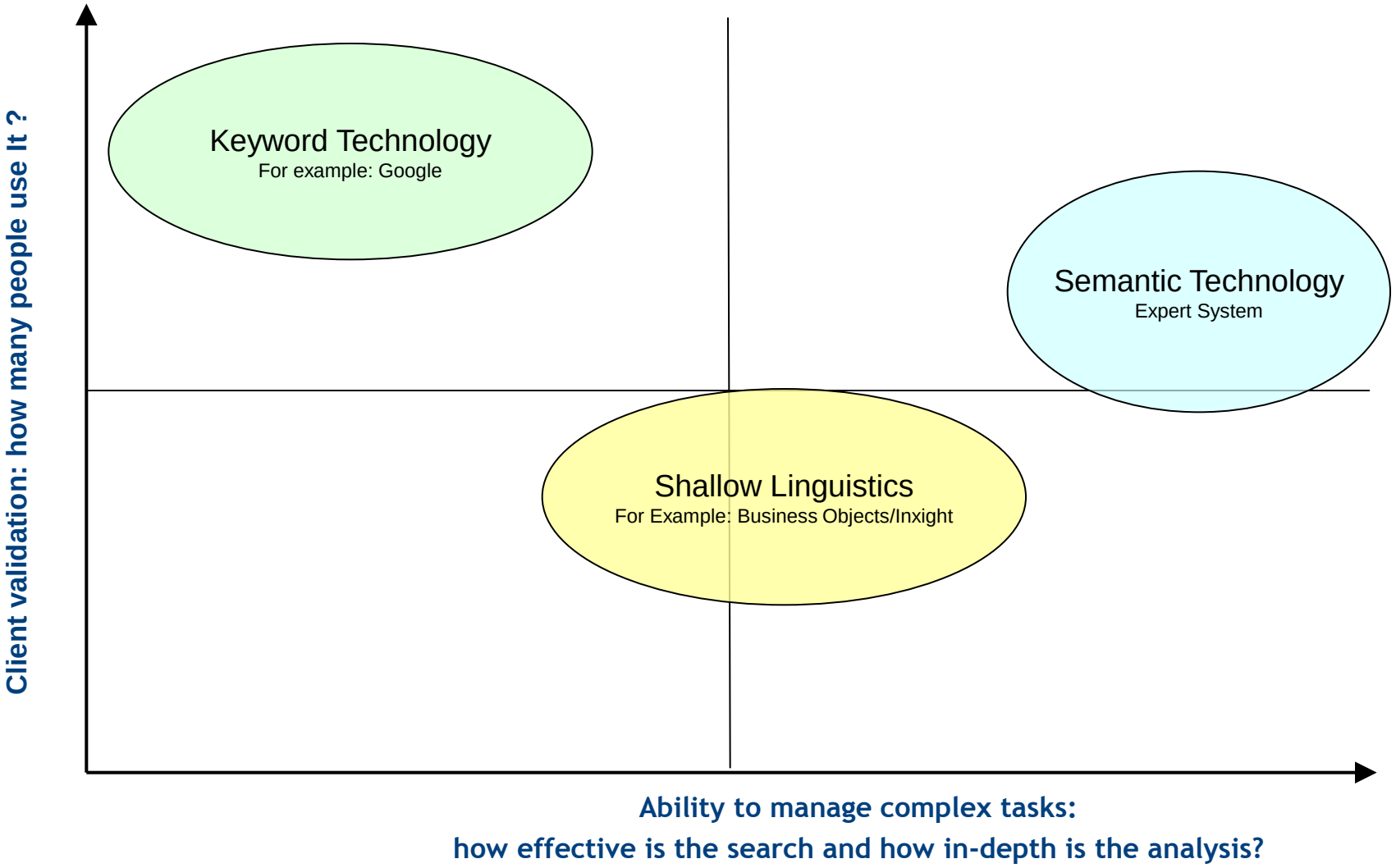
  Luke (subject) has eaten (verb) a chicken (object).

- **It understands the meaning of words.**

  To eat (chicken); to consume (oil); to destroy (sweater); to spend (money); to rust (the tower), etc.



Three technologies compared



DISAMBIGUATION: EXAMPLE

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

The original text:

(Reuters) - Prime Minister Benjamin Netanyahu's U.N. speech about Iranian nuclear advances has dampened speculation in Israel that he could order a war this year.

Analyzing Thursday's address in which Netanyahu literally drew a "red line" on a cartoon bomb to show how close Iran was to building nuclear weaponry, commentators saw his deadline for any military action falling in early 2013, well after U.S. elections in November and a possible snap Israeli poll.

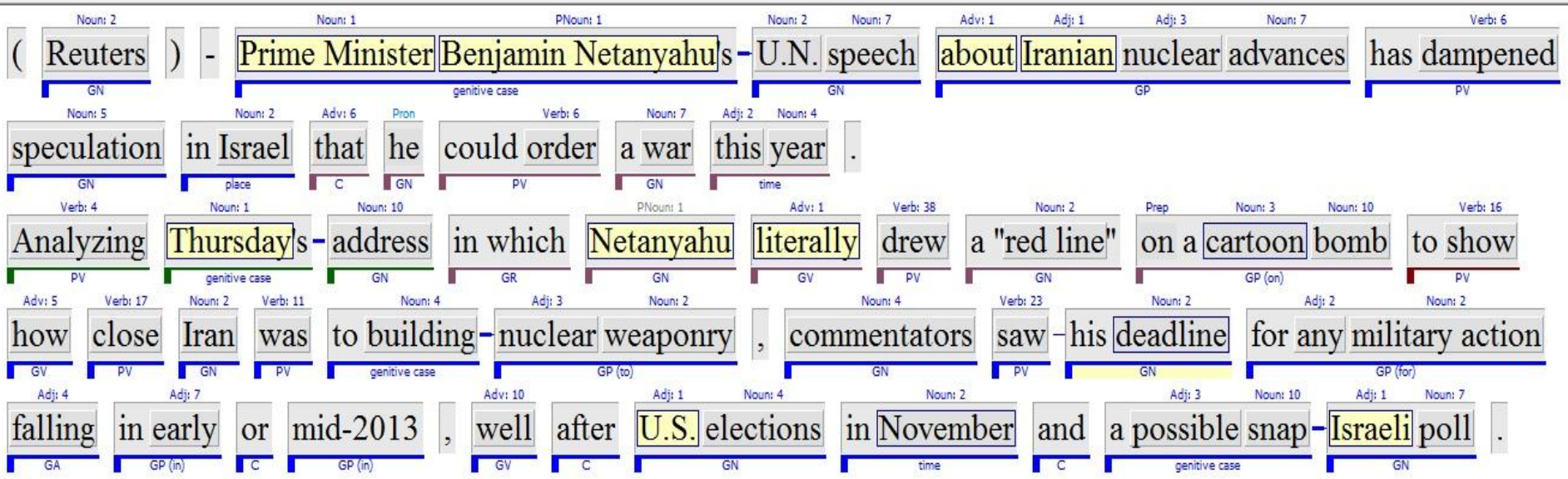
September 29-2012

DISAMBIGUATION:
GRAMMAR ANALYSIS

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

(Reuters) - Prime Minister Benjamin Netanyahu's U.N. speech about Iranian nuclear advances has dampened speculation in Israel that he could order a war this year.

Analyzing Thursday's address in which Netanyahu literally drew a "red line" on a cartoon bomb to show how close Iran was to building nuclear weaponry, commentators saw his deadline for any military action falling in early or mid-2013, well after U.S. elections in November and a possible snap Israeli poll.

