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

Oil
Natural Gas
Downstream Products
Power Generation
Engineering & Construction
Petrochemicals

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

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

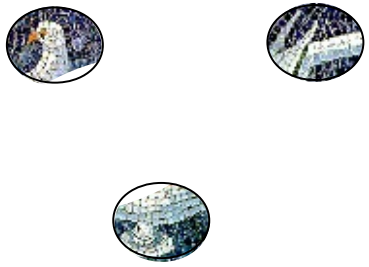
TO SEARCH FOR INFORMATION — THE FIRST TASKS

TO PROVIDE KNOWLEDGE — THE ULTIMATE IMPERATIVE

Semantic Search: the “mosaic” metaphore

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

Keyword search

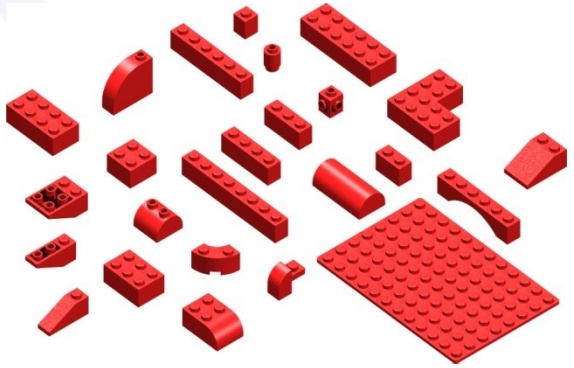


Semantic search

Semantic Search: the “mosaic” metaphore

Semantic processing

Syntactic and Semantic tiles



