

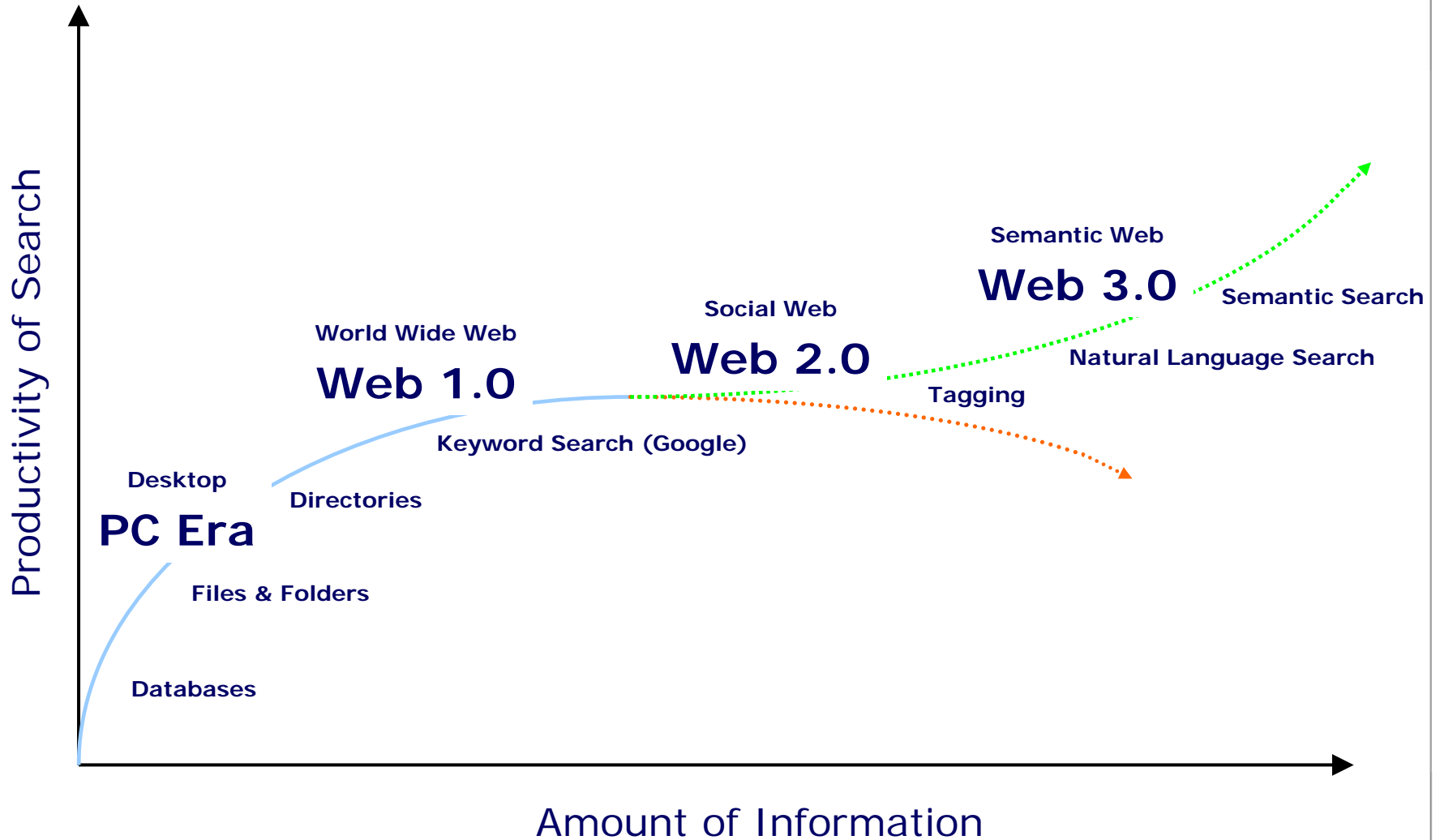


Whales & Cat Fur: Using A Semantic Net To Improve Precision & Recall

Semantic Intelligence

**Nell'universo della conoscenza
con intelligenza linguistica**

Current Problems



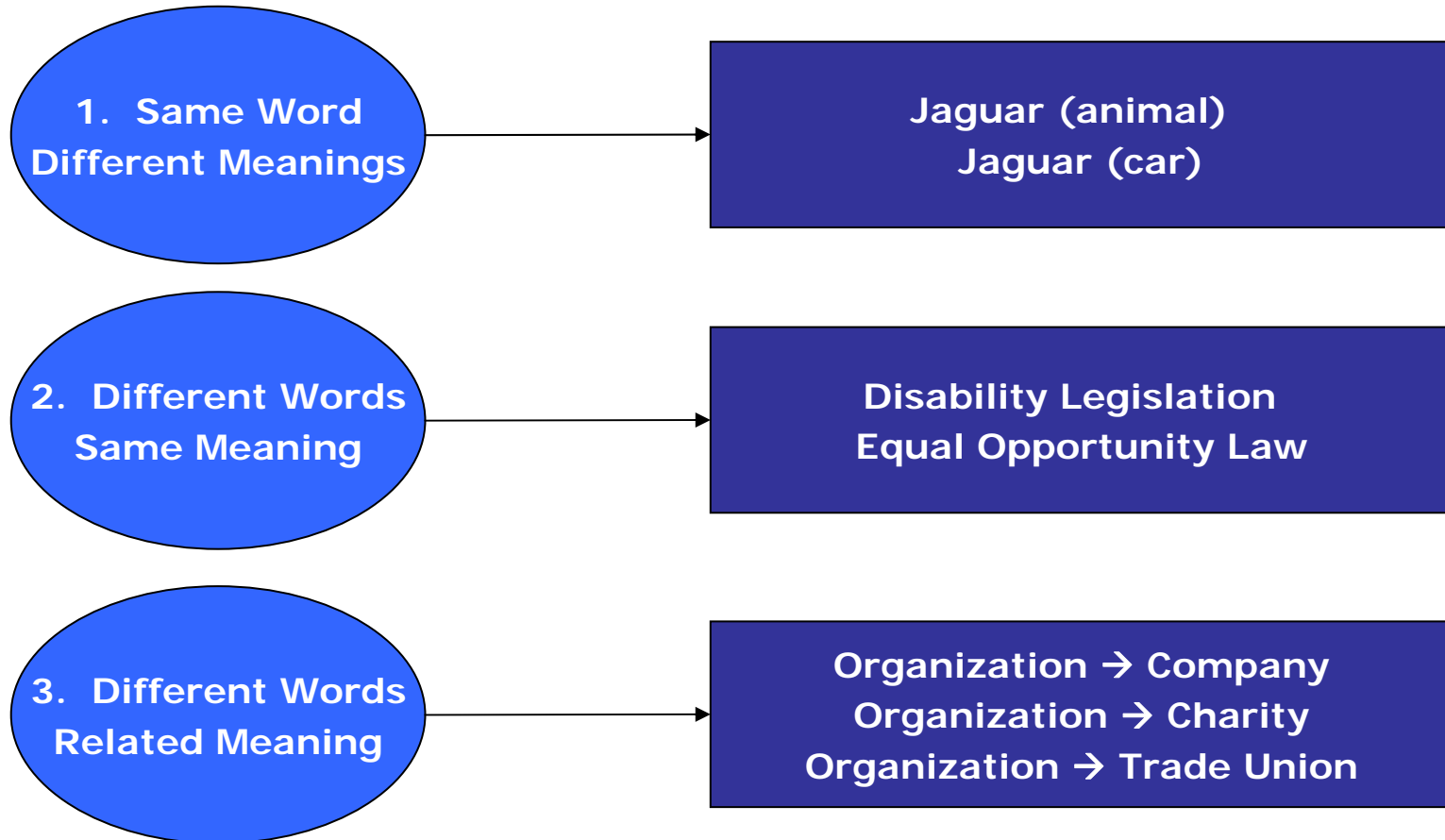
Current Problems

- Information Overload
 - Too many answers
 - Inability to classify or arrange effectively
 - Lack of time to read all documents
- Information Underload
 - Few or no answers
 - No classification or poor selection power of system



Current Problems

- 3 Problems with Search Technology;



Current technologies

Two technologies are currently used to manage unstructured information:

- **keyword (including statistics)**
- **shallow linguistics**

Other technologies (like *latent semantic* or *dynamic reasoning*) are simple variations of the two integrated with statistical elements or other “tricks”



- no understanding of the real content of the documents; 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:

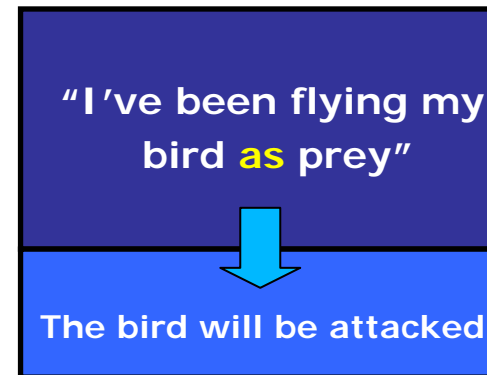
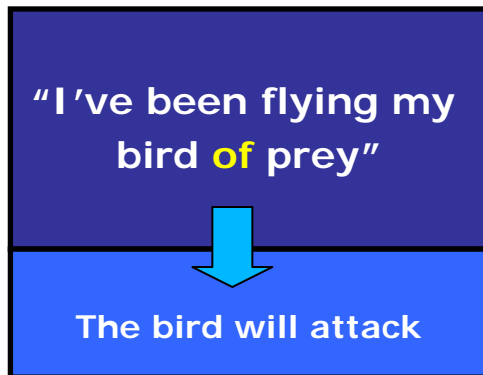
Internet users A the new credible results enable capability Web to information through technology called on advanced will in a new access way -- providing them with Utopia most search the search possible existing any whole