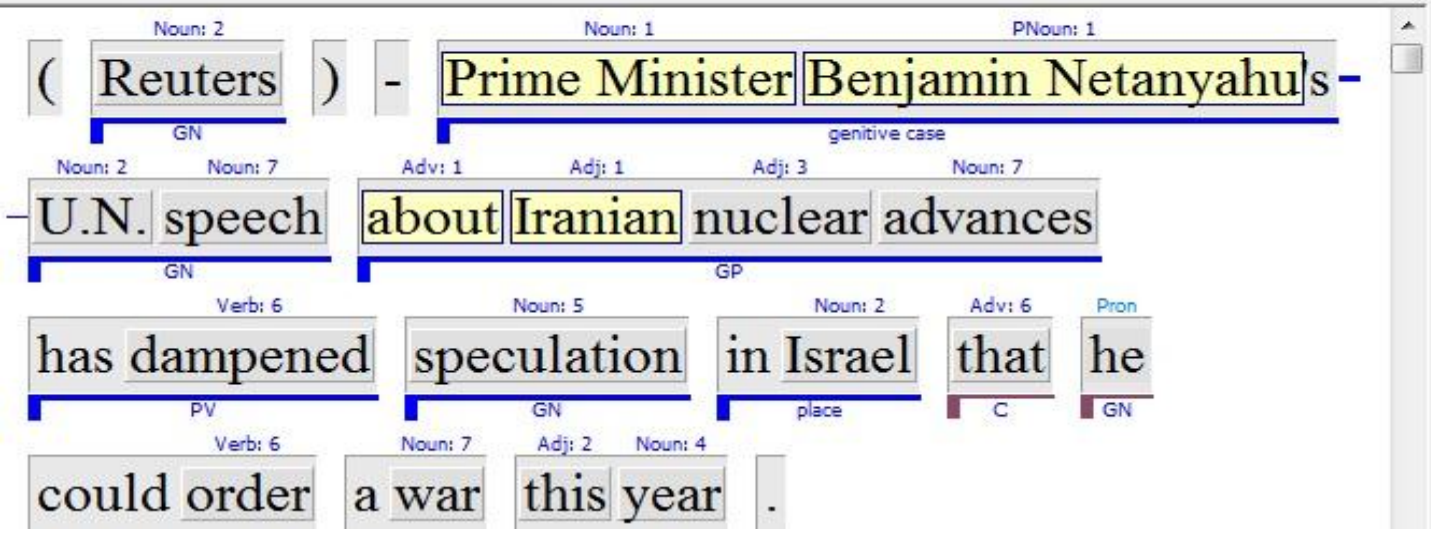


DISAMBIGUATION:
LOGICAL ANALYSIS

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

(Reuters) - Prime Minister Benjamin Netanyahu's U.N. speech about Iranian nuclear advances has dampened speculation in Israel that he could order a war this year.

Analyzing Thursday's address in which Netanyahu literally drew a "red line" on a cartoon bomb to show how close Iran was to building nuclear weaponry, commentators saw his deadline for any military action falling in early or mid-2013, well after U.S. elections in November and a possible snap Israeli poll.

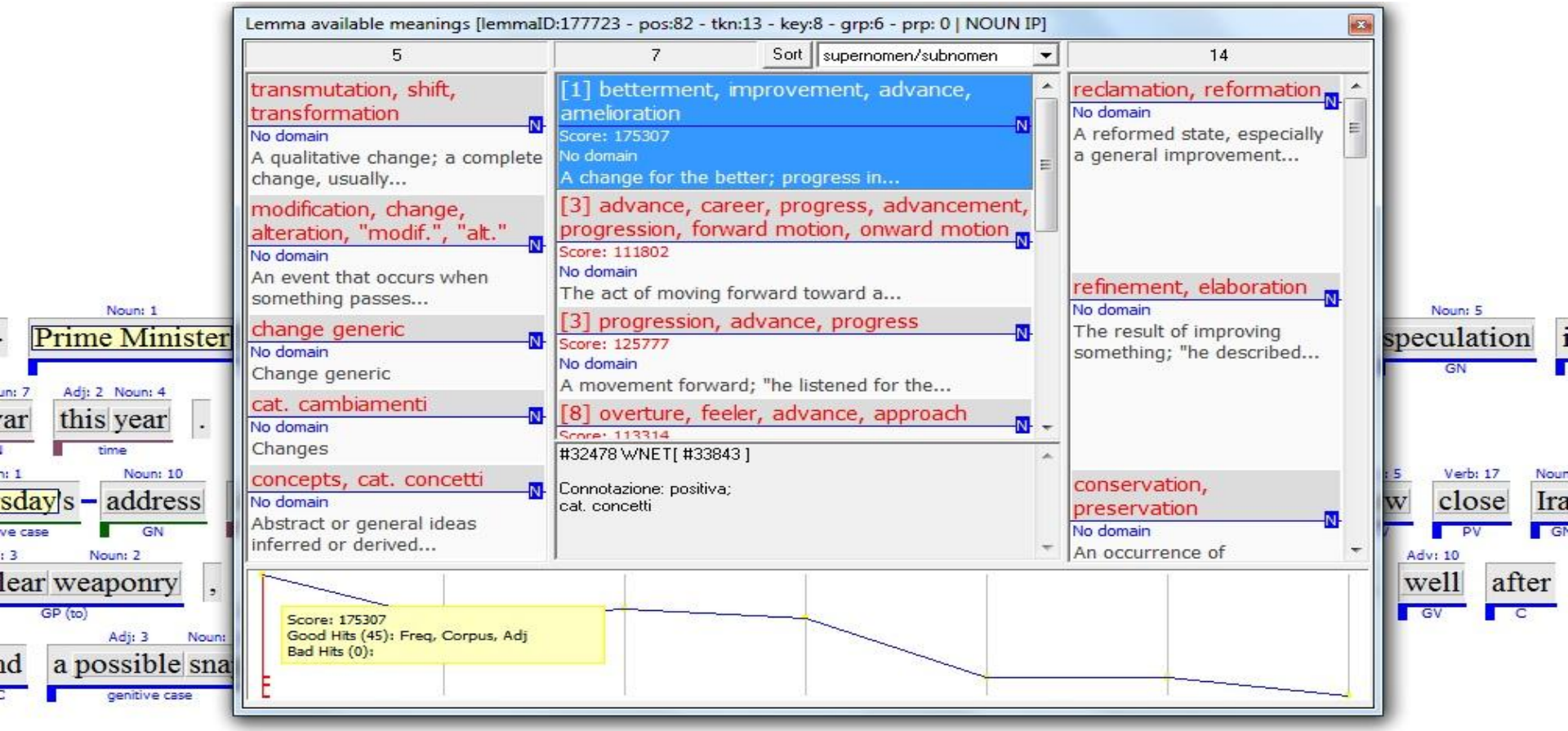


- military 2.0% (2)
  - nuclear #84660#
  - war #3951#
  - nuclear #84660#
  - military action #3882#
- politics 1.1% (2)
- institutes 0.2% (1)
- diplomacy 0.2% (1)
- rhetoric 0.1% (1)

- dampen
  - SBJ:"U.N. speech"
  - OBJ:"speculation"
  - C.PLACE:"in Israel"
- order
  - SBJ:"he"
  - OBJ:"a war"
  - C.TIME:"this year"
- analyze
  - OBJ:"address"
- draw

DISAMBIGUATION:
SEMANTIC ANALYSIS

THE COGITO
SEMANTIC NETWORK



WHAT IS A SEMANTIC NETWORK?

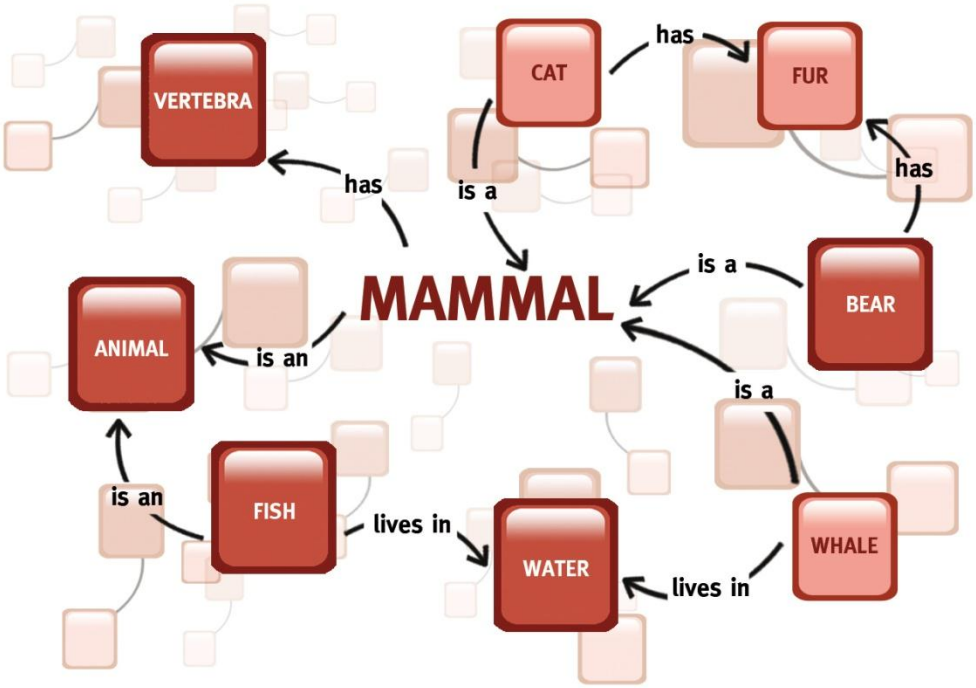
A rich map of associations and meanings of words.

- Includes all definitions of all words.
- Includes relationships between all words.

The quality of results is derived from the richness and complexity of the semantic network.