- 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 representation

*Knowledge is **connection**, not collection*

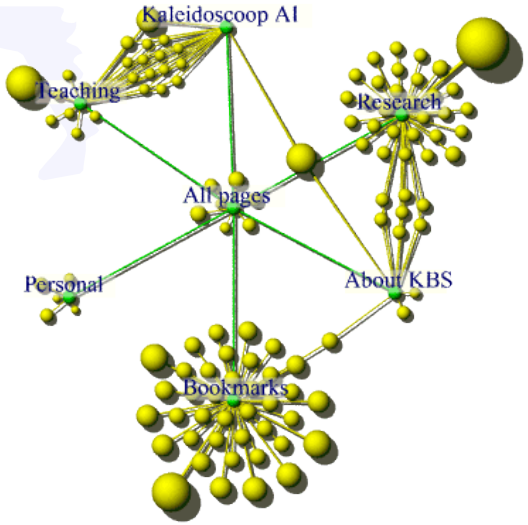
Standard Information retrieval

collection

Semantic engine

Semantic search

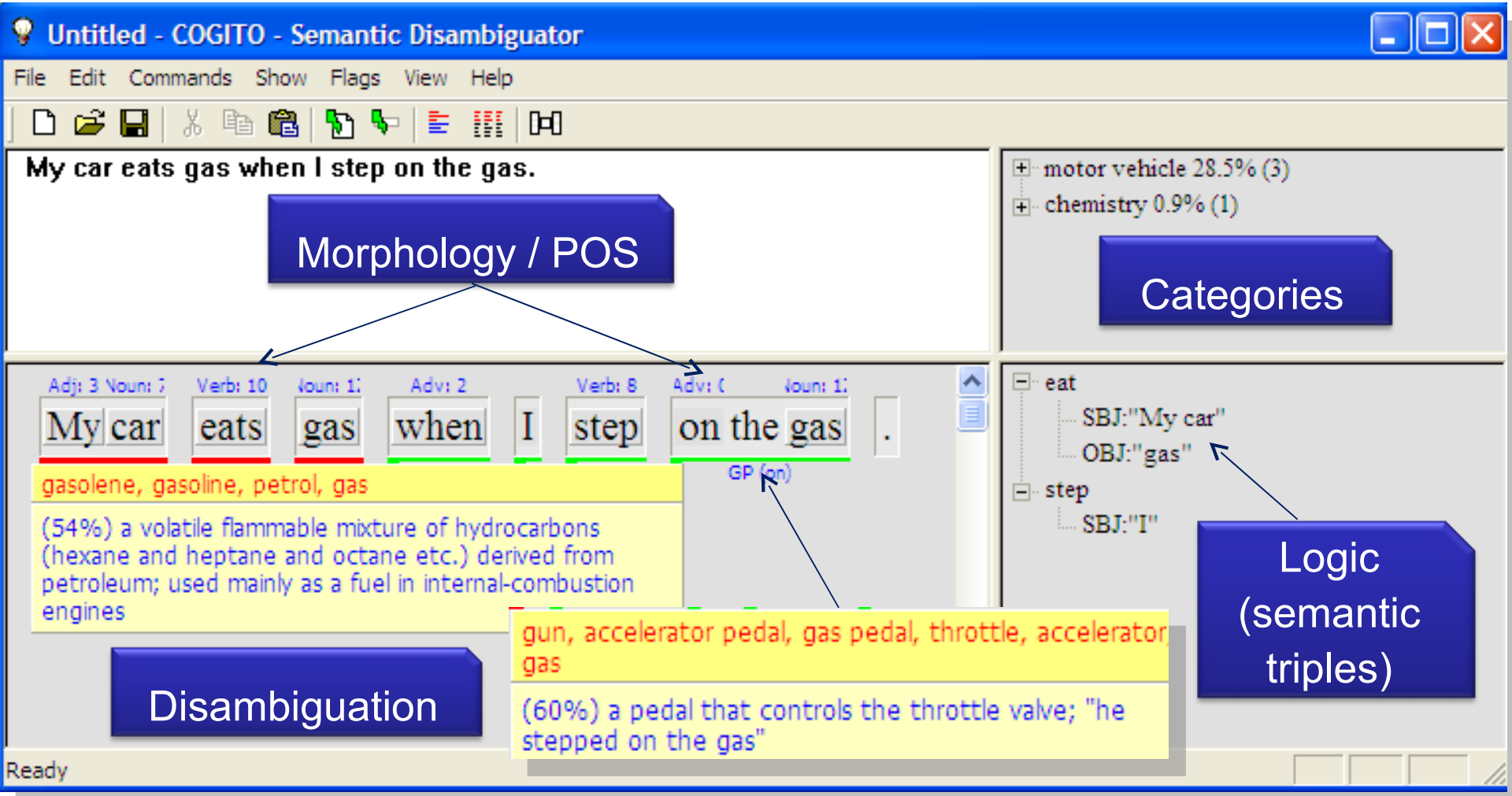
connection



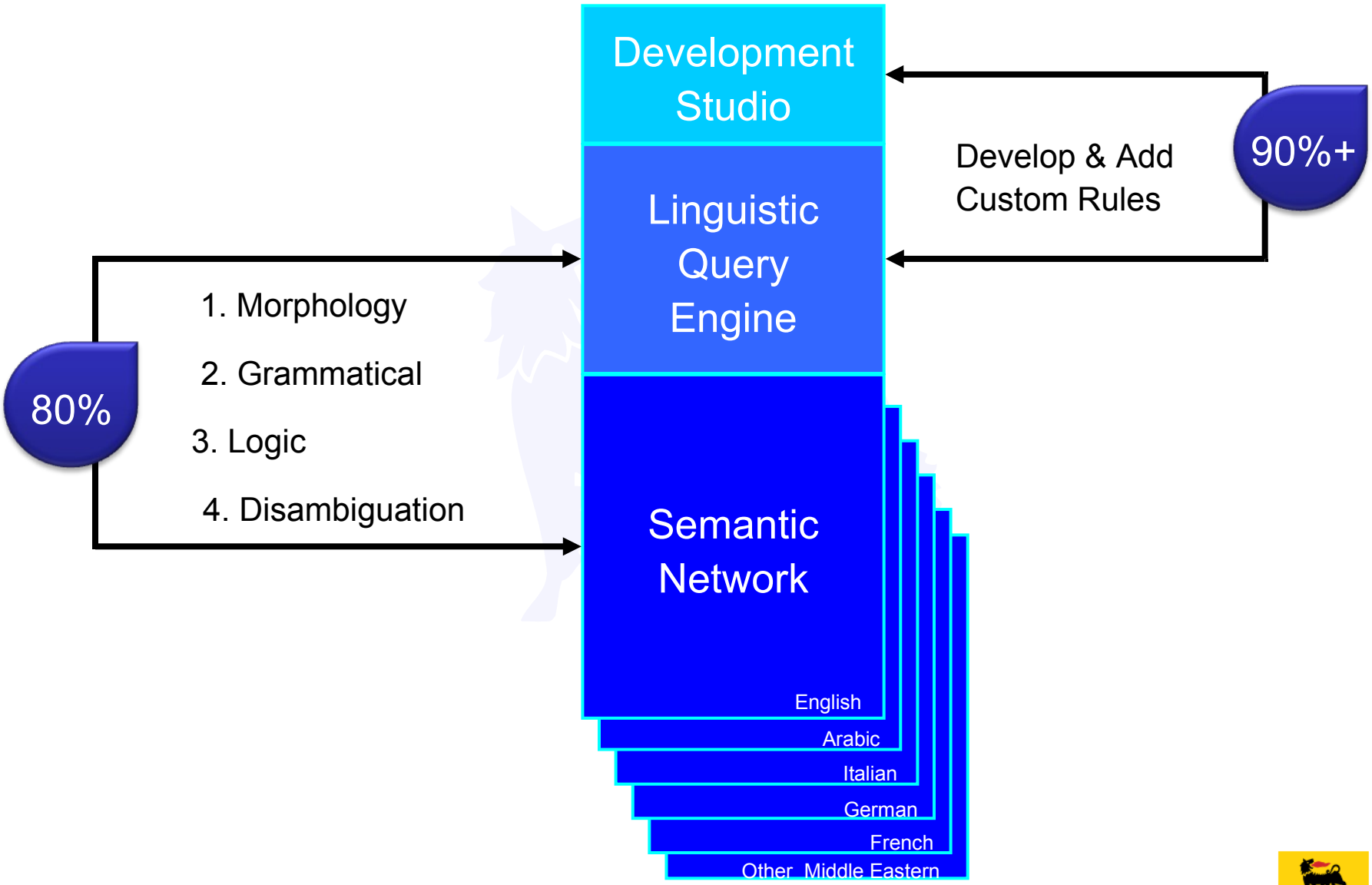
Technology - Semantic Methods

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

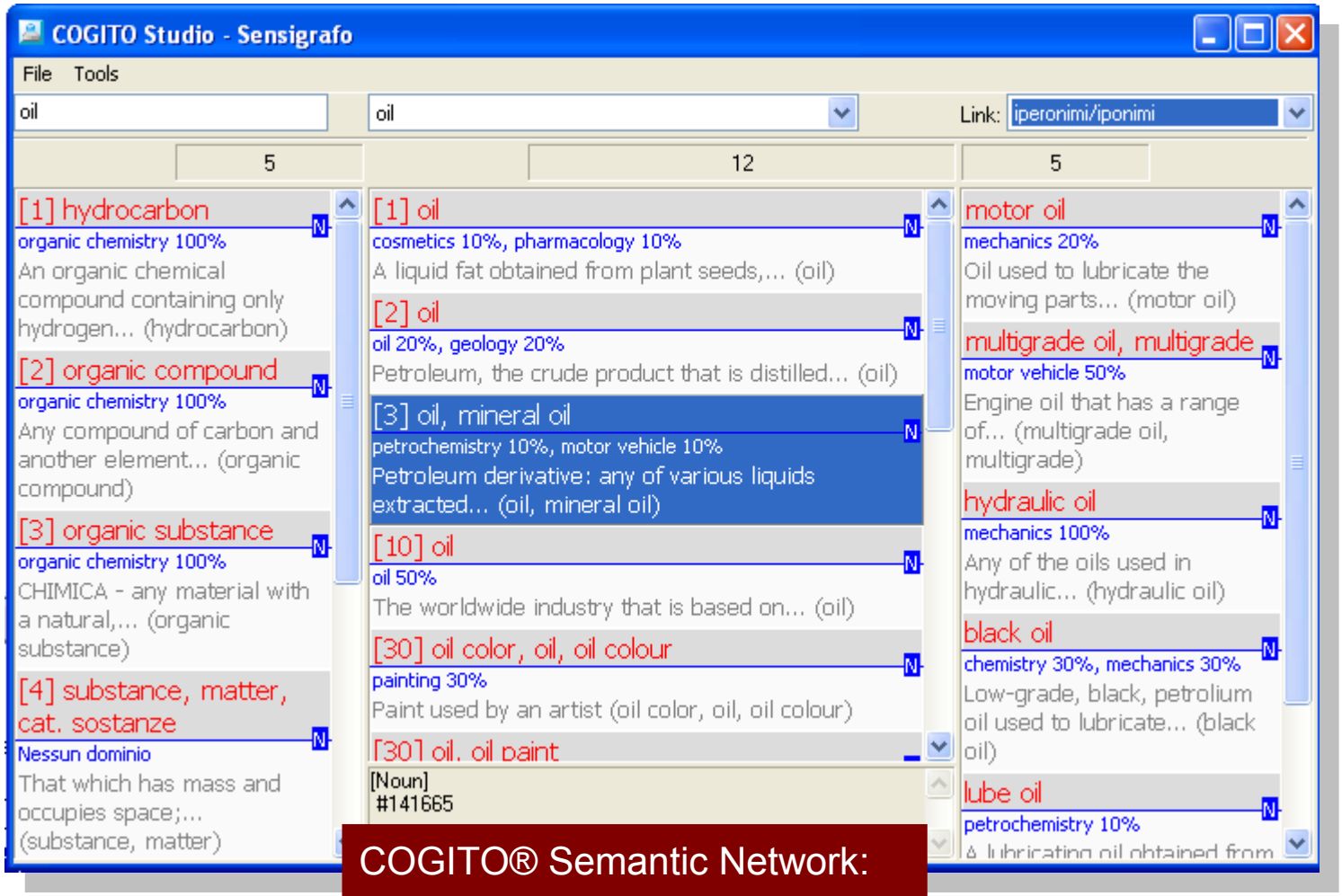
Technology - Semantic Methods



Technology Stack & Precision



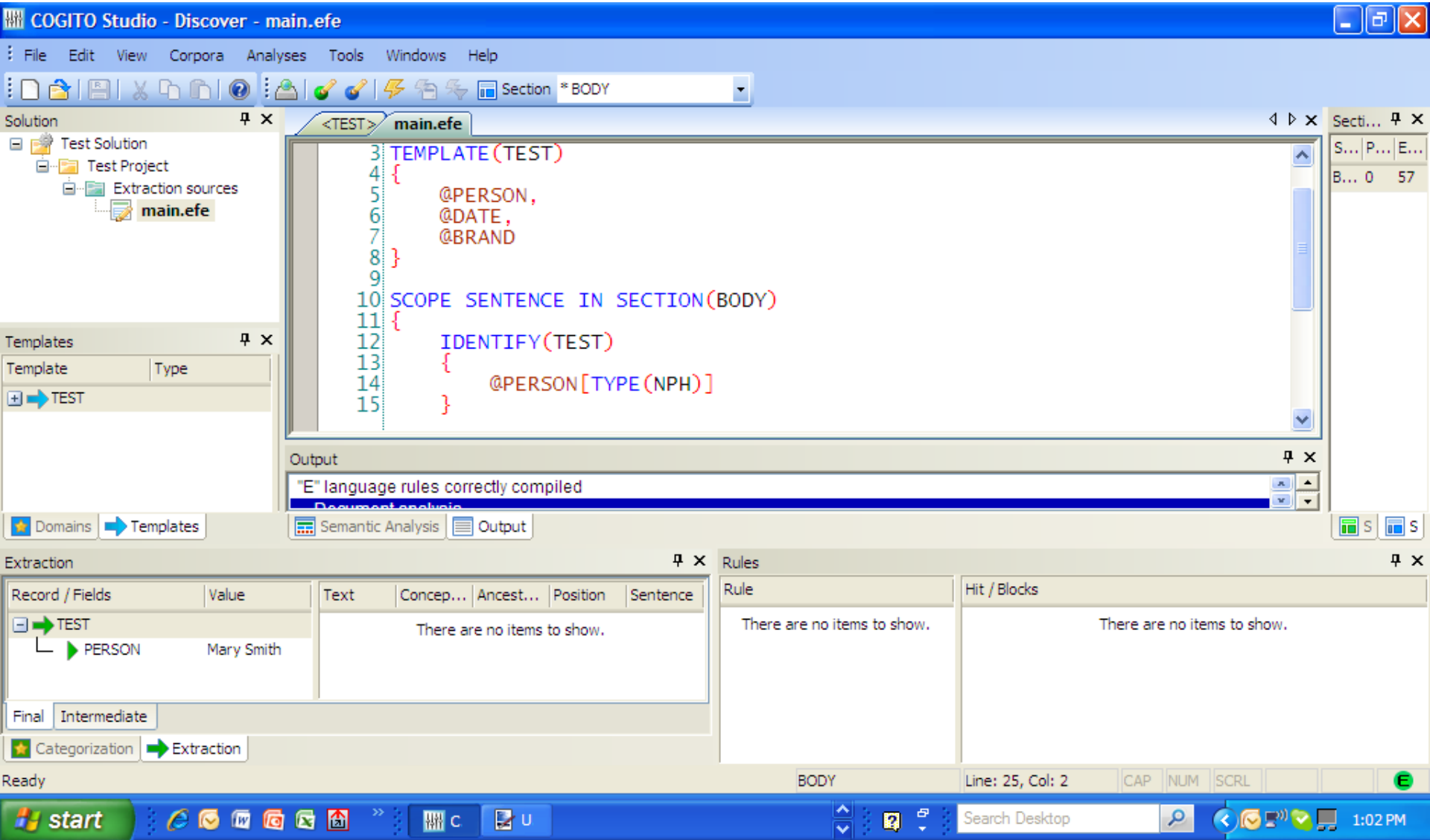
Technology - Semantic Network



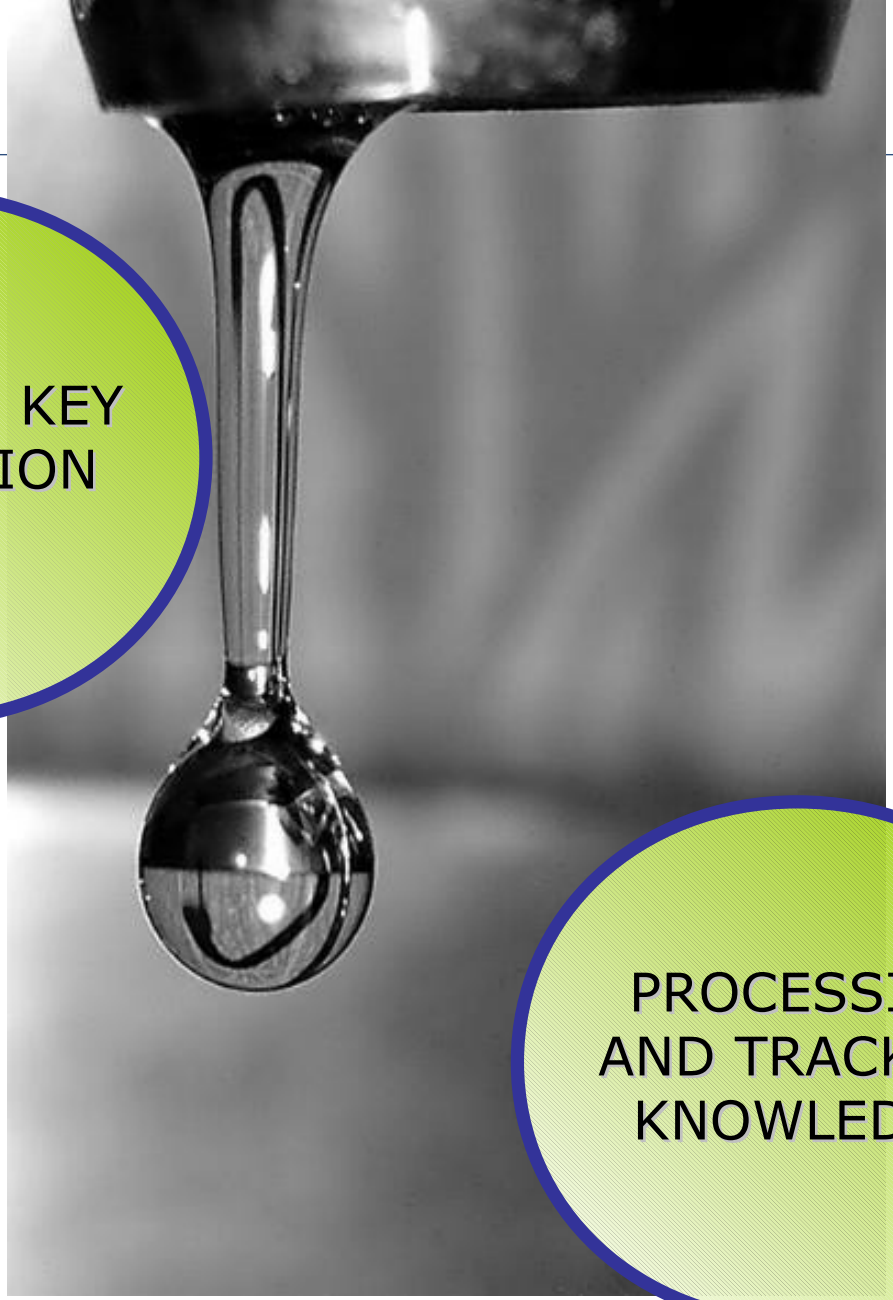
COGITO® Semantic Network:

