



***Linguistic-semantic  
Applications for Text Filtering  
Cogito®***

*the Linguistic Platform of Expert System*

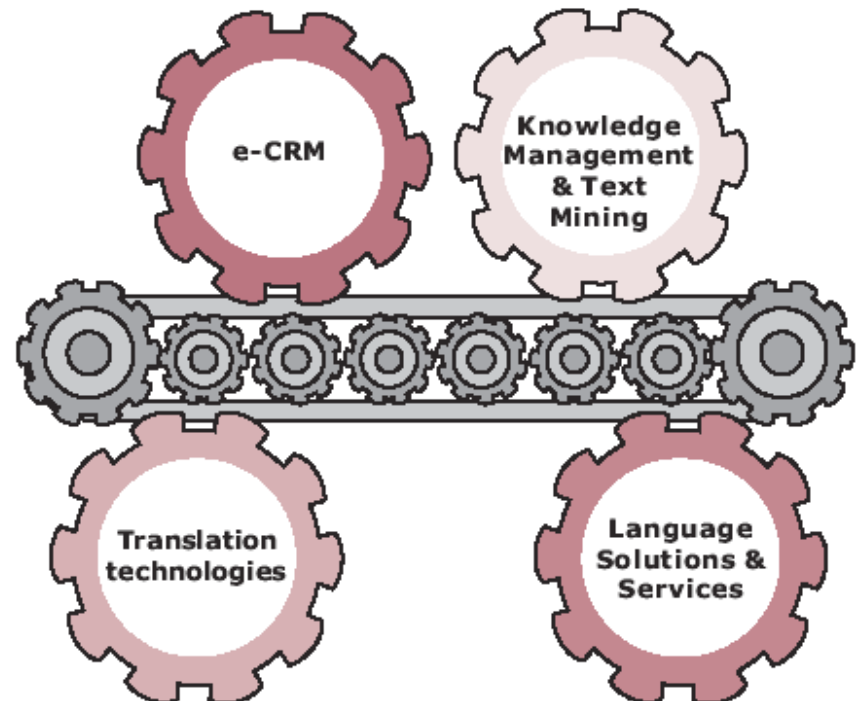
*Massimo Vologni  
Expert System*

## ***Who is Expert System***

Expert System has been operating in the field of linguistics since 1989. After many years' experience in the development of software for the consumer market, it has turned to the market of business solutions.

Today, the company consists of three divisions:

- **eCRM,**
- **Knowledge Management & Text Mining**
- **Language solutions and services**





Cogito® is able to elaborate in an intelligent way the knowledge contained in texts written in the everyday language, for the retrieval of meanings, the comprehension of natural language, the sharing and circulation of knowledge, etc...



Thanks to the most advanced linguistic technologies, Cogito® recognizes all the aspects:

- structure
- vocabulary
- semantics

in the texts, through a deep analysis and disambiguation of all the found elements, thanks to a rich and complex representation of the real world.

Cogito® is:

- **scalable**
- **customizable**
- **transversal**



**About... Scalability**



**About... Customization**



**About... transversality**



## ***Some Cogito® components***

- 1) ITALIAN / ENGLISH SENSIGRAFO**
- 2) MORPHOLOGICAL, GRAMMATICAL AND SYNTACTICAL PARSER**
- 3) THESAURUS**
- 4) ADVANCED GRAMMAR CHECK**
- 5) ITALIAN-ENGLISH / ENGLISH-ITALIAN DICTIONARY**
- 6) ITALIAN VERTICAL DICTIONARIES**
- 7) CORPUS**
- 8) MEANINGS DISAMBIGUATOR**

# ***The Behavior of Semantic Network and Parsing***

The components of Cogito® can work both alone and in collaborative mode.

## **Sensigrafo Parsing**

### **Steps of the analysis:**

- creation and maintenance of a semantic network to code knowledge;**
- reading of the text and extraction of the elements of the sentences (engine);**
- comparison with the network of meanings;**
- retrieval of the concepts from the analyzed text;**
- returning of the meaning (not ambiguous) of the text.**



## *Focus on: the Disambiguator*

If you want to ask a program to distinguish between:

**"The rust eats the tower"**  
and **"The knight eats the castle"**

You must first teach a software to think "like a person".

### **Disambiguation:**

this is the true problem with  
automatic text interpretation.





## ***How the disambiguator works***

Es: the word CARRY can have various interpretations.