COGITO® English Semantic Network:

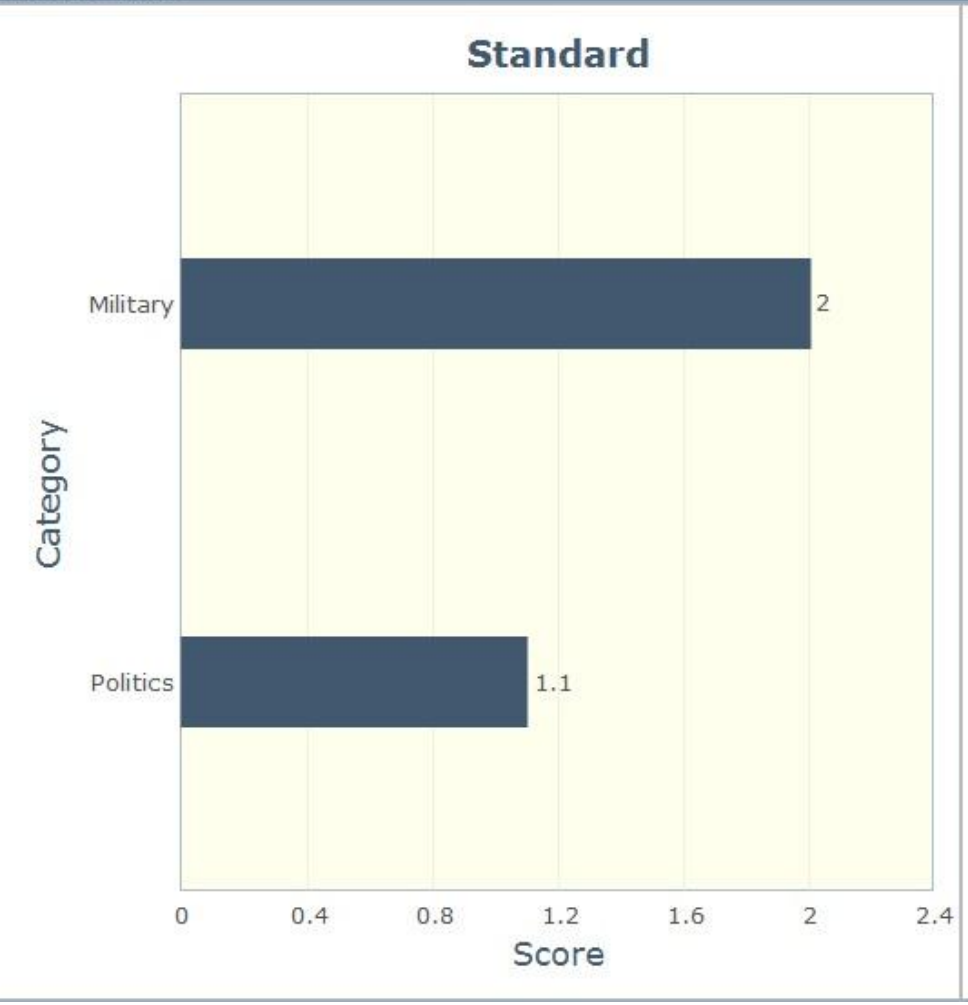
- 350,000 words
- 2.8M relationships

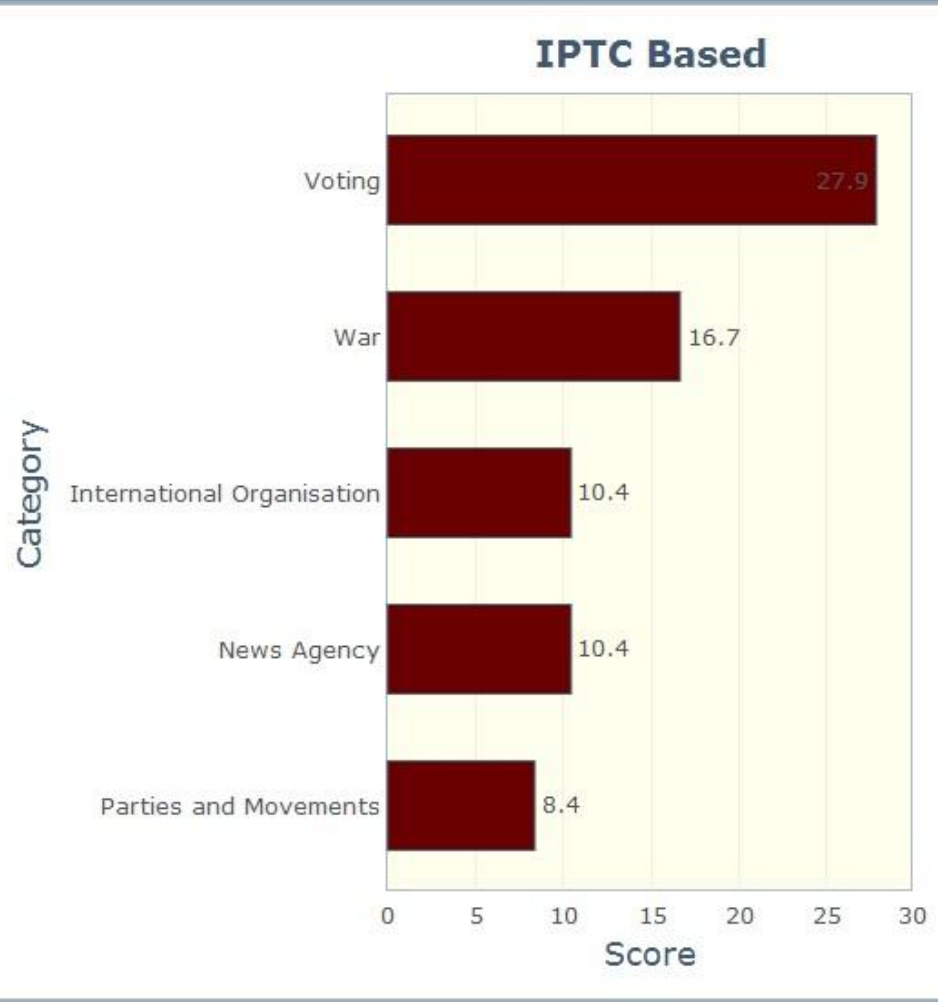


CUSTOM TAXONOMIES

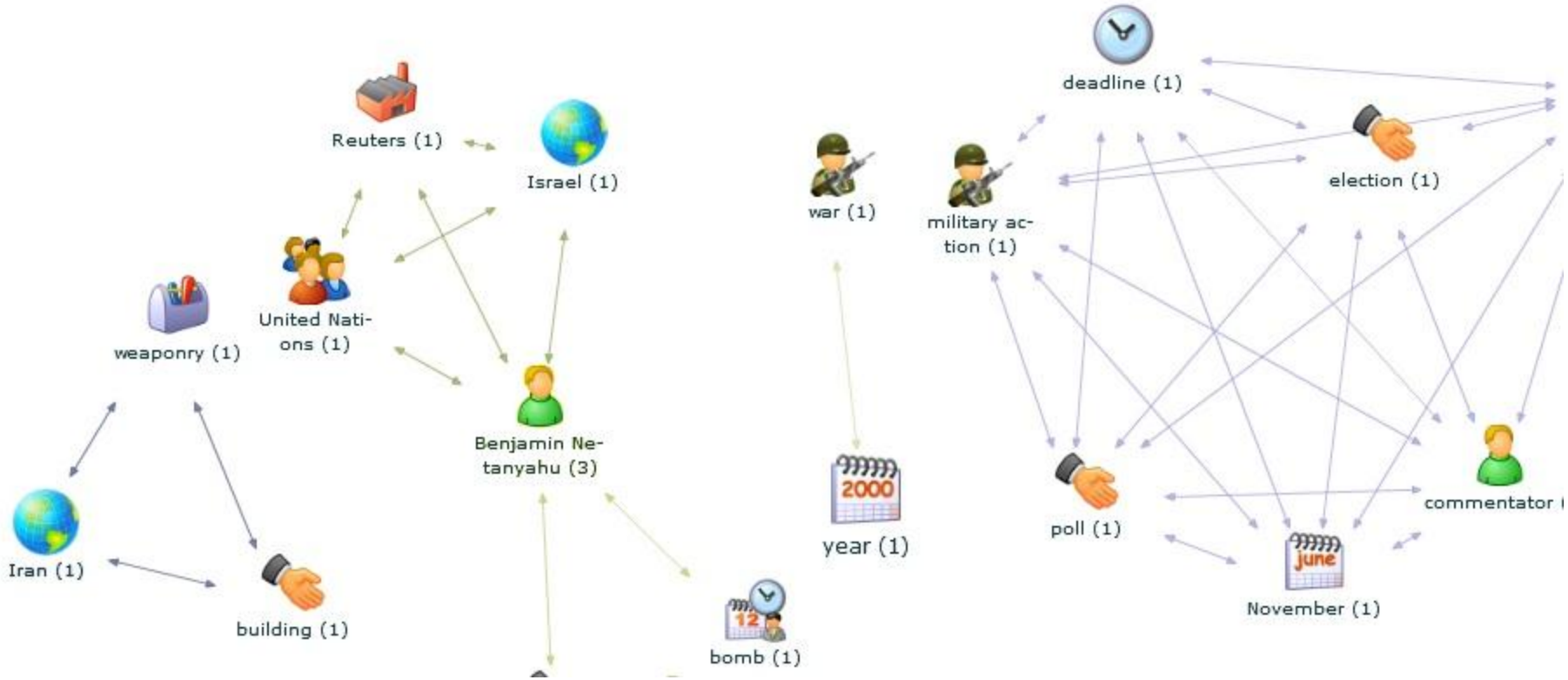
EXPERT SYSTEM
SEMANTIC INTELLIGENCE

Categorization





WHO: RELATIONSHIPS BETWEEN ENTITIES



WHAT: CONTEXT AND CONCEPTS OF INTEREST

What's happening?