- 350,000 words
- 2.8m relationships

Semantic Development Environment



Use Case One



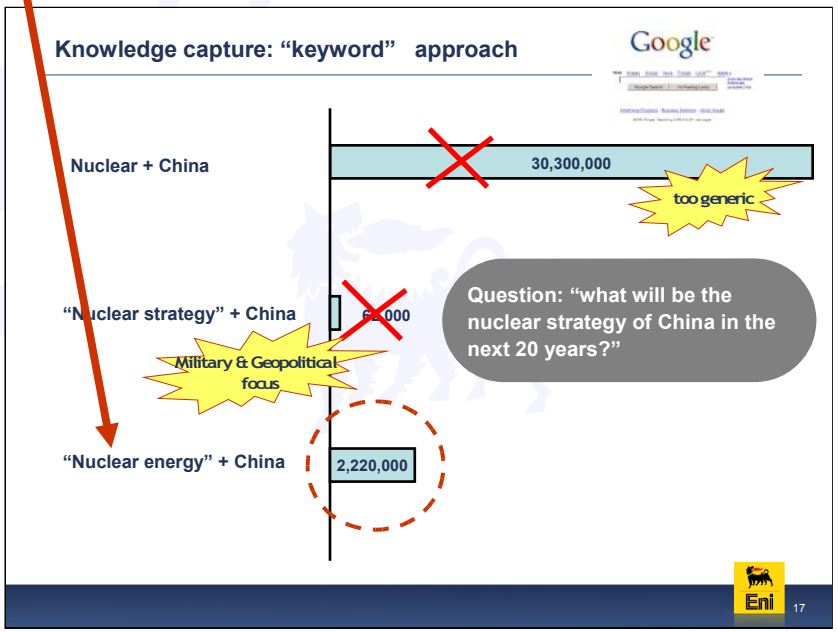
DISTILLING KEY INFORMATION

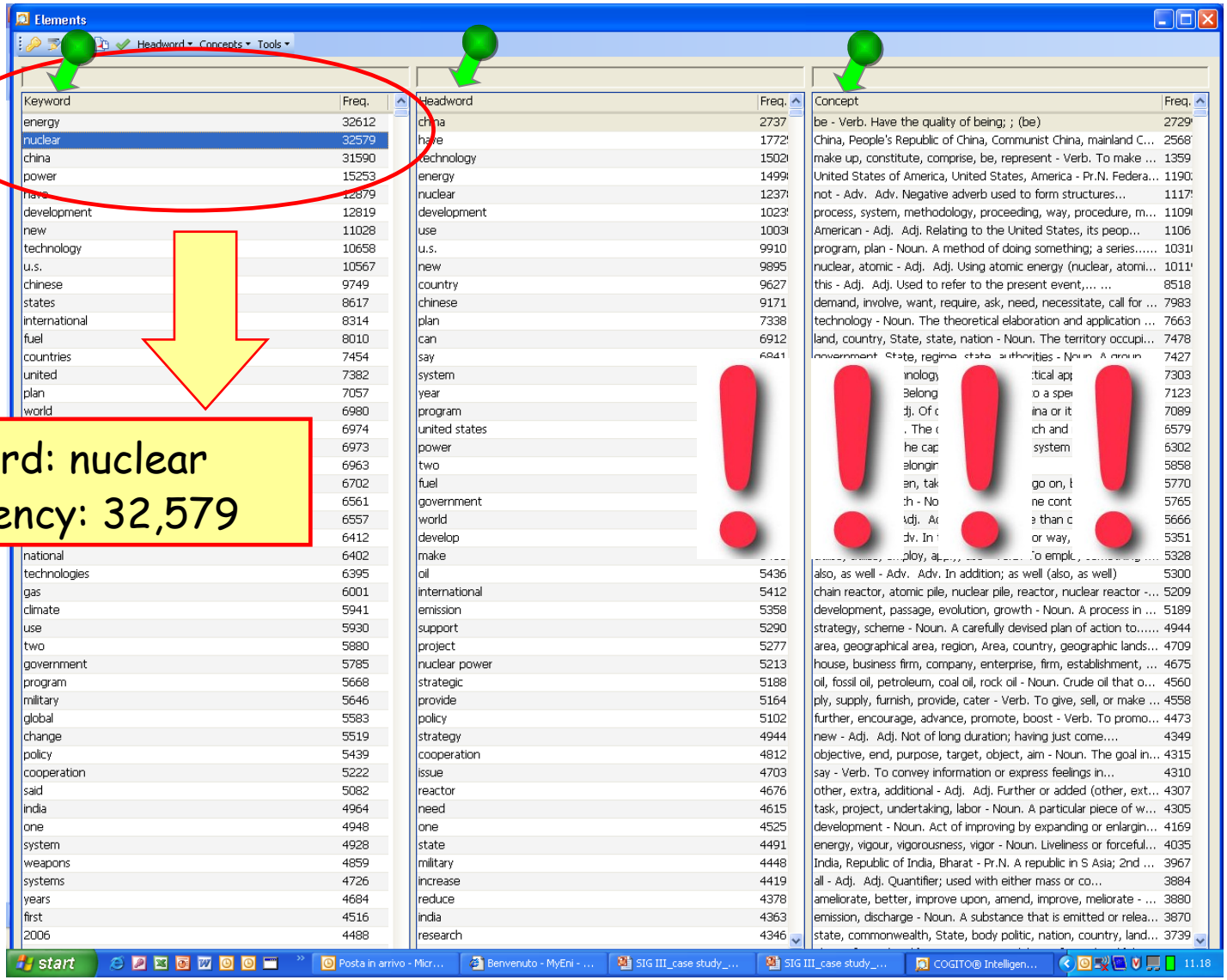
PROCESSING AND TRACKING KNOWLEDGE

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

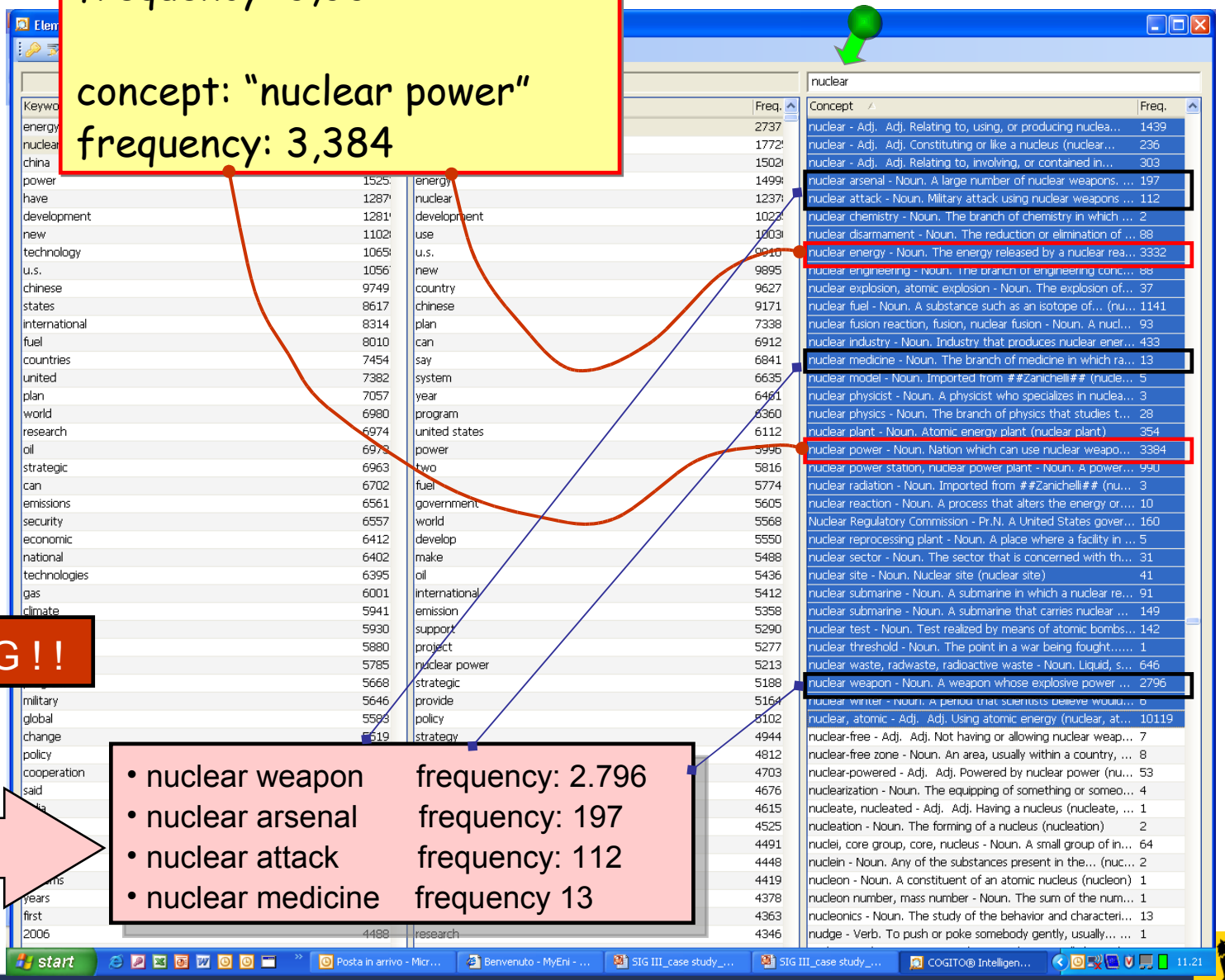



Elements
Headword
Concepts
Tools
Keyword
Freq.
energy 32612
nuclear 32579
china 31590
power 15253
development 12819
new 11028
technology 10658
u.s. 10567
chinese 9749
states 8617
international 8314
fuel 8010
countries 7454
united 7382
plan 7057
world 6980
national 6402
technologies 6395
gas 6001
climate 5941
use 5930
two 5880
government 5785
program 5668
military 5646
global 5583
change 5519
policy 5439
cooperation 5222
said 5082
india 4964
one 4948
system 4928
weapons 4859
systems 4726
years 4684
first 4516
2006 4488
Headword
Freq.
china 2737
technology 1502
energy 1499
nuclear 1237
development 1023
use 1003
u.s. 9910
new 9895
country 9627
chinese 9171
plan 7338
can 6912
system
year
program
united states
power
two
fuel
government
world
develop
make
oil 5436
international 5412
emission 5358
support 5290
project 5277
nuclear power 5213
strategic 5188
provide 5164
policy 5102
strategy 4944
cooperation 4812
issue 4703
reactor 4676
need 4615
one 4525
state 4491
military 4448
increase 4419
reduce 4378
india 4363
research 4346
Concept
Freq.
be - Verb. Have the quality of being; ; (be) 2729
China, People's Republic of China, Communist China, mainland C... 2568
make up, constitute, comprise, be, represent - Verb. To make ... 1359
United States of America, United States, America - Pr.N. Federa... 1190
not - Adv. Adv. Negative adverb used to form structures... 1117
process, system, methodology, proceeding, way, procedure, m... 1109
American - Adj. Adj. Relating to the United States, its peop... 1106
program, plan - Noun. A method of doing something; a series..... 1031
nuclear, atomic - Adj. Adj. Using atomic energy (nuclear, atomi... 1011
this - Adj. Adj. Used to refer to the present event,... ... 8518
demand, involve, want, require, ask, need, necessitate, call for ... 7983
technology - Noun. The theoretical elaboration and application ... 7663
land, country, State, state, nation - Noun. The territory occupi... 7478
7427
7303
7123
7089
5579
6302
5858
5770
5765
5666
5351
5328
also, as well - Adv. Adv. In addition; as well (also, as well) 5300
chain reactor, atomic pile, nuclear pile, reactor, nuclear reactor -... 5209
development, passage, evolution, growth - Noun. A process in ... 5189
strategy, scheme - Noun. A carefully devised plan of action to..... 4944
area, geographical area, region, Area, country, geographic lands... 4709
house, business firm, company, enterprise, firm, establishment, ... 4675
oil, fossil oil, petroleum, coal oil, rock oil - Noun. Crude oil that o... 4560
ply, supply, furnish, provide, cater - Verb. To give, sell, or make ... 4558
further, encourage, advance, promote, boost - Verb. To promo... 4473
new - Adj. Adj. Not of long duration; having just come.... 4349
objective, end, purpose, target, object, aim - Noun. The goal in... 4315
say - Verb. To convey information or express feelings in... 4310
other, extra, additional - Adj. Adj. Further or added (other, ext... 4307
task, project, undertaking, labor - Noun. A particular piece of w... 4305
development - Noun. Act of improving by expanding or enlargin... 4169
energy, vigour, vigorousness, vigor - Noun. Liveliness or forceful... 4035
India, Republic of India, Bharat - Pr.N. A republic in S Asia; 2nd ... 3967
all - Adj. Adj. Quantifier; used with either mass or co... 3884
ameliorate, better, improve upon, amend, improve, meliorate - ... 3880
emission, discharge - Noun. A substance that is emitted or relea... 3870
state, commonwealth, State, body politic, nation, country, land... 3739
keyword: nuclear
frequency: 32,579
start
Posta in arrivo - Micr...
Benvenuto - MyEni - ...
SIG III_case study_...
SIG III_case study_...
COGITO® Intelligen...
11.18

Semantic area of "nuclear"

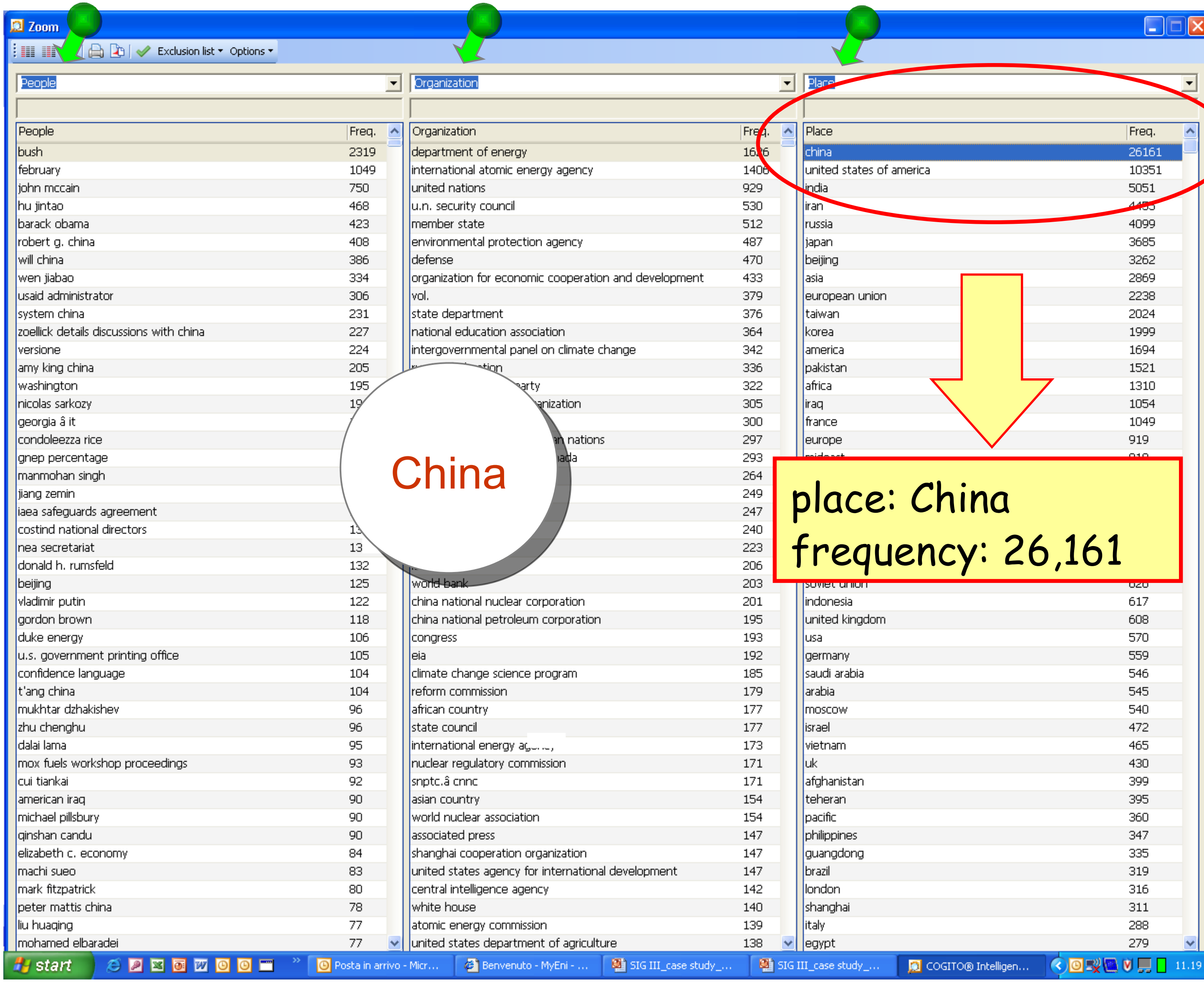
concept: "nuclear energy"
frequency: 3,332