***The students carry law textbooks to the lesson***

***Some libraries carry law textbooks***

***The new carry law is not deterring crime***





## DISAMBIGUATING, and then...?

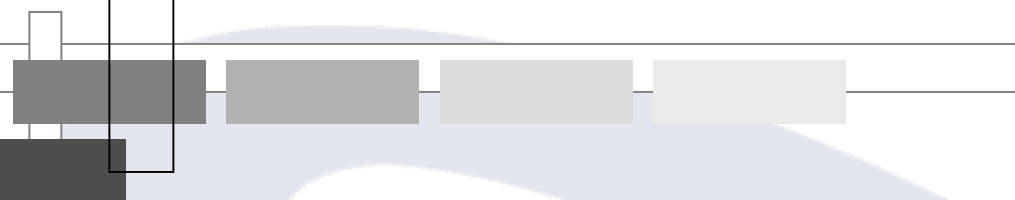
A linguistic technology platform like COGITO is a key element to develop products for the

**text and information mining  
In function of filtering.**



# **Text and information Mining**

Semantics for the filtering of contents  
from a sea of documents.



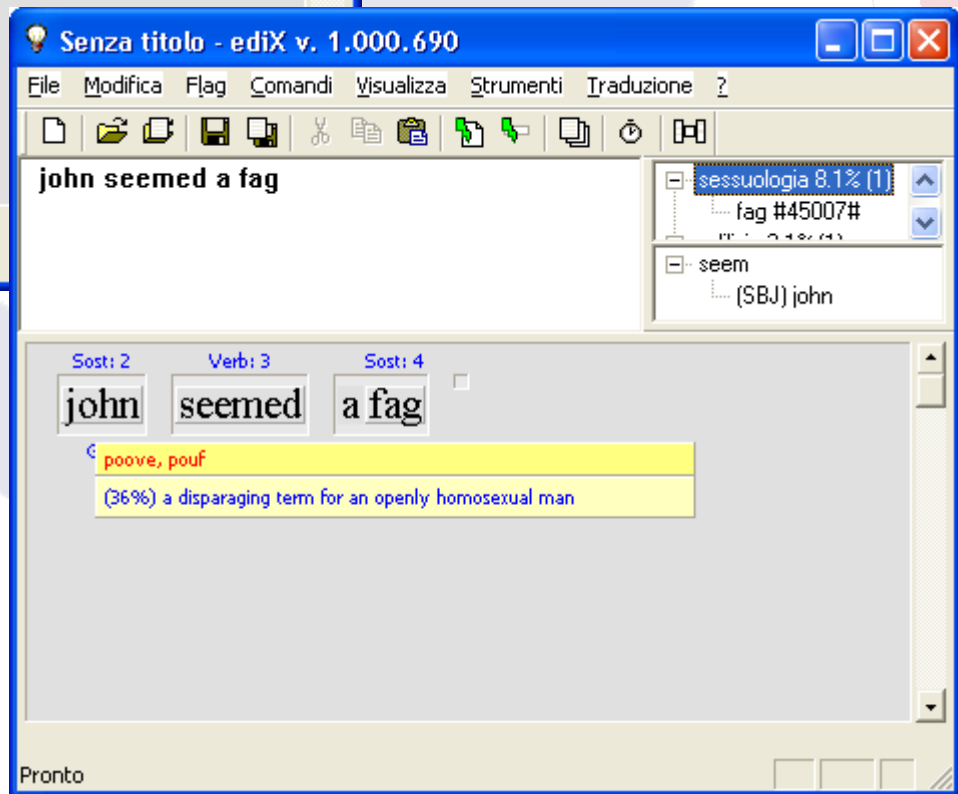
With the ever increasing amount of potentially useful information available on Internet and Intranet, the solutions for Information Mining offered by Expert System make it possible

**to find and filtering information required quickly**

The solutions offered by Expert System for Information Mining make use of language technology to look for, to organise, to elaborate and filtering unstructured contents.

The problem is to understand the double sense of a word and correctly stop only the vulgar/sexual talk and contents.