- **are the same!**



Shallow Linguistic

- Words are recognized as belonging or not to a dictionary and inflected forms are attributed to root words; all the elements are given a grammatical type (noun, verb, adjective, adverbs...)
- Lack of meaning based processing
- The text becomes a set of tagged words with the elimination of some stop words



Current Problems

- Traditional systems have a superficial processing level no understanding of the relation between elements and of the meaning, for example, in 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!



What is a Semantic technology: deep analysis

- Semantic technologies recognize the inner structure of the text and the relations between the elements, and can manage the content with much higher precision and relevance
- for example, in a sentence like this:

The cake that John gave to me was very tasty.

the system understands that it is the cake that is tasty (and not John even if it is nearer in the text), that John gave it to me and that the main clause is "The cake was very tasty"



Semantic understanding

Working on the semantic part of the content, Cogito can easily understand that this sentence:

A fast car that eats gas.

has many words in common with this sentence:

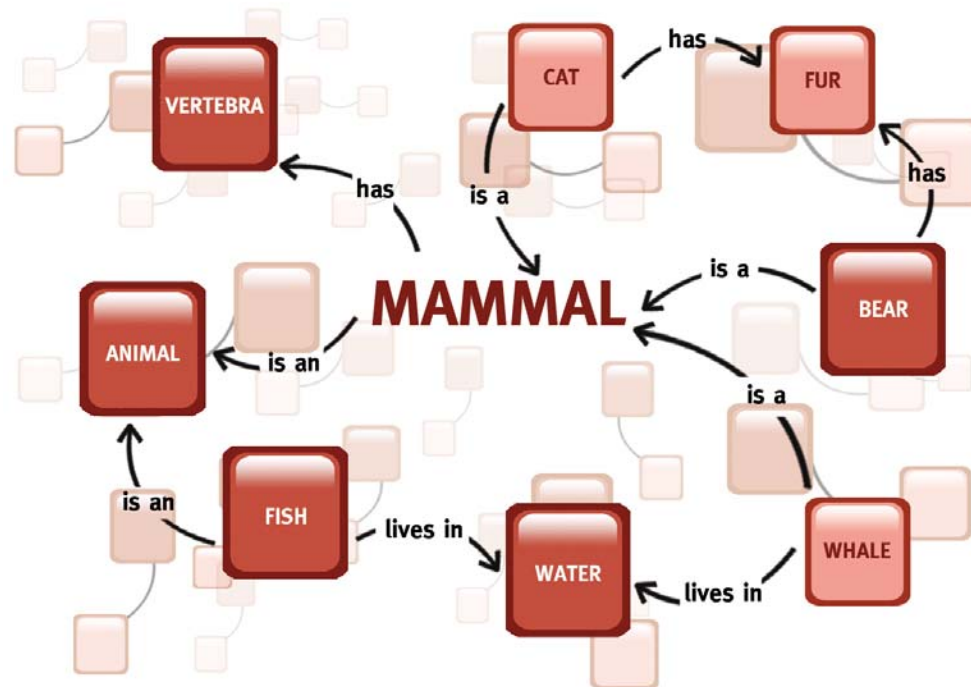
A car goes fast when we step on the gas.

but the meaning of gas (and of the whole sentence) is completely different. In the first one, we are talking of *gasoline* while in the second of the *accelerator*