concept: "nuclear power"
frequency: 3,384



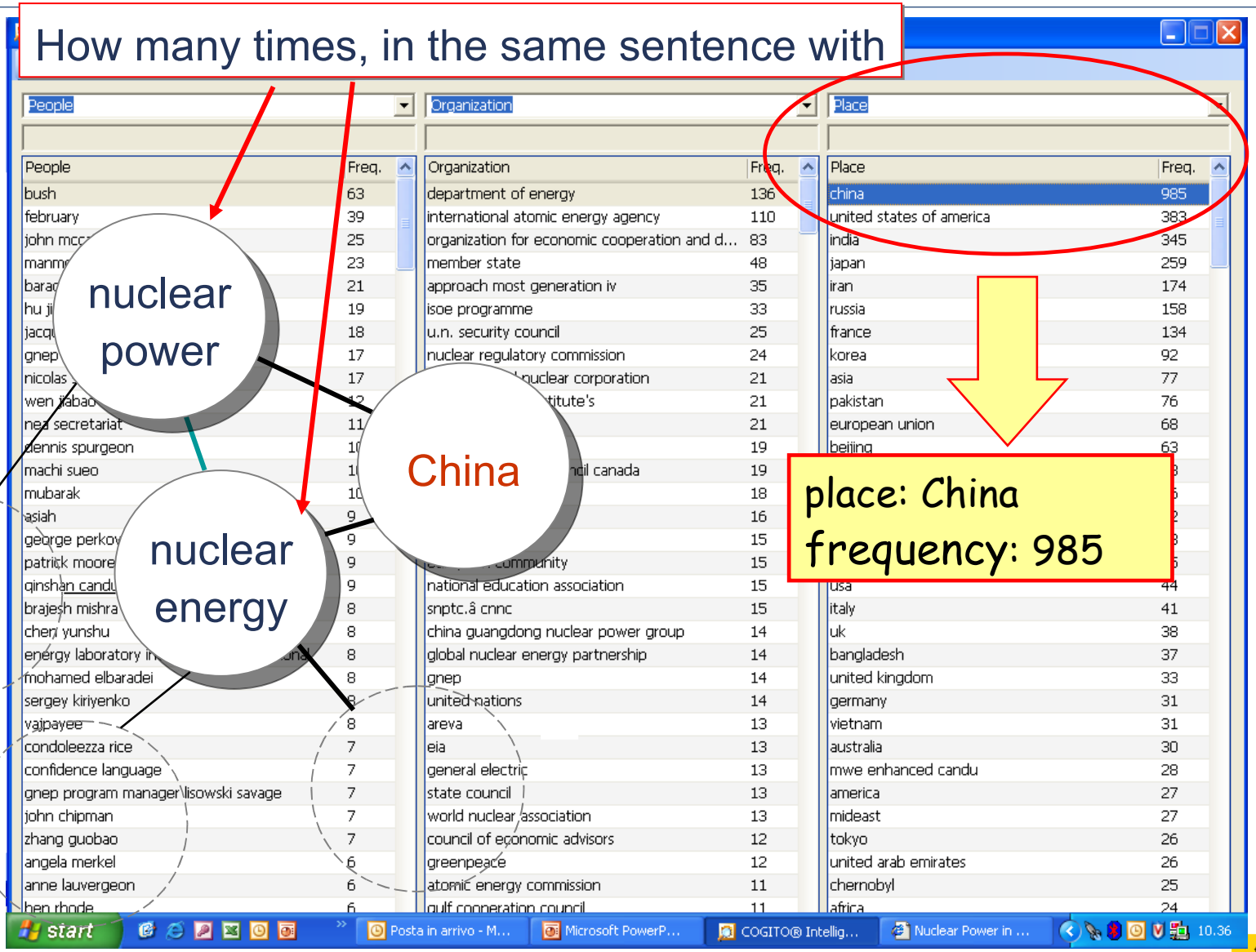
WARNING ! !

- nuclear weapon frequency: 2.796
- nuclear arsenal frequency: 197
- nuclear attack frequency: 112
- nuclear medicine frequency 13



People	Freq.	Organization	Freq.	Place	Freq.
bush	2319	department of energy	16[illegible]6	china	26161
february	1049	international atomic energy agency	1408	united states of america	10351
john mccain	750	united nations	929	india	5051
hu jintao	468	u.n. security council	530	iran	44[illegible]
barack obama	423	member state	512	russia	4099
robert g. china	408	environmental protection agency	487	japan	3685
will china	386	defense	470	beijing	3262
wen jiabao	334	organization for economic cooperation and development	433	asia	2869
usaid administrator	306	vol.	379	european union	2238
system china	231	state department	376	taiwan	2024
zoellick details discussions with china	227	national education association	364	korea	1999
versione	224	intergovernmental panel on climate change	342	america	1694
amy king china	205	[illegible]tion	336	pakistan	1521
washington	195	[illegible]arty	322	africa	1310
nicolas sarkozy	19[illegible]	[illegible]nization	305	iraq	1054
georgia â it			300	france	1049
condoleezza rice		[illegible]n nations	297	europe	919
gnep percentage		[illegible]ada	293	[illegible]	[illegible]
manmohan singh			264		
jiang zemin			249		
iaea safeguards agreement			247		
costind national directors	13[illegible]		240		
nea secretariat	13[illegible]		223		
donald h. rumsfeld	132		206		
beijing	125	world bank	203	soviet union	[illegible]
vladimir putin	122	china national nuclear corporation	201	indonesia	617
gordon brown	118	china national petroleum corporation	195	united kingdom	608
duke energy	106	congress	193	usa	570
u.s. government printing office	105	eia	192	germany	559
confidence language	104	climate change science program	185	saudi arabia	546
t'ang china	104	reform commission	179	arabia	545
mukhtar dzhakishev	96	african country	177	moscow	540
zhu chenghu	96	state council	177	israel	472
dalai lama	95	international energy ag[illegible]	173	vietnam	465
mox fuels workshop proceedings	93	nuclear regulatory commission	171	uk	430
cui tiankai	92	snptc.â cnnc	171	afghanistan	399
american iraq	90	asian country	154	teheran	395
michael pillsbury	90	world nuclear association	154	pacific	360
qinshan candu	90	associated press	147	philippines	347
elizabeth c. economy	84	shanghai cooperation organization	147	guangdong	335
machi sueo	83	united states agency for international development	147	brazil	319
mark fitzpatrick	80	central intelligence agency	142	london	316
peter mattis china	78	white house	140	shanghai	311
liu huaqing	77	atomic energy commission	139	italy	288
mohamed elbaradei	77	united states department of agriculture	138	egypt	279

How many times, in the same sentence with

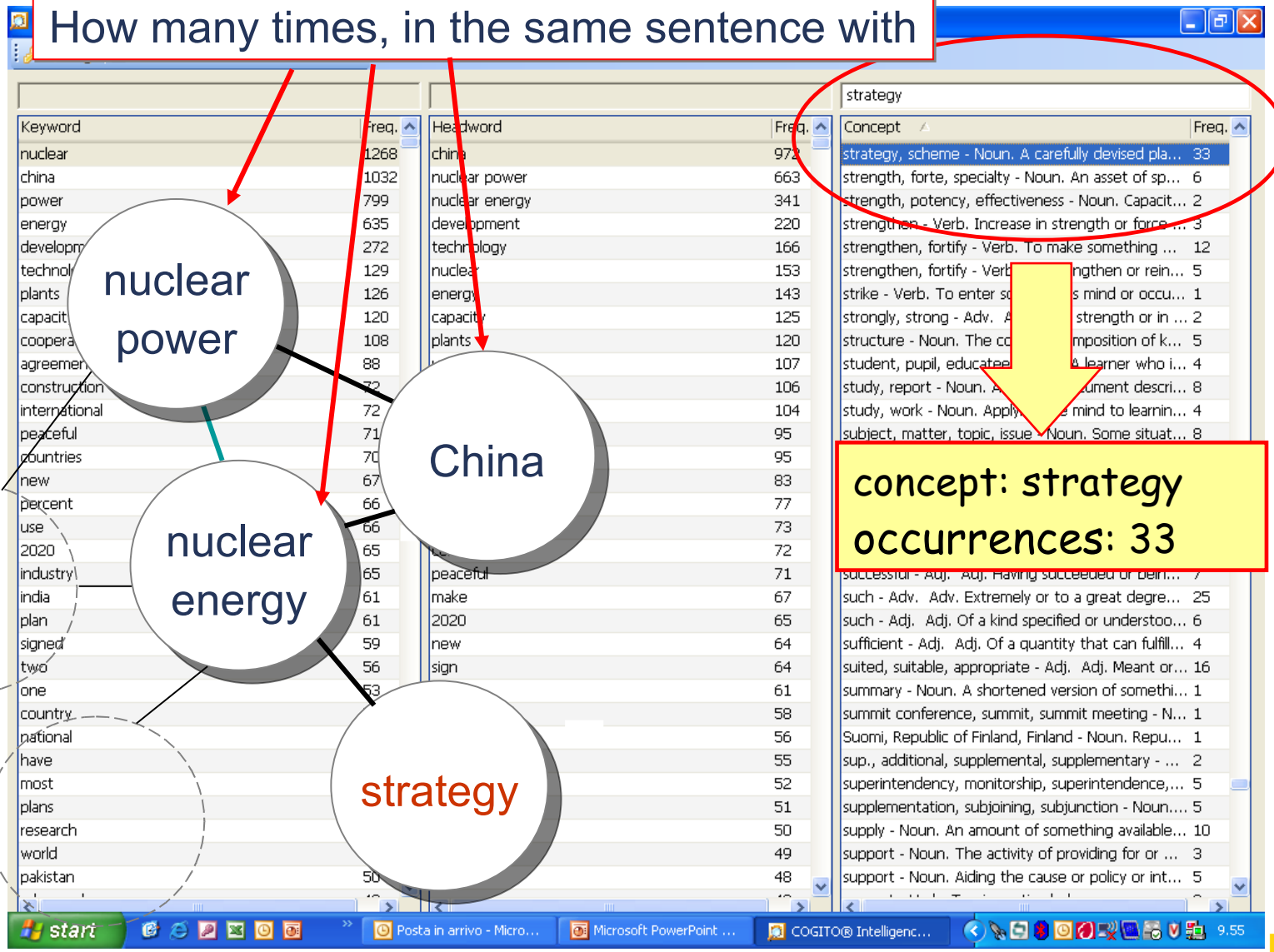


People	Freq.
bush	63
february	39
john mcc	25
manm	23
barac	21
hu ji	19
jacq	18
gnep	17
nicolas	17
wen jiabao	12
nea secretariat	11
dennis spurgeon	10
machi sueo	10
mubarak	10
asiah	9
george perkov	9
patrick moore	9
qinshan candu	9
brajesh mishra	8
chen yunshu	8
energy laboratory in	8
mohamed elbaradei	8
sergey kiriyenko	8
vajpayee	8
condoleezza rice	7
confidence language	7
gnep program manager lisowski savage	7
john chipman	7
zhang guobao	7
angela merkel	6
anne lauvergeon	6
ben rhode	6