The filter must be applied on the meaning of the word and not on a specific sequence of signs





# The semantic-net of Cogito ®

**S #25876**  
**shaft, tool, cock,**  
**"pecker", "peter",**  
**"prick", "dick"**  
 Nessun dominio  
 obscene terms for penis

**S #7589**  
**fowl, poultry,**  
**domestic fowl**  
 zoologia  
 a domesticated  
 gallinaceous bird

**S #7600**  
**chicken, "Gallus**  
**gallus", "chook"**  
 zoologia  
 bred for flesh or  
 eggs; believed to

**S #7603**  
**cock, rooster**  
 ornitologia  
 adult male  
 chicken

**V #69433**  
**cock**  
 armi  
 set the trigger of  
 a firearm back for  
 firing

**S #94635**  
**"Chanticleer"**  
 letteratura  
 a rooster,  
 especially in fairy  
 tales (literary)

**S #7605**  
**cockrel**  
 zoologia  
 a young domestic  
 cock; not older  
 than one year

**S #141870**  
**"chantecler",**  
**"chanticleer"**  
 storia medievale  
 a rooster: used as  
 a proper name in

Seleziona

cock    cock    Attiva    Link: iperonimi/iponimi    100%



# ***Examining Real-Time streams of data***

E-mails, SMS, contributions to forums, chats and news feeds can be object of intense real-time analysis.

The Beagle Monitor is able **to “sniff out” pieces of information** deemed to be particularly critical or dangerous, thus constituting a valuable support tool for these needs.

## **semantic net**

so it is possible to isolate people and/or situations with the help of even very complex lexical structures of synonyms and hyponyms.



**Beagle - Monitor**

SMS Entrati:       SMS Filtrati:       SMS Usciti:

Ciao gaetano io dgt da cagliari  
 X Mod: Ma fino a che ora si chatta stasera?  
 ENTRA AMICI4  
 coglione sarai tu! fuck yuo  
 mi avete stancato siete dei coglioni  
 dicono che pamela anderson è una maiala e  
 anche se è un culattone reeves è il più figo  
 sono incazzato nero, devo tirarmi su, c'è qu  
 stasera nessuna raga? Modche dici vado a d  
 leccami tutta  
 ma vai a farti una sega  
 penso anch'io che non se ne salvi nessuno,  
 che vadano a quel paese, berlusconi, fini, d'  
 Mod ok ho capito esco qui sono tutti impeg  
 NN MI DICHI NULLA DELLA MIA PROPOSTA?  
 cico79 meglio l'arcuri della ferilli è + giovar  
 vincent cassel sa di SESSO  
 leonardo di caprio leonardo di caprio e anco  
 anch'io sono di milano, dove abiti, io vicino  
 sono del leone e tu?  
 io xo' il mio organo te ti metto nel culo

**Personaggi**

---

**Volgarità:**

sono incazzato nero, devo tirarmi su, c'è qu  
leccami tutta  
ma vai a farti una sega  
 penso anch'io che non se ne salvi nessuno,  
 che vadano a quel paese, berlusconi, fini, d'  
 Mod ok ho capito esco qui sono tutti impeg  
 cico79 meglio l'arcuri della ferilli è + giovar  
 vincent cassel sa di SESSO  
 io xo' il mio organo te lo metto nel culo

X Mod: Ma fino a che ora si chatta stasera?  
 ENTRA AMICI4  
 stasera nessuna raga? Modche dici vado a d  
 NN MI DICHI NULLA DELLA MIA PROPOSTA?  
 leonardo di caprio leonardo di caprio e anco  
 anch'io sono di milano, dove abiti, io vicino  
 sono del leone e tu?  
 Sempre il mio pensiero è rivolto a te realtà c  
 un salutone a tutti quelli di roma  
 ma ci sono io, chattiamo subito subito  
 anche rutelli non scherza  
 Il mio lui deve essere:DOLCE ROMANTICO S  
 qualcuno di roma??  
 ti ricoprirei tutta di baci  
 sono gaia, vi va di parlare con me?  
 anche se ha giocato male la juve rimane la r  
 ancora x max, 6 veramente simpatico  
 oh no, bertinotti sicuramente  
 cerco ragazze per parlare  
 ciao, sono gio79, sono di milano e tu?  
 ciao a tutti chi vuole chiaccherare con me?

Il tuo SMS:

Input of sms

filtering

Output of sms

## Conceptual Information Search

Expert System's language technology is able to **extrapolate concepts** and deal with the information within a document in an intelligent manner.

Text analysis is based on the recognition of concepts and is independent of the terms used to define them, **thus forming cognitional maps of texts.**



**Filtering information  
using  
Natural Language**



[www.expertsystem.it](http://www.expertsystem.it)

Address:

Via Virgilio, 56/Q  
41100 Modena - Italy

Phone: +39 059 894011

Fax: +39 059 894099

e-Mail: [mvologni@expertsystem.it](mailto:mvologni@expertsystem.it)

Business information: [info@expertsystem.it](mailto:info@expertsystem.it)