People, Organizations and Places

- **Places (2 items)**
  - Iran
  - Israel
- **People (1 item)**
  - Benjamin Netanyahu (Prime Minister)
- **Organizations (2 items)**
  - Reuters
  - United Nations

Main concepts

Semantic Tagging

Main elements

Benjamin Netanyahu	(16.5%)
red line	(6.2%)
United Nations	(6.1%)
deadline	(5.8%)
Iran	(5.4%)
speculation	(5.4%)
speech	(5.4%)

Concepts of interest

Elements	
Company	
Reuters (#225910)	News agency: a London news agency providing...
Country	
Iran (#10128582)	unitary state (Asia)
Israel (#10291909)	unitary state (Asia)
Event	
bomb (#138261)	A performance that is a commercial or...
Month	
November (#65628)	The month following October and preceding December...
Tool	
weaponry (#22244)	Weapons considered collectively
Year	
year (#65576)	A period of time containing 365 (or...
MilitaryAction	
military action (#3882)	A military or naval action (as a...
war (#3951)	The waging of armed conflict against an...

Note that the original text was "a cartoon bomb" therefore, the word bomb has been disambiguated in this way

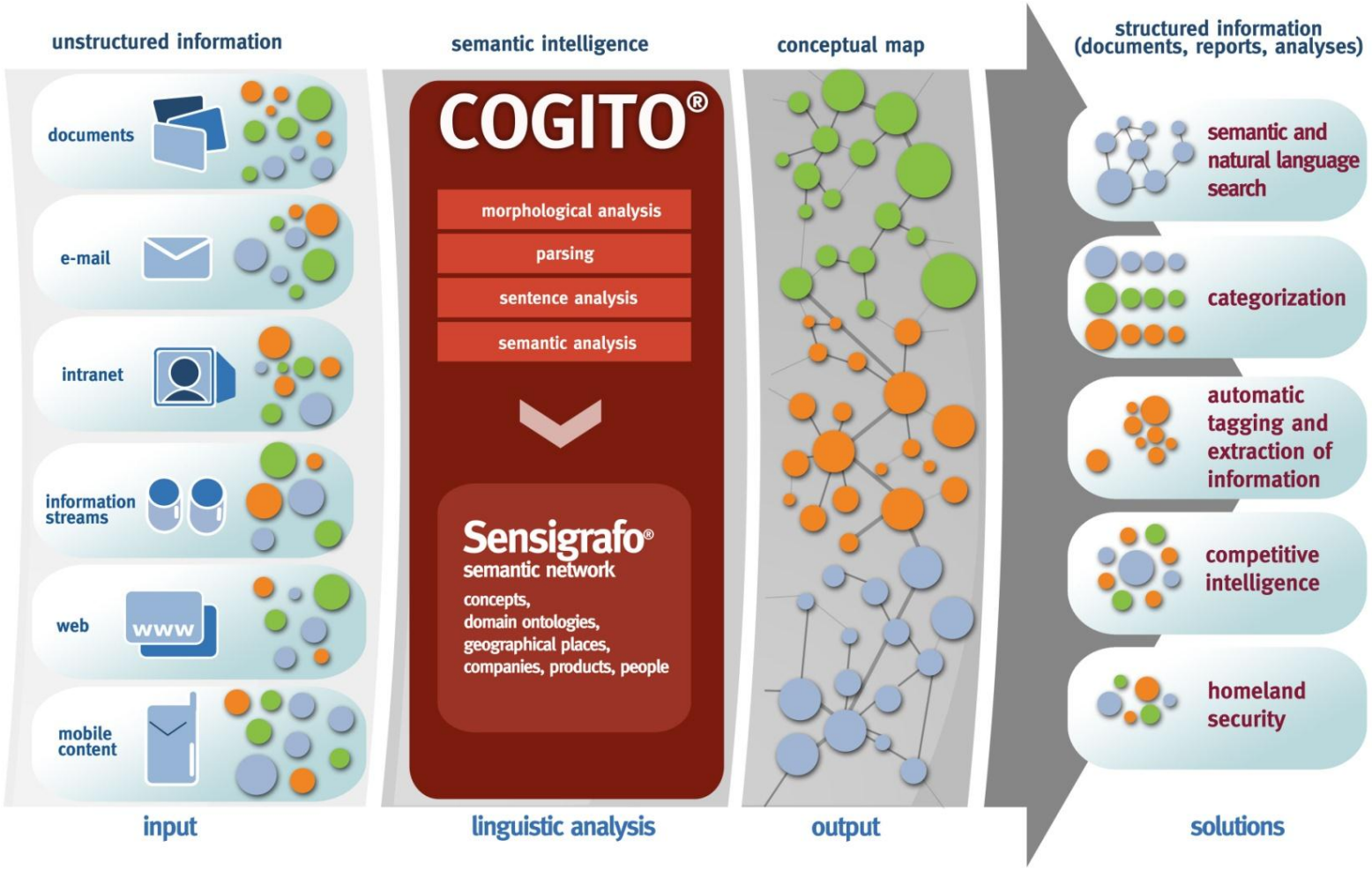
WHERE: INTEGRATION WITH MAPS

WHEN and HOW

Automatic identification of the ***timeline*** of events and ***how*** the different entities were involved (type of entities, role in the event, etc.).

see (#68528)		
+ sbj	People	commentator (#46365)
+ obj	Time	deadline (#65455)
+ other	MilitaryAction	military action (#3882)
+ other	Date	2013 (#v299992)
+ other	Action	election (#672)
+ time	Month	November (#65628)