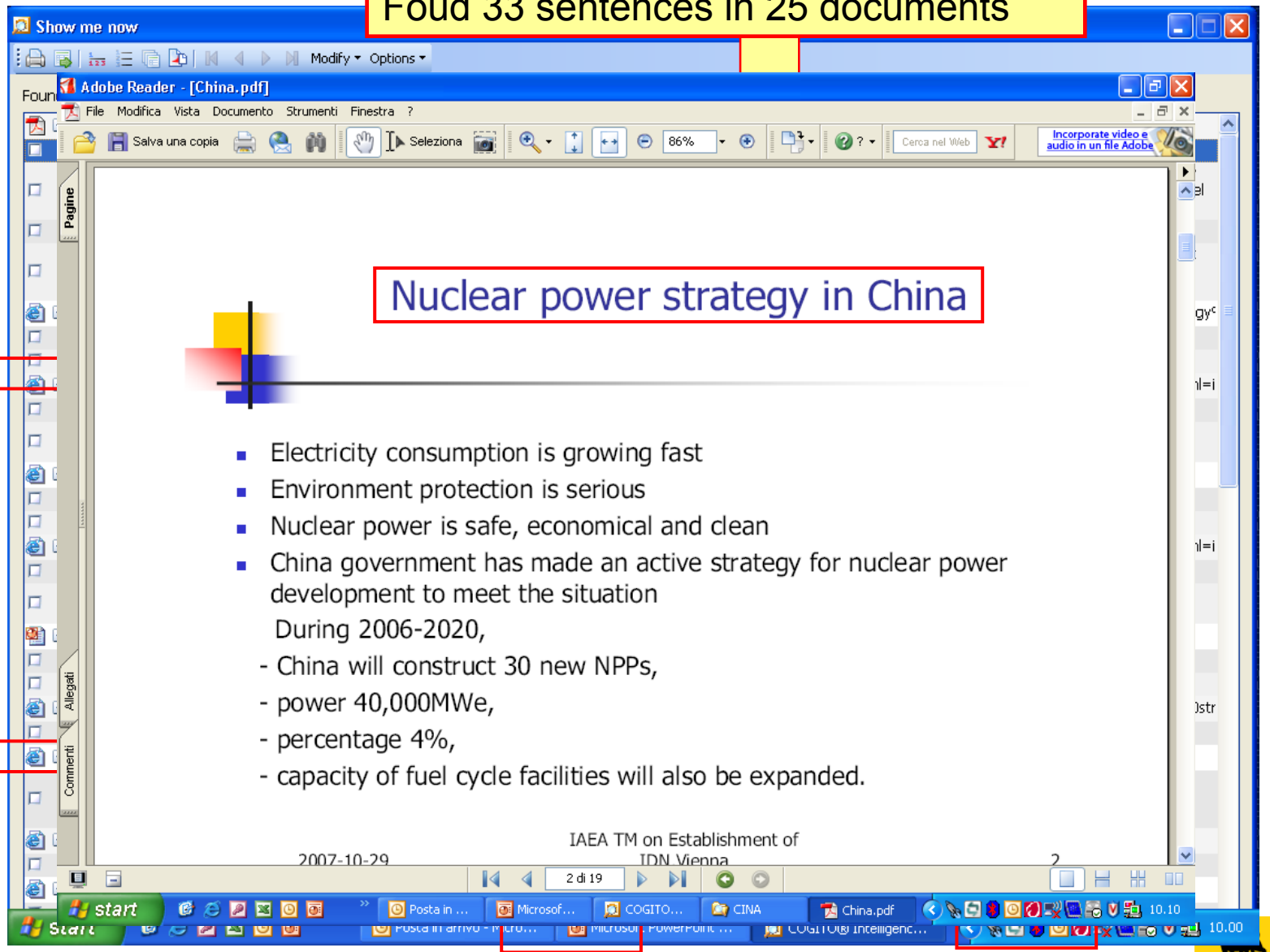
Organization	Freq.
department of energy	136
international atomic energy agency	110
organization for economic cooperation and d...	83
member state	48
approach most generation iv	35
isoe programme	33
u.n. security council	25
nuclear regulatory commission	24
nuclear corporation	21
titute's	21
	21
	19
ndl canada	19
	18
	16
	15
community	15
national education association	15
snptc.â cnnc	15
china guangdong nuclear power group	14
global nuclear energy partnership	14
gnep	14
united nations	14
areva	13
eia	13
general electric	13
state council	13
world nuclear association	13
council of economic advisors	12
greenpeace	12
atomic energy commission	11
gulf cooperation council	11

Place	Freq.
china	985
united states of america	383
india	345
japan	259
iran	174
russia	158
france	134
korea	92
asia	77
pakistan	76
european union	68
beijing	63
	3
	6
	2
	3
	5
usa	44
italy	41
uk	38
bangladesh	37
united kingdom	33
germany	31
vietnam	31
australia	30
mwe enhanced candu	28
america	27
mideast	27
tokyo	26
united arab emirates	26
chernobyl	25
africa	24

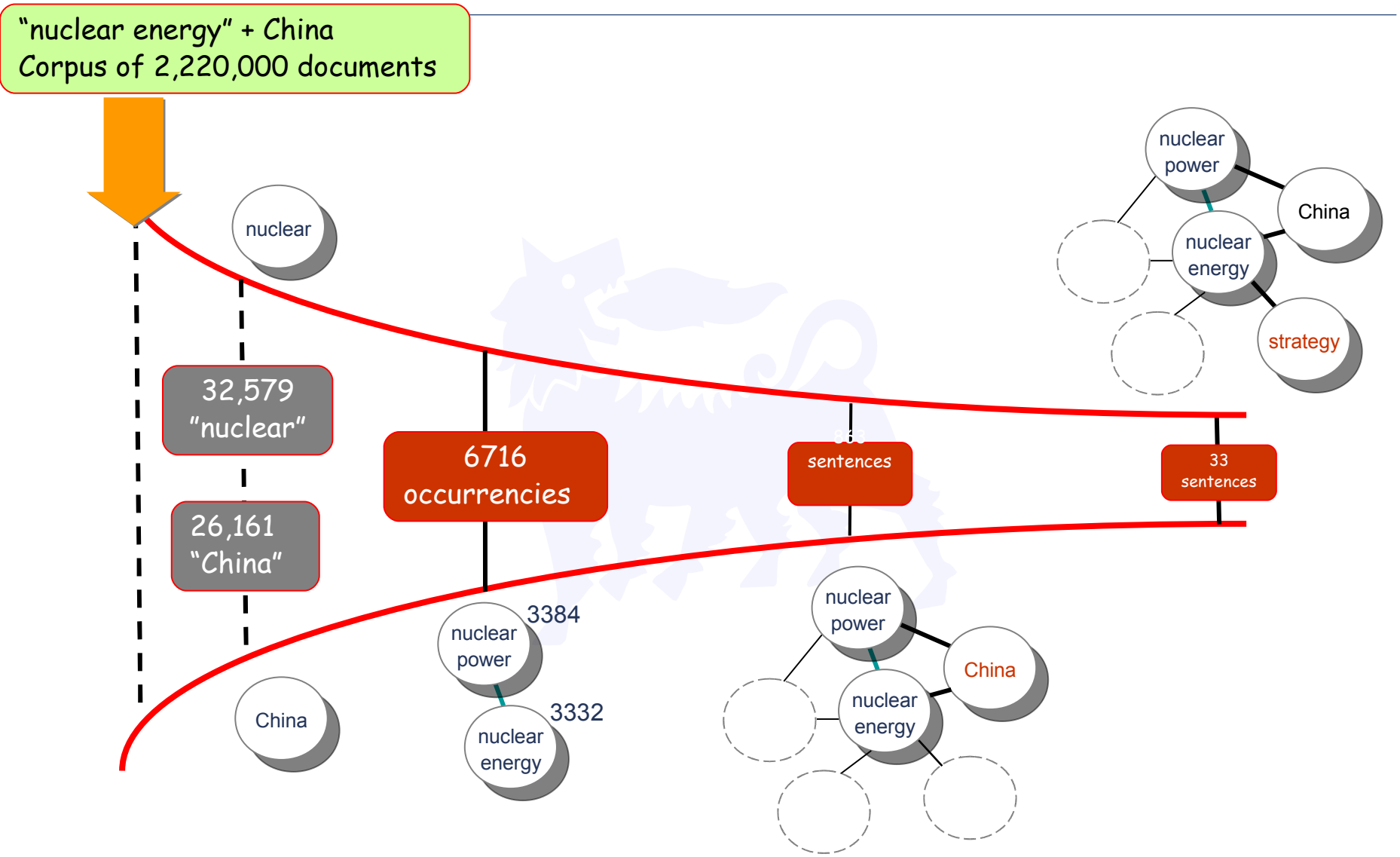
How many times, in the same sentence with