HOW COGITO WORKS



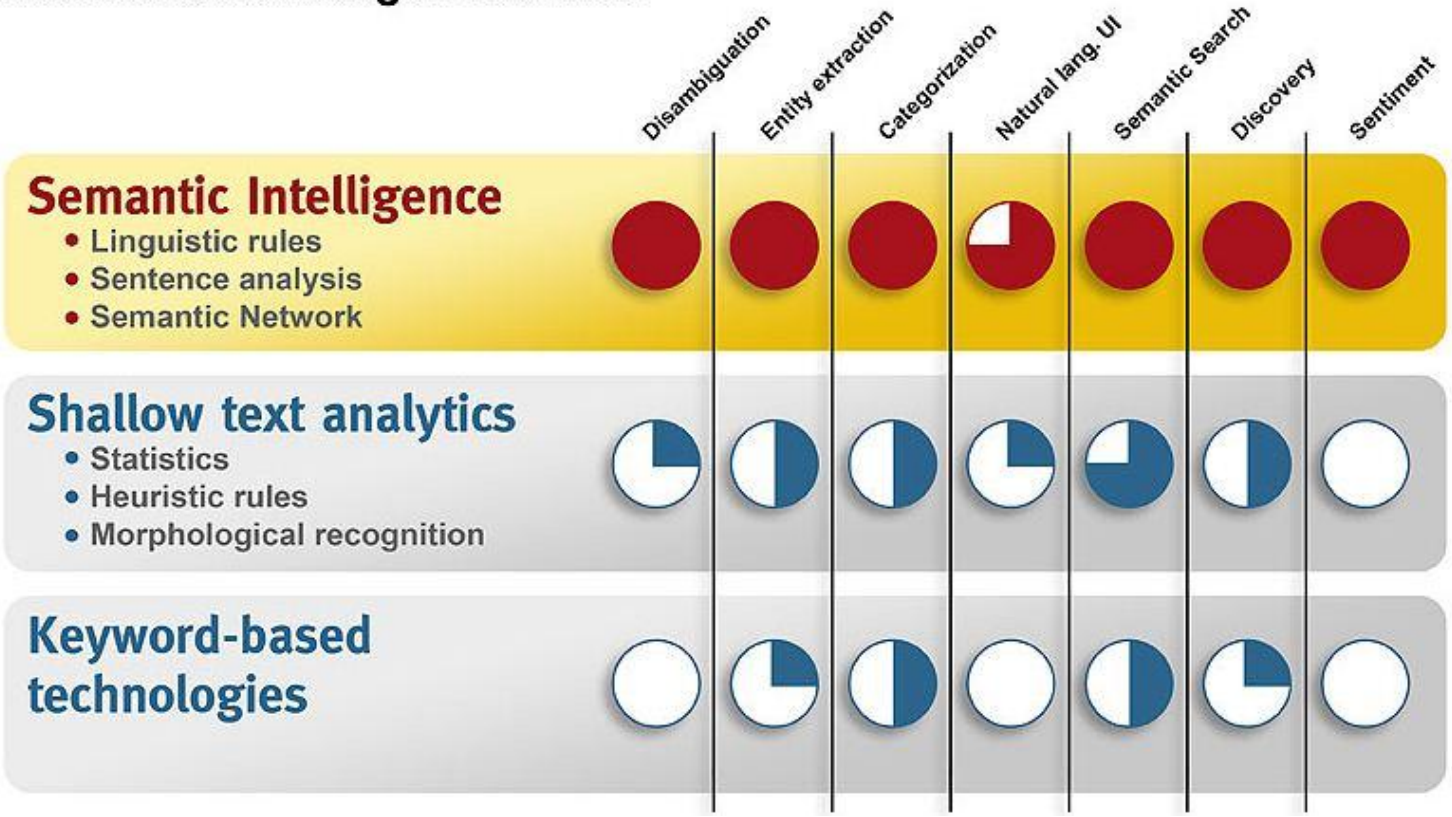
NEXT GENERATION TECHNOLOGY

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

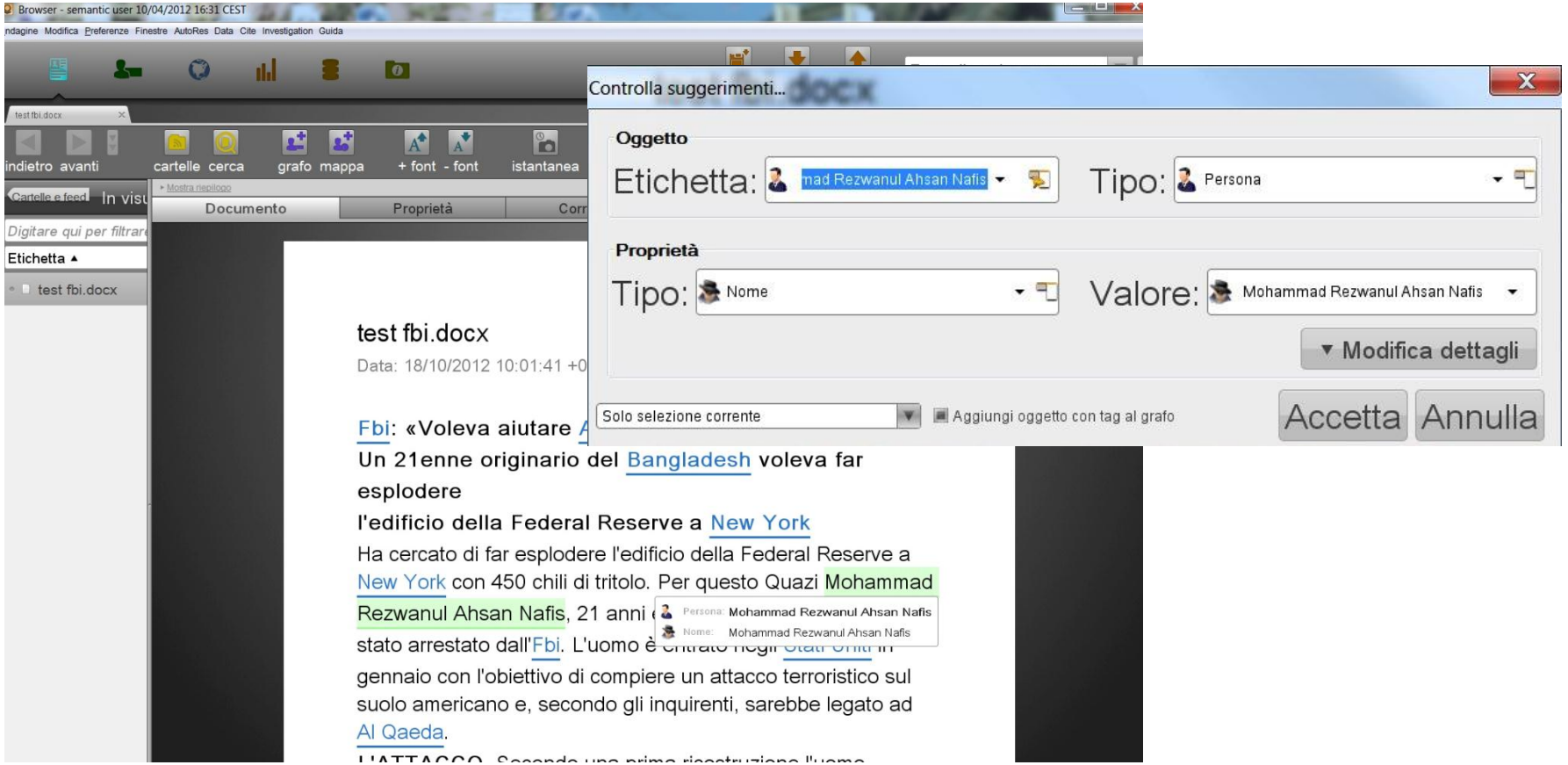
100% semantic technology for next-gen text analysis

Excels in the areas where traditional technologies fall short

COGITO®



PALANTIR © INTEGRATION EXAMPLE



COGITO's real time entities extration is made available to the user who can immediately validate it

WHO WE ARE

Expert System is the largest, fastest growing semantic software company in the world.

We develop technology, applications and solutions to extract, understand and share information more effectively.

ESTABLISHED MARKET PRESENCE

- Expert System was established in Modena, Italy by three young programmers with an idea. A few months later, Expert System's software was integrated into the Microsoft Office suite.
- **US subsidiary** since 2007.
- Private and Profitable with Revenue doubled in the last three years to **over $15 million in 2011** and EBITDA above 20%.
- **30% of resources devoted to R&D** and over $14 million invested in the last 3 years, with $5M more planned for the next 2 years.
- **More than 100 employees** and offices in Italy, London, Washington, D.C. and Chicago.

RECOGNIZED FOR MATURE AND PROVEN TECHNOLOGY

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

Gartner

Identified among the world's leading information access technology developers.

IDC
Analyze the Future

Selected one of the "Innovative Information Access Companies Under $100M to Watch."

FORRESTER®

Recognized for text analytics and superior SharePoint integration capabilities.

Microsoft®

One of the few non-Microsoft technologies in the MS Office suite.

CONTACT

Thank You!

Gian Piero Oggero
goggero@expertsystem.net

www.expertsystem.net

COGITO®
INTELLIGENCE PLATFORM

Demonstration

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

WHAT WE DO

The Cogito Intelligence Platform supports analysts in all phases of the intelligence cycle:

- ***Acquisition***: Crawler to acquire data from different sources
- ***Exploitation***: Interact with data using semantic analysis
- ***Evaluation***: Easily produce reports using the exploited data
- ***Distribution***: Sharing data with other authorized users

COGITO INTELLIGENCE PLATFORM

Combines superior text analytics and domain ontology capabilities with the ability to search and manage large quantities of data from:

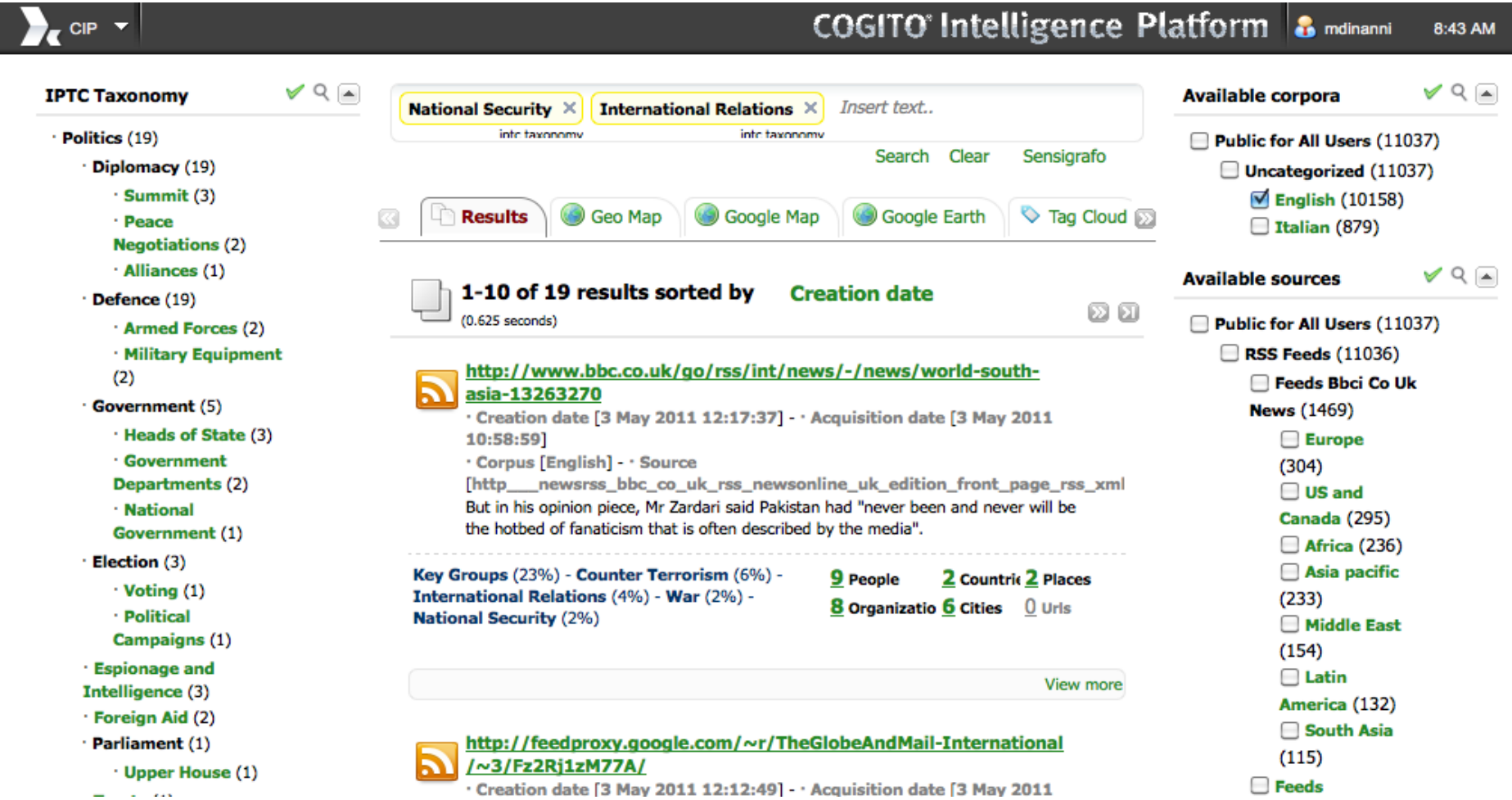
- Documents
- Multimedia
- Audio streams
- Web pages
- Social networks

Incorporates a proven approach for intelligence data management, and features the best software components for:

- Speech analysis
- GEO mapping
- Deductive algorithms
- Advanced visualization systems

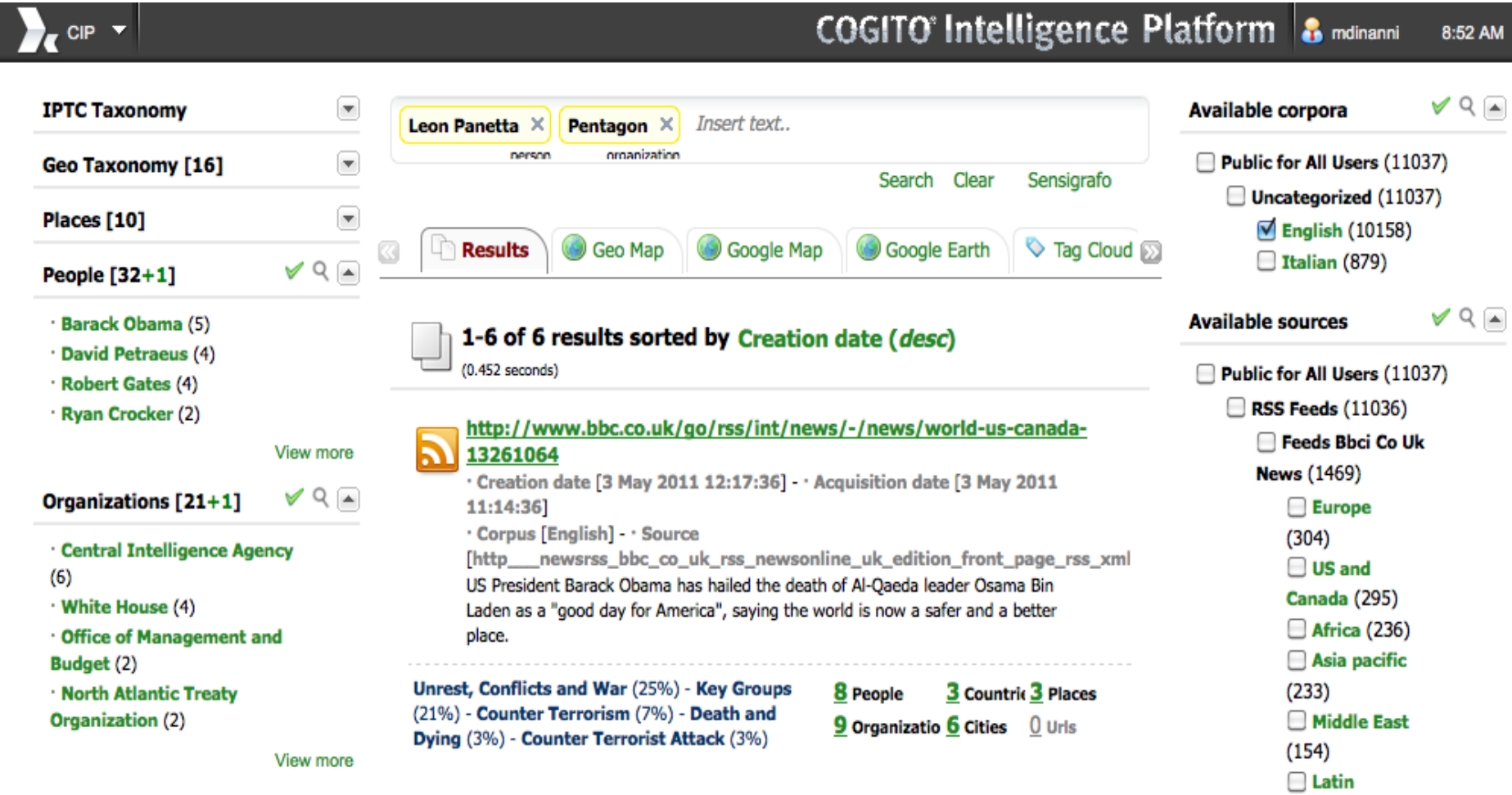
WORKING WITH CONCEPTS, KEYWORDS AND LEMMAS

WORKING WITH TAXONOMIES AND CATEGORIES



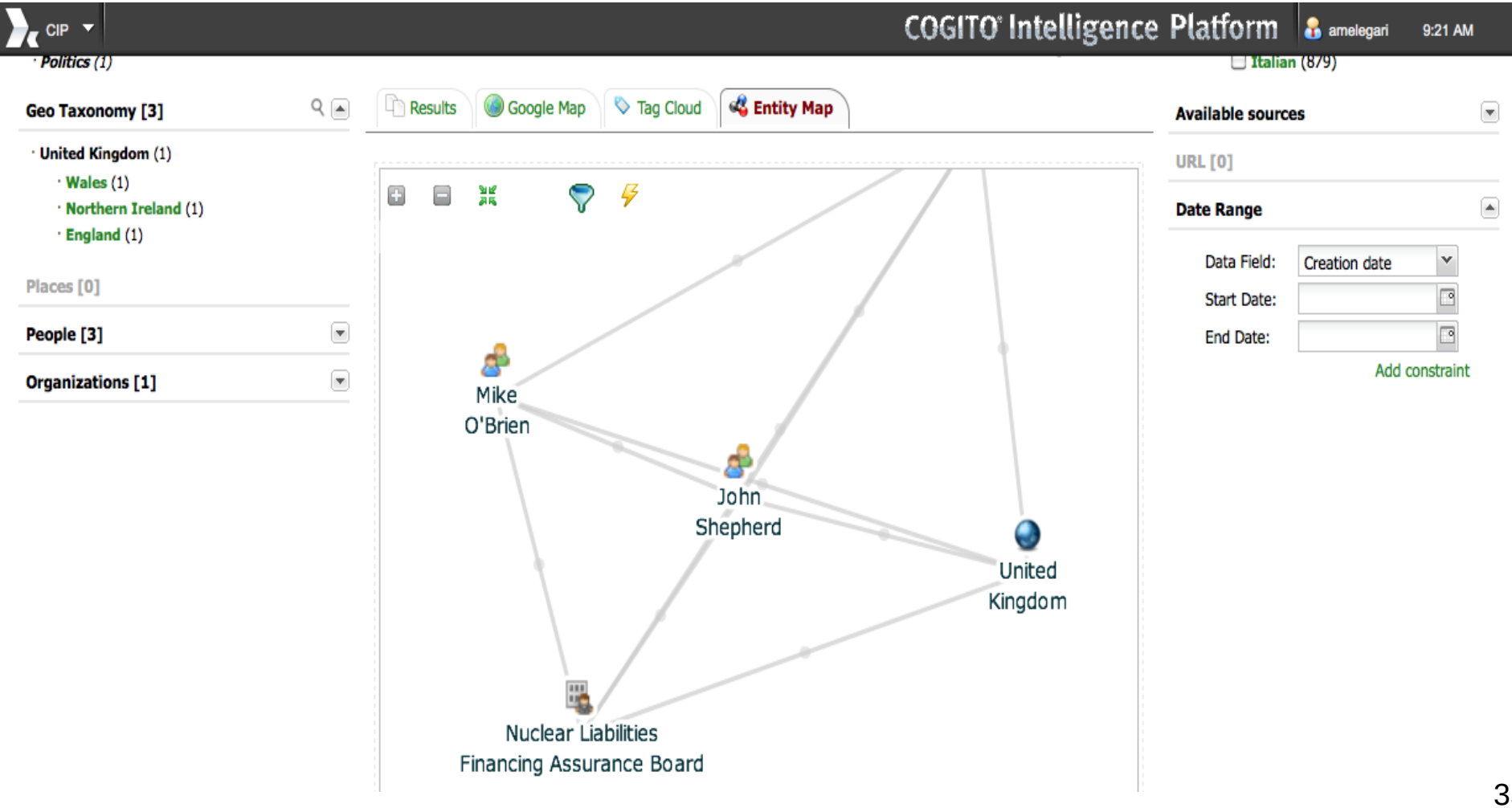
USING ENTITIES FOR INTERACTING WITH DATA

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

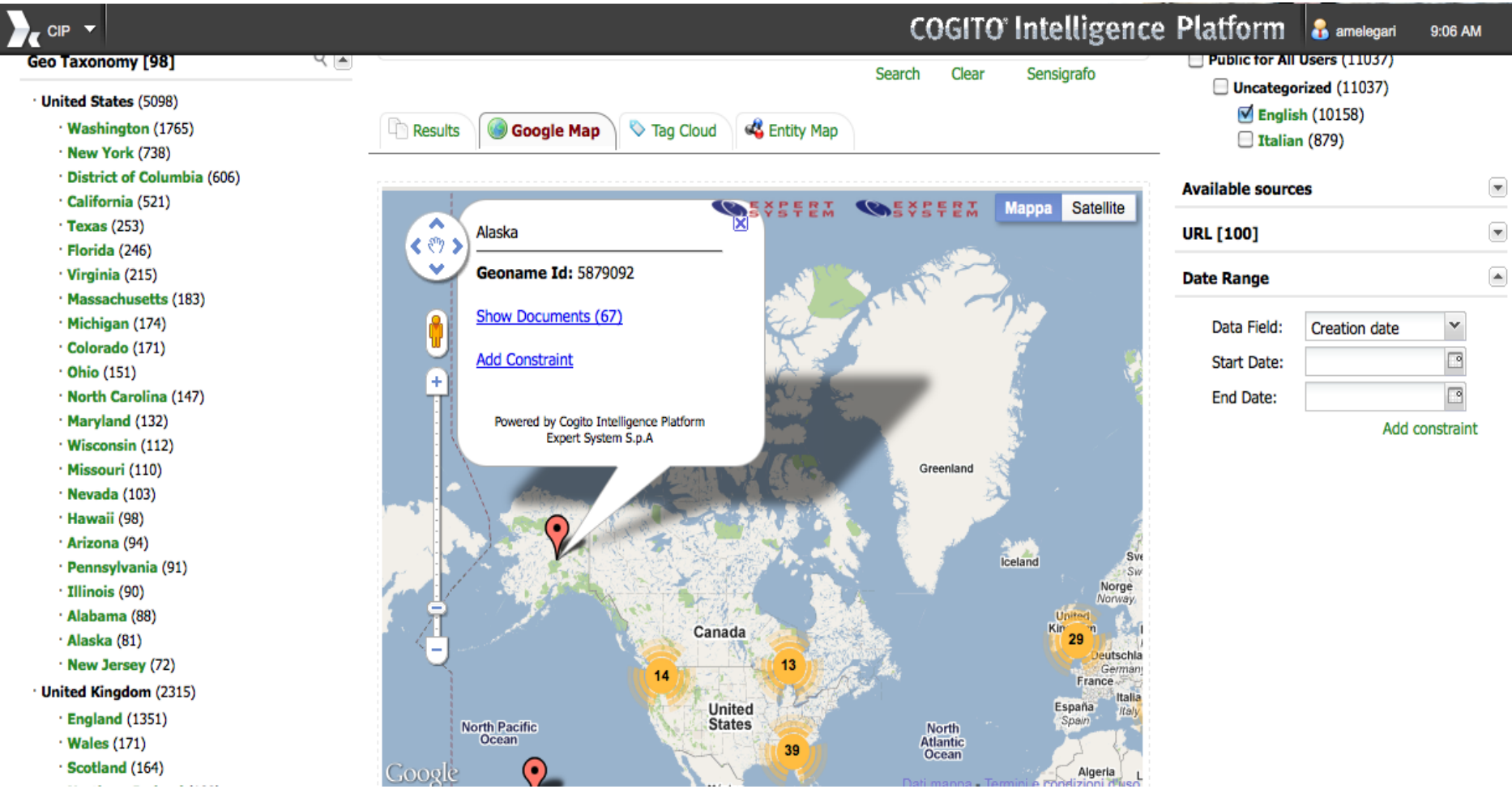


CREATING QUERIES USING THE SEMANTIC TAG CLOUD

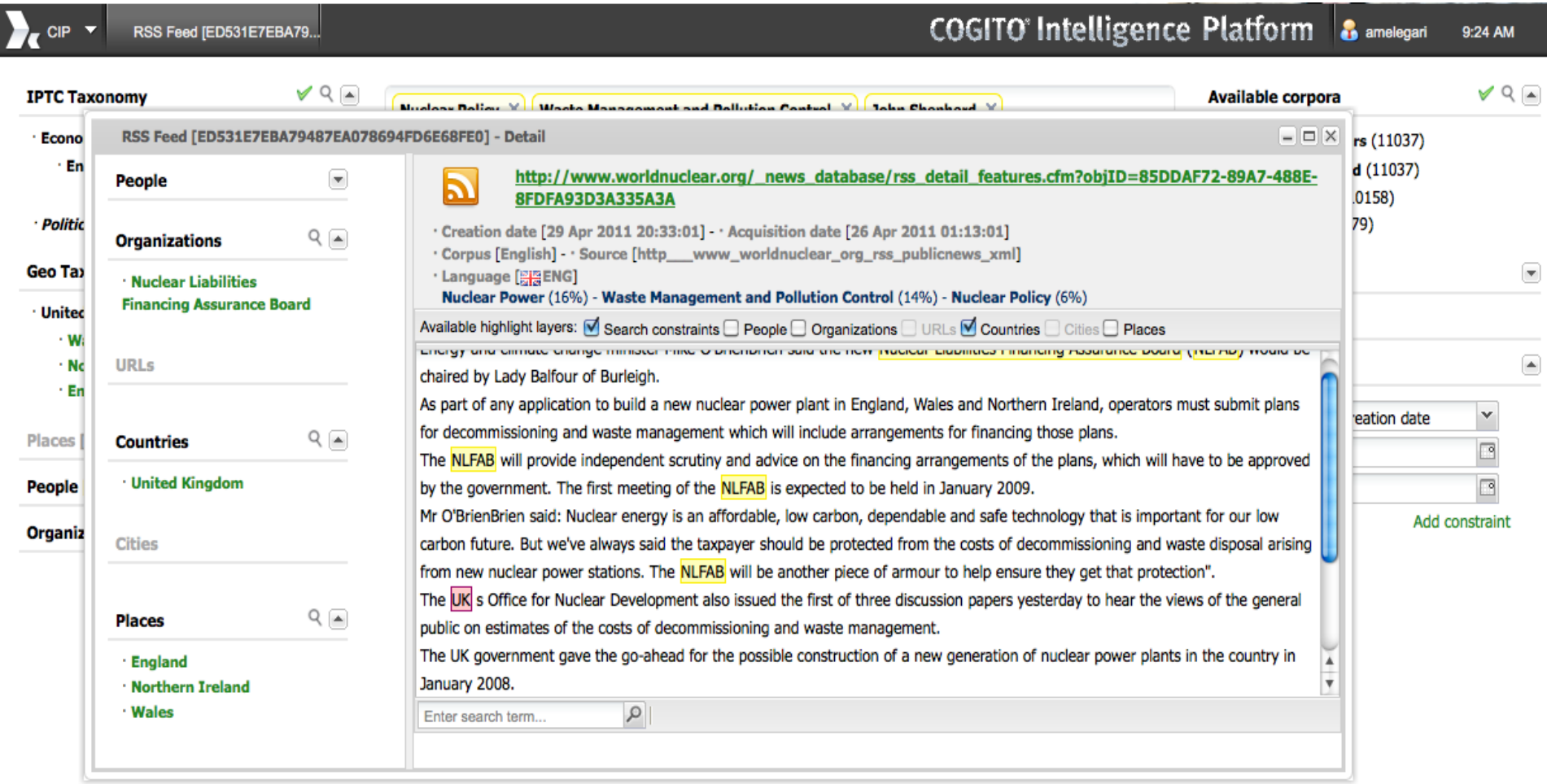
VIEW ENTITIES USING E-R DIAGRAM