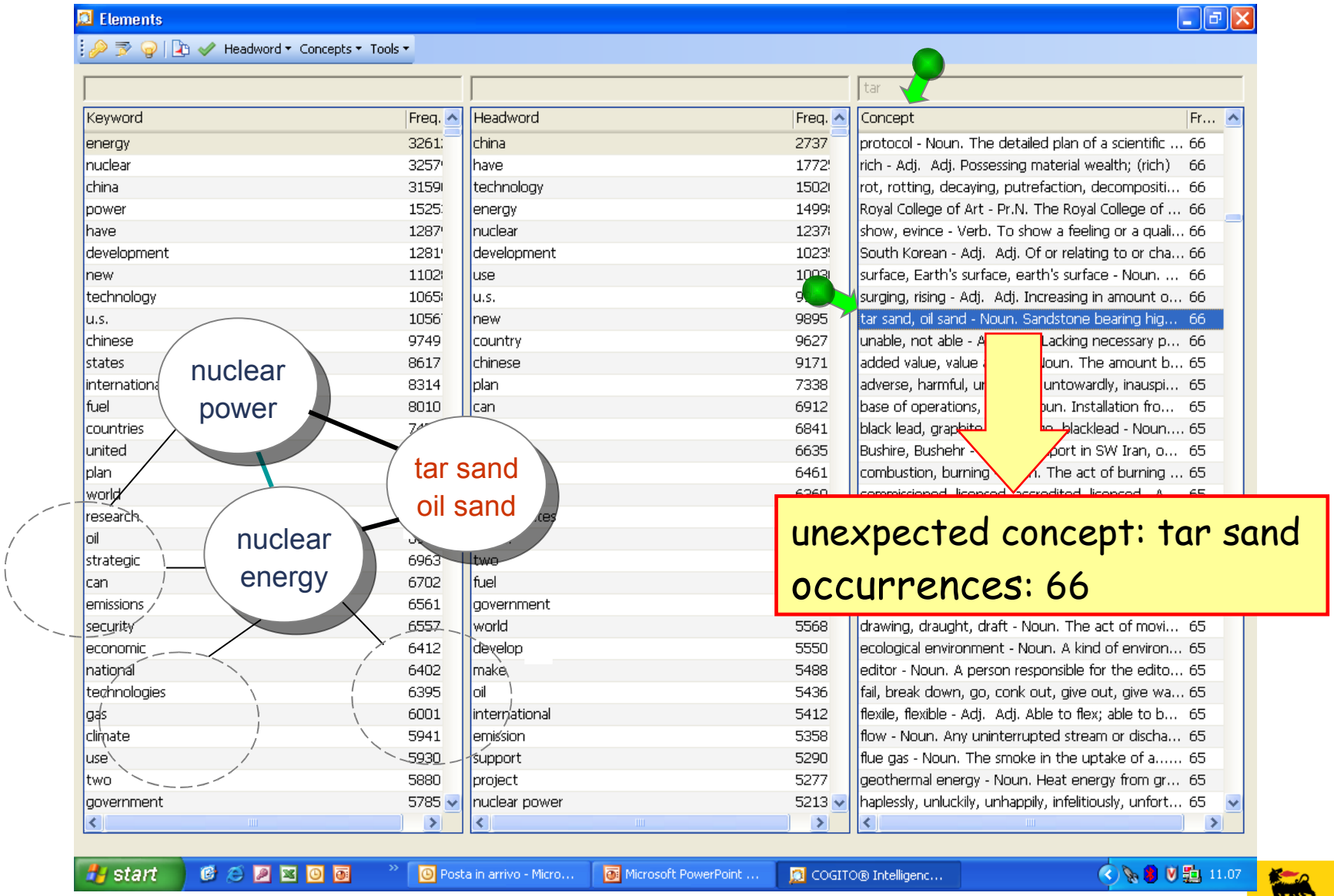
Foud 33 sentences in 25 documents

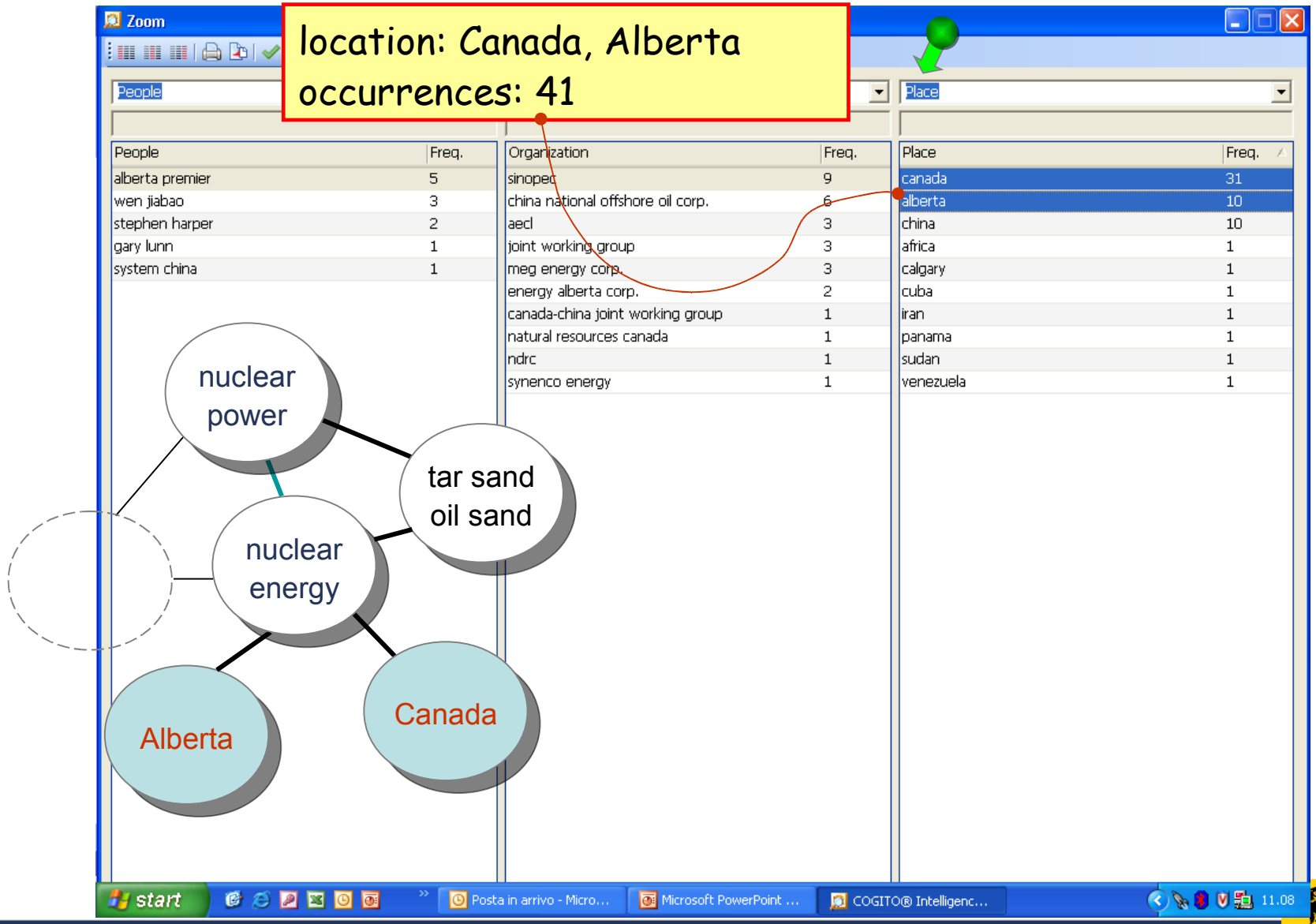


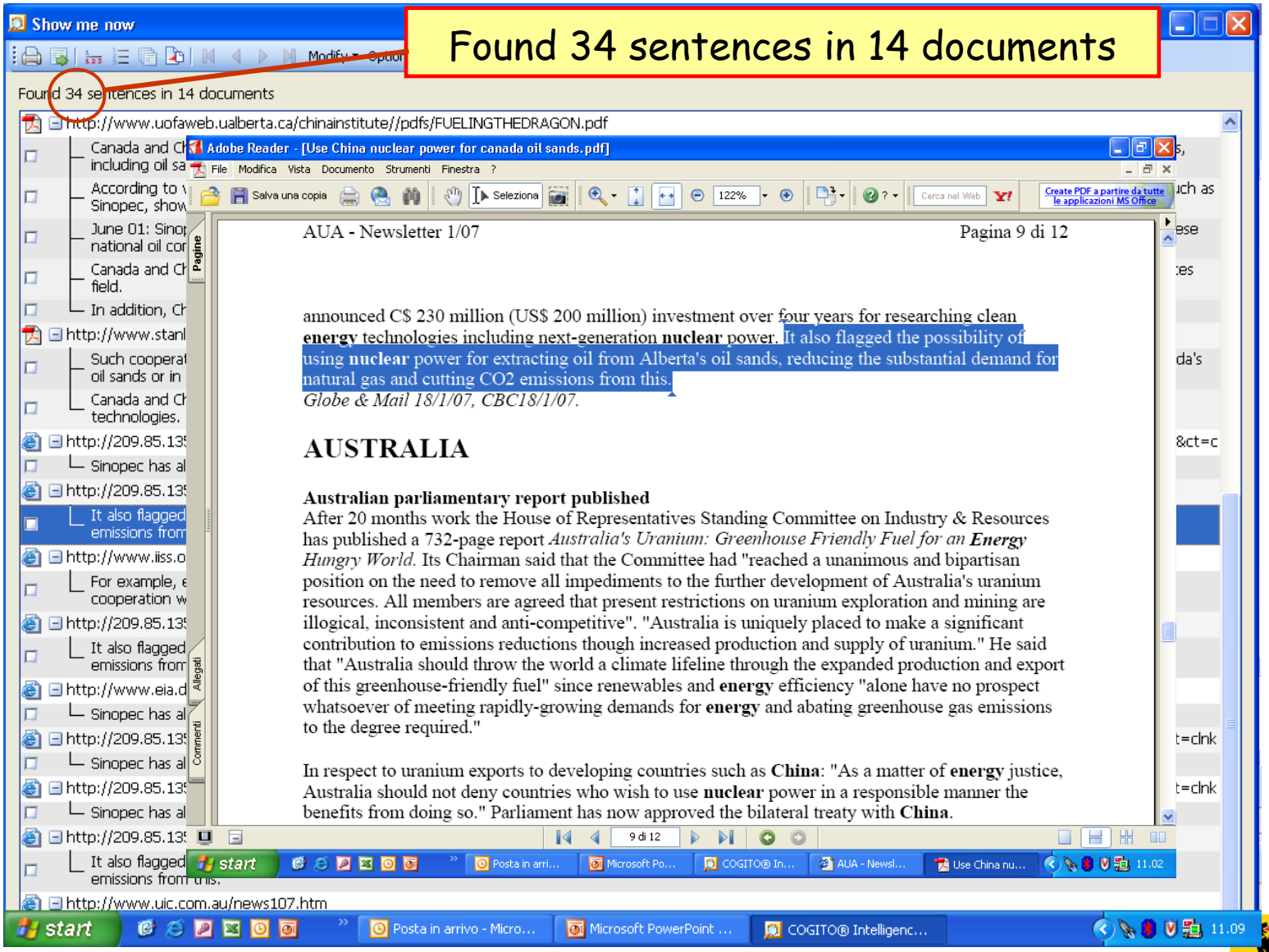
The virtuous funnel: from collection to connection



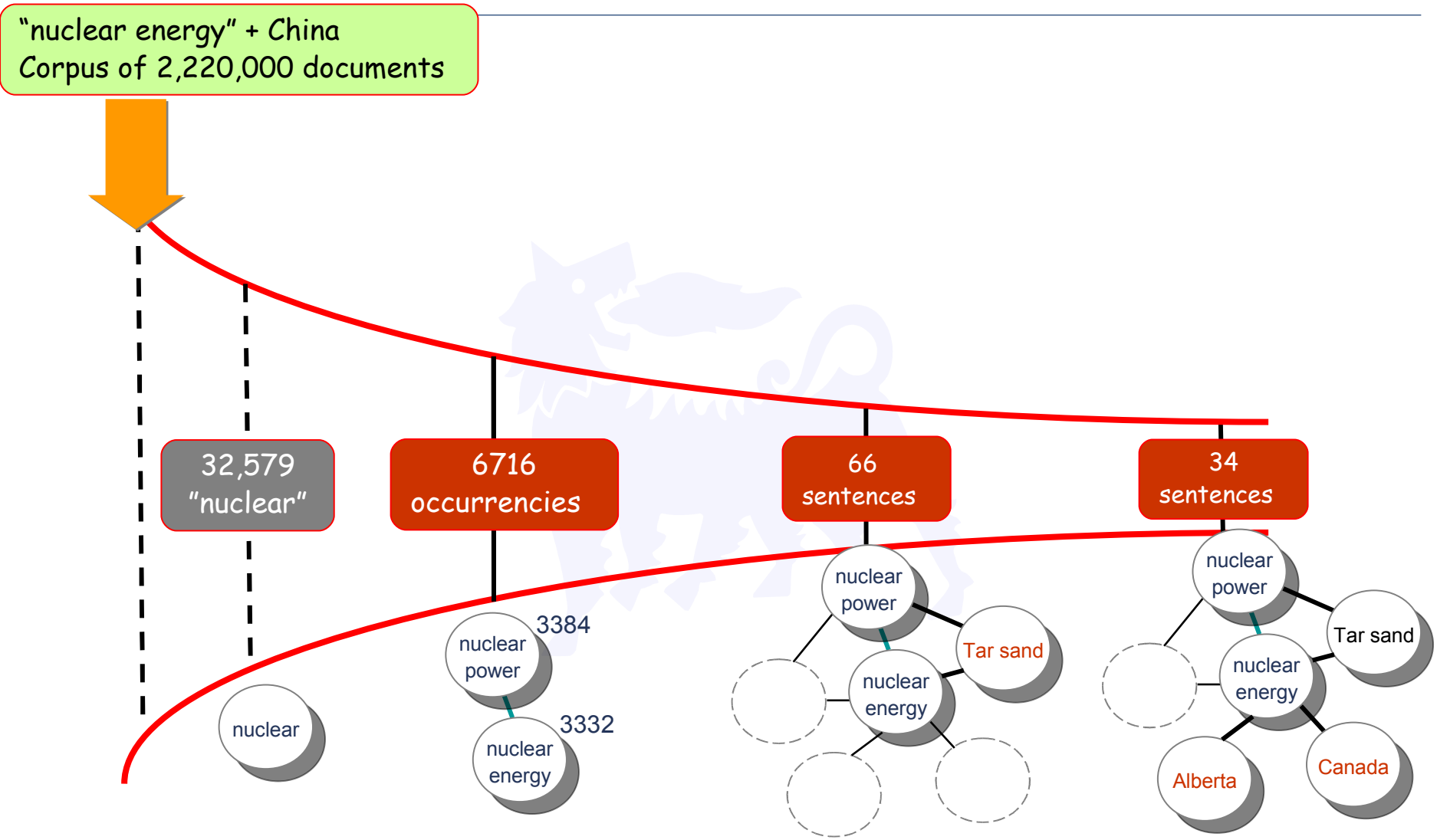
Serendipity, maybe




Zoom
location: Canada, Alberta
occurrences: 41
People
Place
People Freq.
alberta premier 5
wen jiabao 3
stephen harper 2
gary lunn 1
system china 1
Organization Freq.
sinopec 9
china national offshore oil corp. 6
aecl 3
joint working group 3
meg energy corp. 3
energy alberta corp. 2
canada-china joint working group 1
natural resources canada 1
ndrc 1
synenco energy 1
Place Freq.
canada 31
alberta 10
china 10
africa 1
calgary 1
cuba 1
iran 1
panama 1
sudan 1
venezuela 1
nuclear power
tar sand oil sand
nuclear energy
Alberta
Canada
start
Posta in arrivo - Micro...
Microsoft PowerPoint ...
COGITO® Intelligenc...
11.08

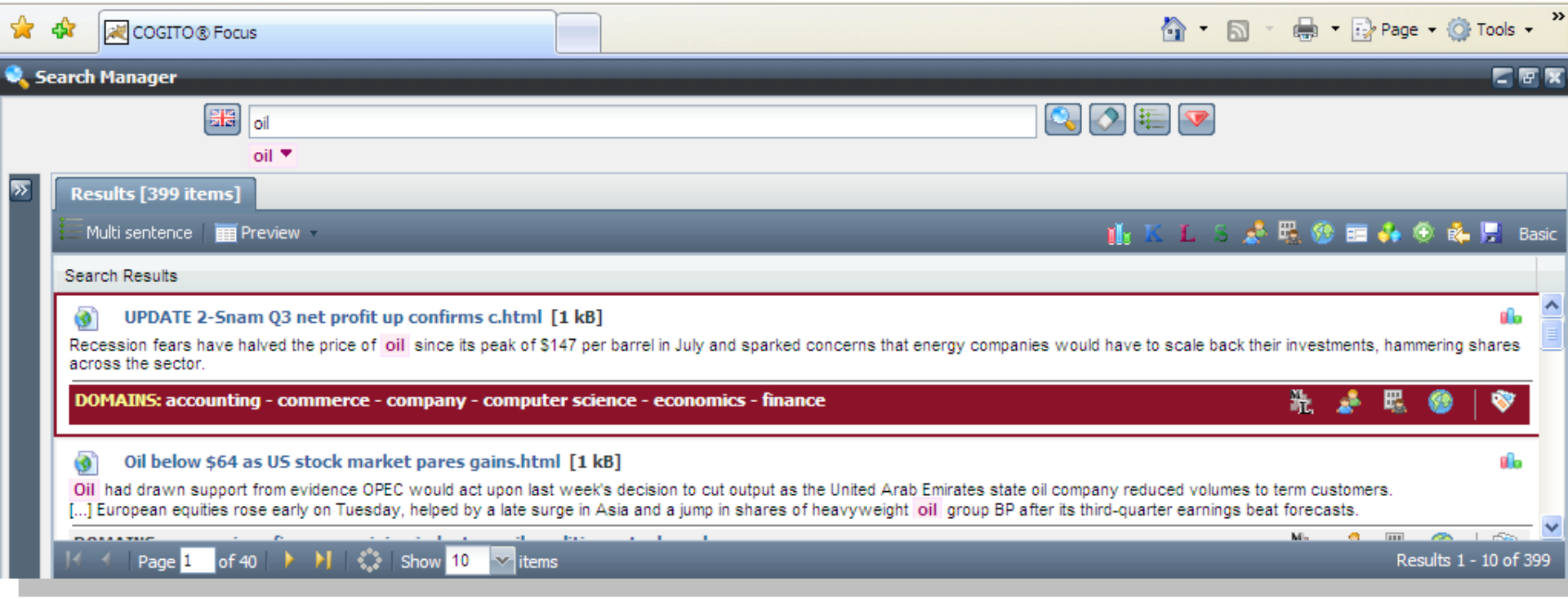

Found 34 sentences in 14 documents
Show me now
Found 34 sentences in 14 documents
http://www.uofaweb.ualberta.ca/chinainstitute//pdfs/FUELINGTHEDRAGON.pdf
Adobe Reader - [Use China nuclear power for canada oil sands.pdf]
AUA - Newsletter 1/07
Pagina 9 di 12
announced C$ 230 million (US$ 200 million) 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, CBC18/1/07.
AUSTRALIA
Australian parliamentary report published
After 20 months work 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 "reached 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 though 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 whatsoever of meeting rapidly-growing demands for energy and abating 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.
http://www.uic.com.au/news107.htm
start
Posta in arrivo - Micro...
Microsoft PowerPoint ...
COGITO® Intelligenc...
11.09

The knowledge funnel



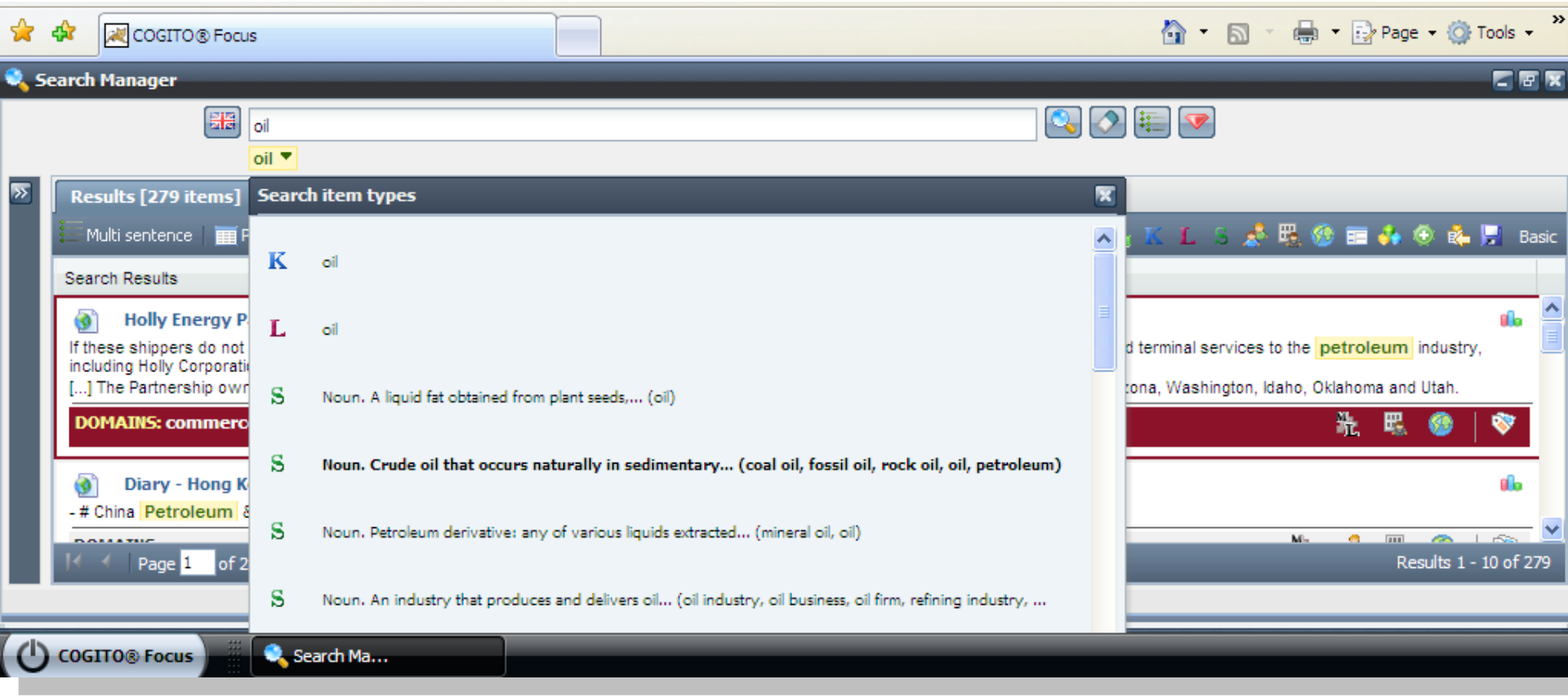
3-Step Walk With Semantic Search

Oil as keyword = 399 results



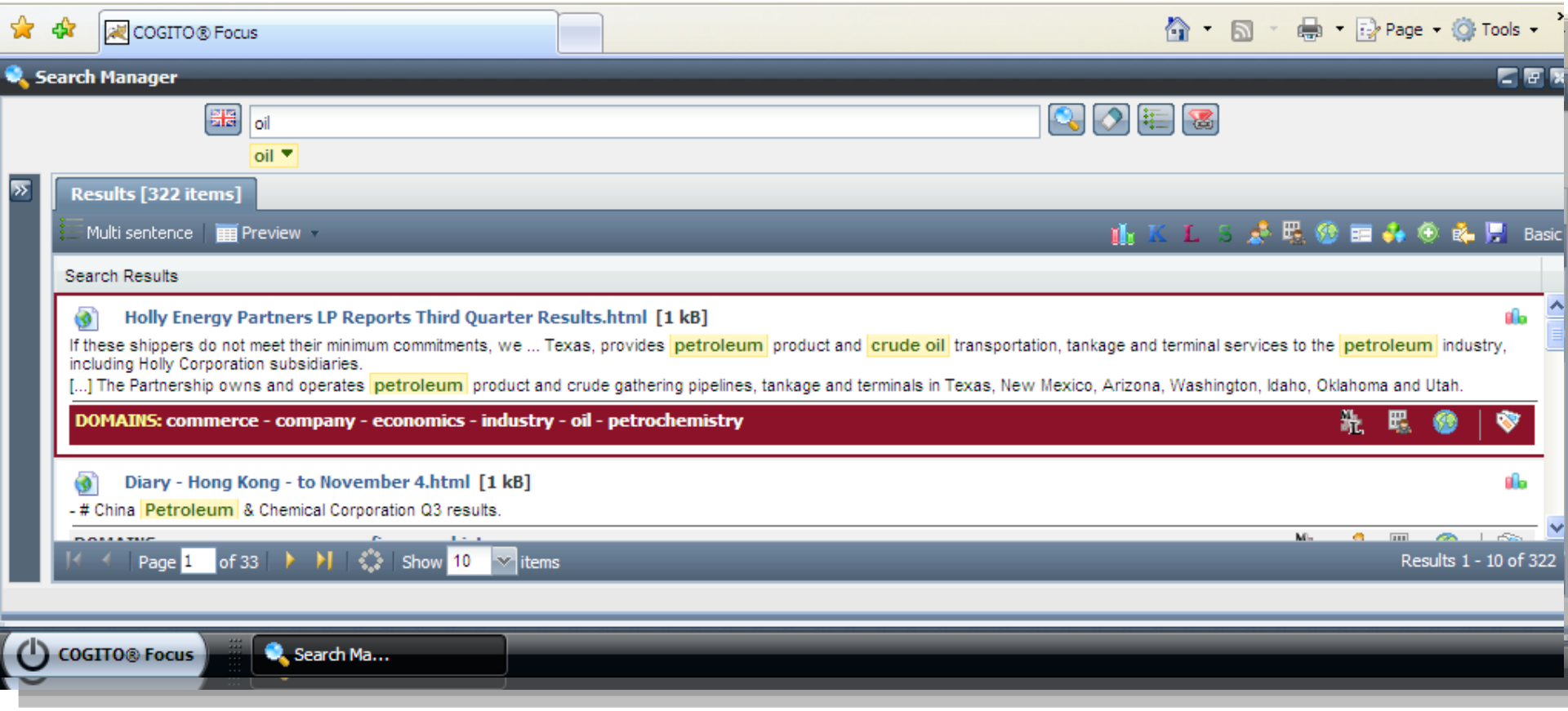
3-Step Walk With Semantic Search

Oil as concept = 279 results



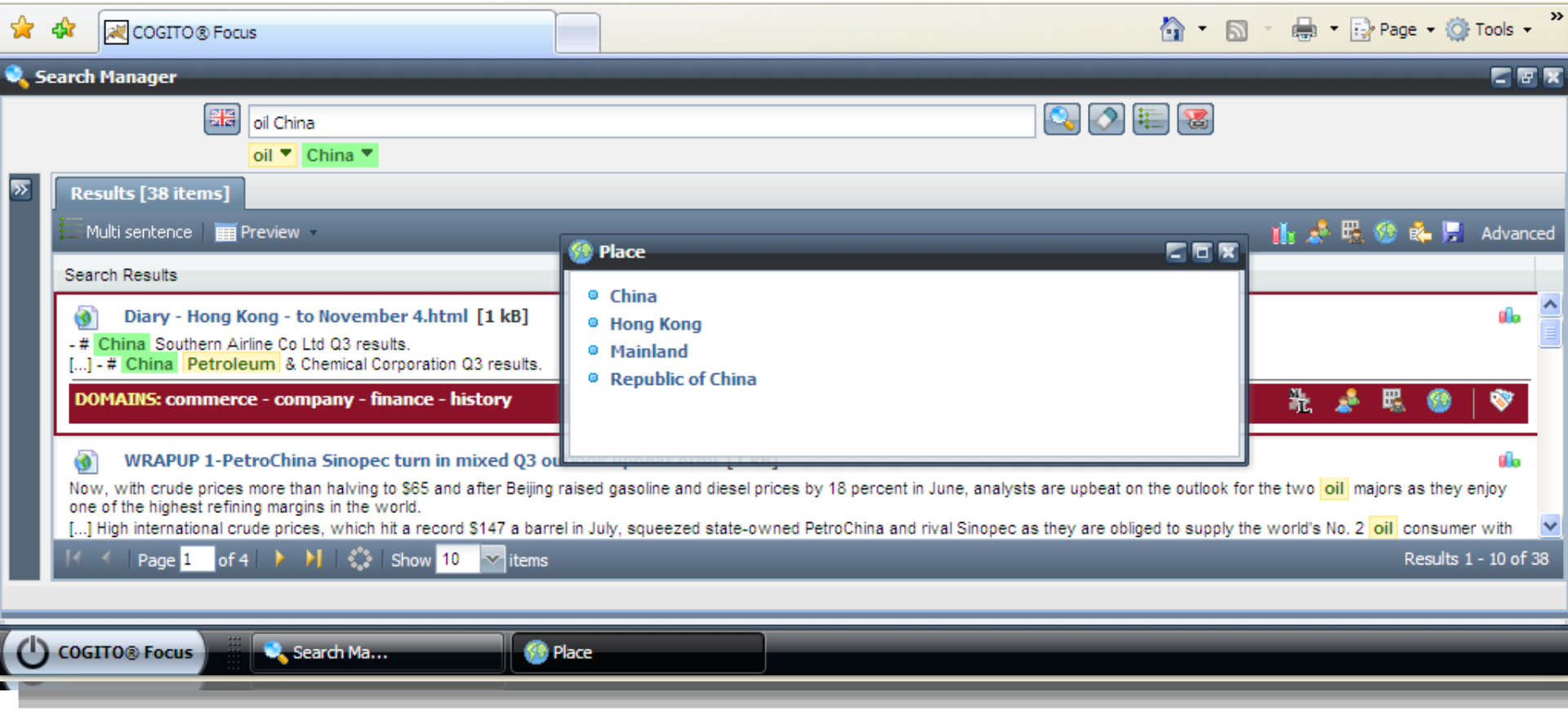
3-Step Walk With Semantic Search

Oil as a semantic concept = 322 results

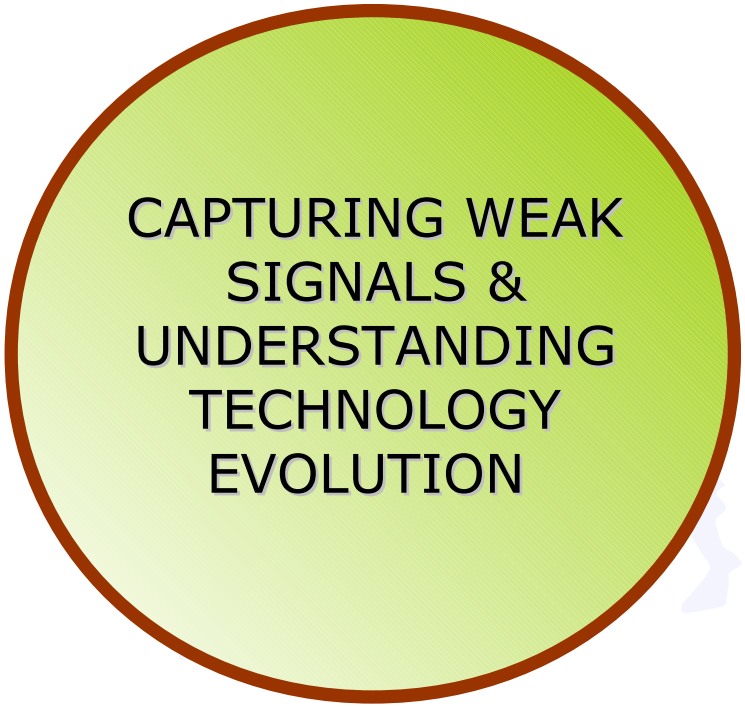


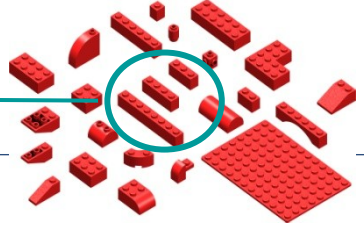
3-Step Walk With Semantic Search

Oil as a semantic concept & objects = 38 results



Use Case Two


CAPTURING WEAK SIGNALS & UNDERSTANDING TECHNOLOGY EVOLUTION



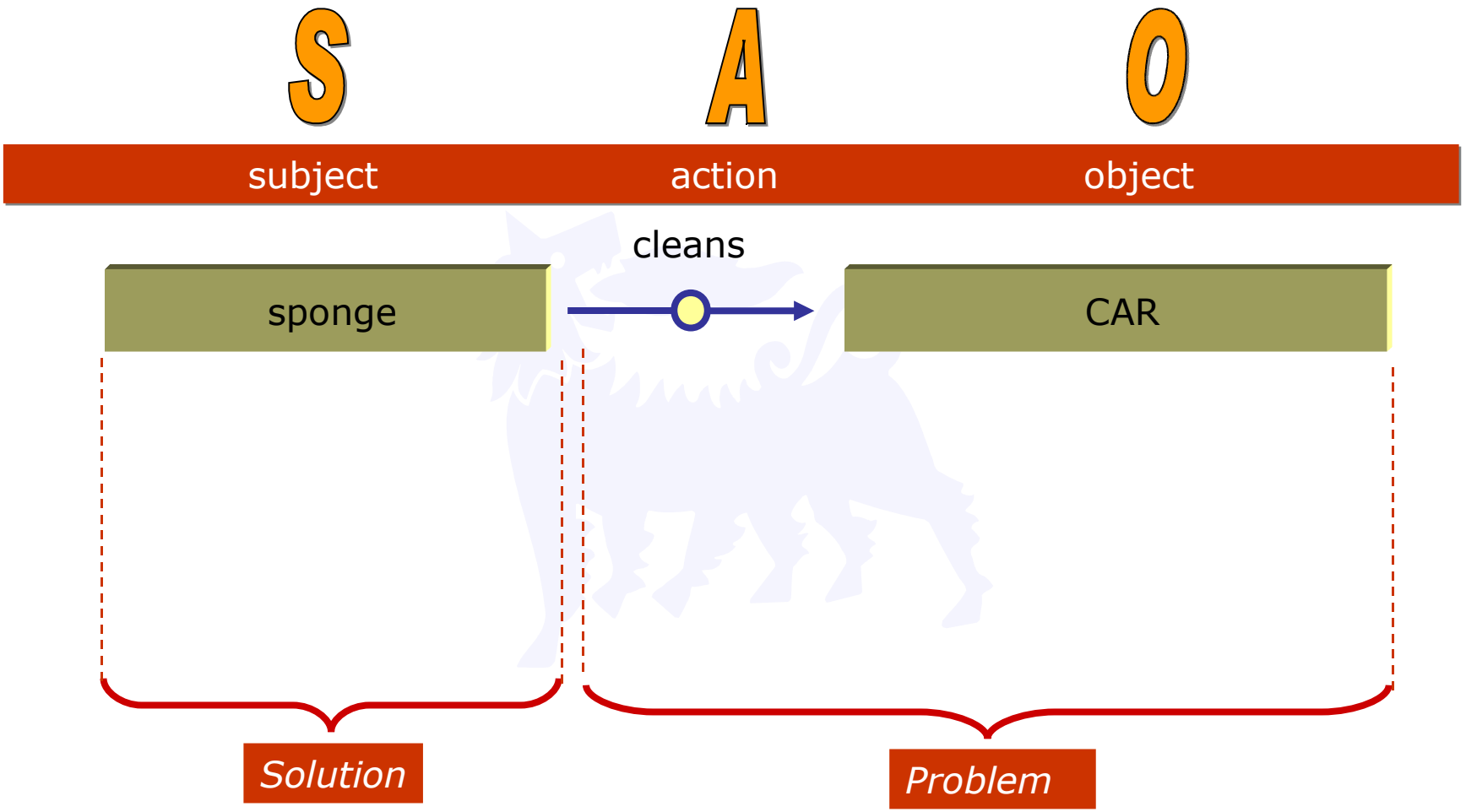