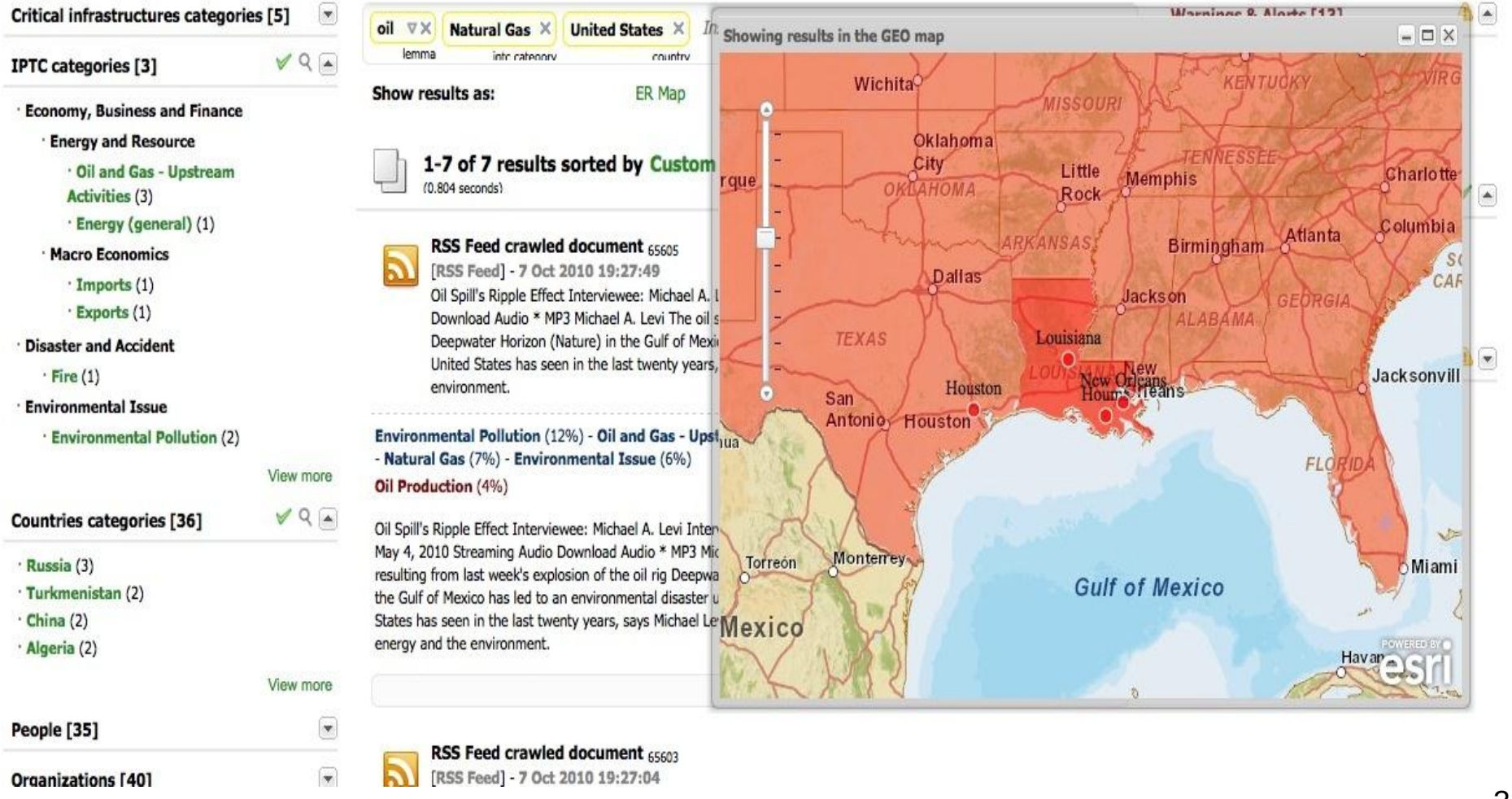
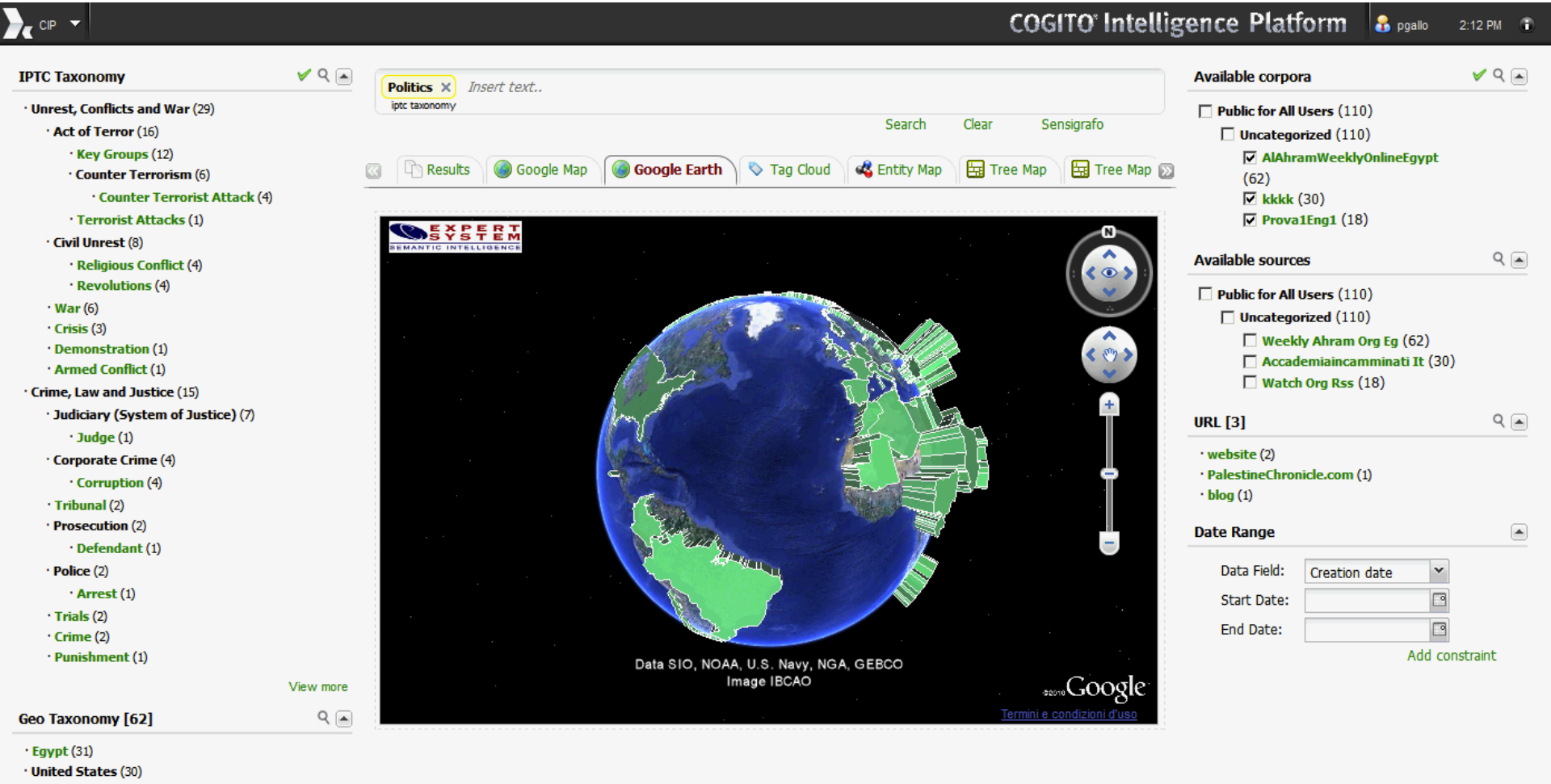
USING GOOGLE MAPS TO GEOREFERENCE AND INTERACT



KEY TARGETS ARE HIGHLIGHTED IN EACH DOCUMENT



INTEGRATION WITH GIS SYSTEMS



USING GOOGLE EARTH TO VISUALIZE GLOBAL PHENOMENA