Playing with a very popular "trio"

Subject

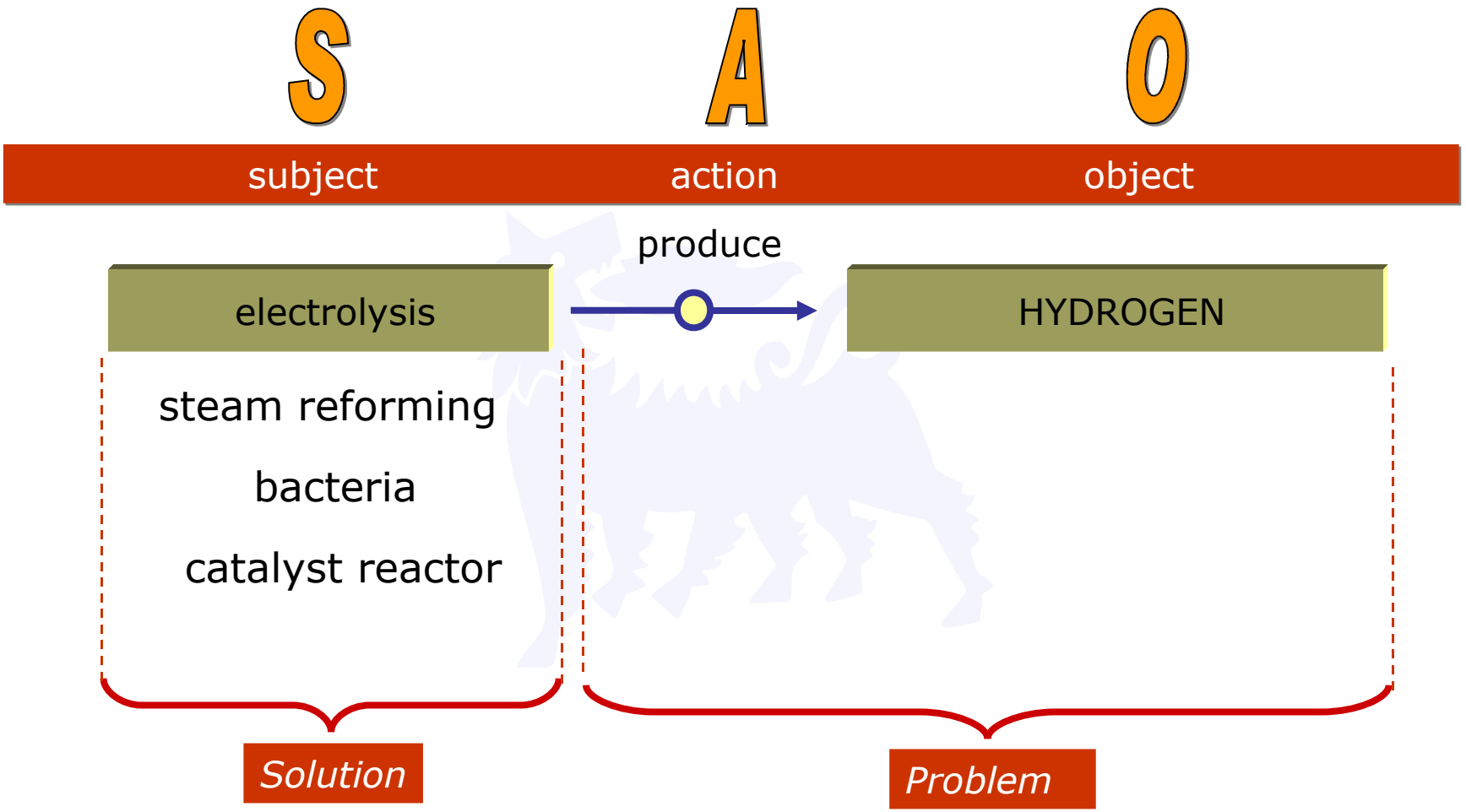
Action

Object

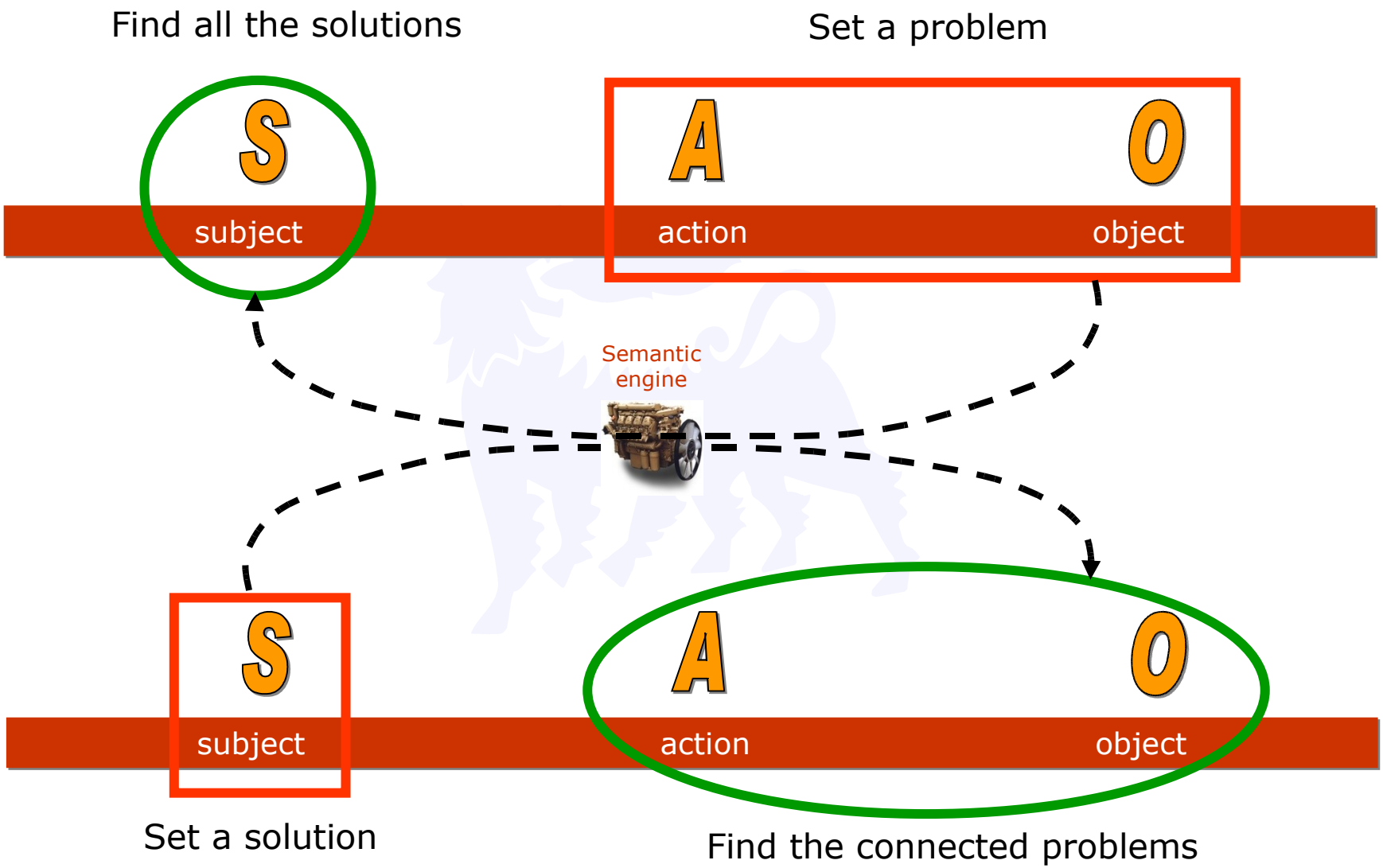
SAO structure identification can boost technical problem solving



SAO structure identification can boost technical problem solving



Extracting knowledge through semantic queries on SAO's



Capturing weak signals of "breakthrough"

S A O

subject | action | object

enhance

increase pressure → OIL RECOVERY

Inject gas

Inject polymers

Problem

Inject criogenic liquids

Ultrasonic stimulation

Magnetic stimulation

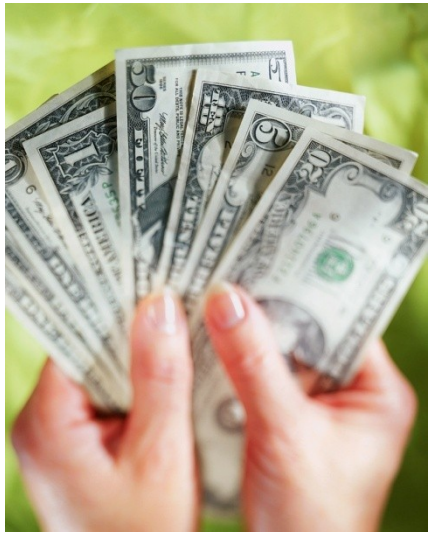
Solution

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