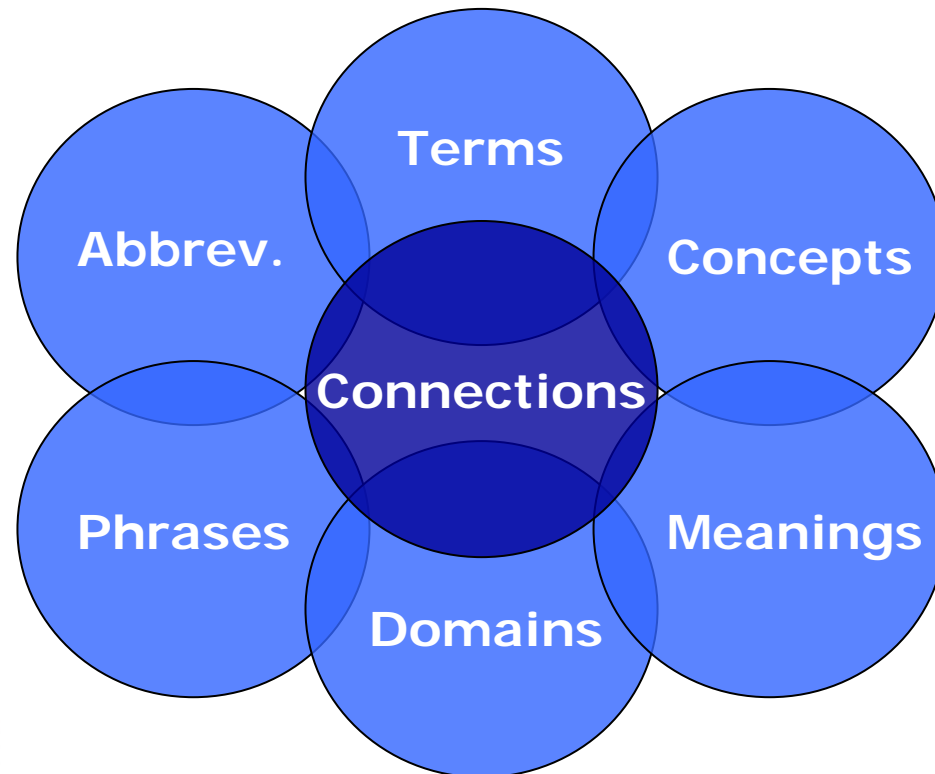
What is a Semantic Network?

- The **heart of semantic technology**;
 - Quality of results derived from the complexity and richness of the network.
 - Includes all definitions of all words.
 - Include relationships among all words.



Semantic Networks

- Traditional technologies can only “guess” the meaning using;
 - keywords, shallow linguistics, & statistics
- Semantic Networks instead indentify;



“San Jose is an American city”

“San Jose is a geographic part of California”

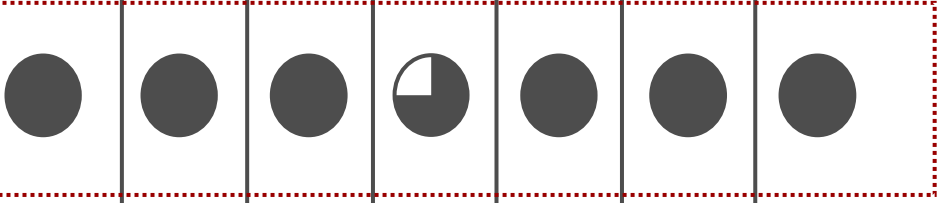


Next Generation Tech.

Disambiguation
Entity extraction
Categorization
Natural lang. I/F
Searching
Discovery
Sentiment extraction


Semantic Intelligence

- Linguistic rules
- Sentence analysis



Shallow text analytics

- Statistics
- Heuristic rules
- Morphological recognition



Keyword-based technologies



Applications & Benefits

1. [Conceptual Search & Analysis](#)
2. [IPTC Categorization](#)
3. [Sentiment Monitor](#)
4. [Question – Answer System](#)
5. [AskWiki](#)





Customer ENI

The Eni Group, one of the major companies operating in the areas of oil and natural gas.

Objectives

Maximize value of intellectual capital
Maximum interoperability (document management systems, PLM, shared drives) and security

Increase effectiveness of strategic marketing by leveraging real time information deriving from external sources and monitoring of competitors' activities

Solution

COGITO® Semantic Search

COGITO® Categorizer

Benefits

Effective searching, question answering and information extraction for research and development team from different sources
Complete organization of the company's information assets
Increased knowledge-sharing and cooperation among all users
Tangible support for strategic business processes like R&D, Marketing, etc.





Customer RCS

Most important publishing company in Italy

Objectives

Reduce Operating Costs for managing classification of incoming newsfeeds

Improve quality and consistency news classification

Solution

COGITO® Categorizer

Benefits

Automatic categorization of all incoming news based on European standard

Increased consistency in categorization process

Reduction in operating costs

Metrics

Before COGITO® Categorizer

15 people manually doing categorization

84% quality in categorization

After COGITO® Categorizer

2 people manually overseeing automatic categorization

Yearly savings of more than 350K€

85% quality in categorization



Customer Ansa

Biggest News Agency in Italy

Objectives

Reduce Operating Costs for managing classification of incoming newsfeeds

Improve quality and consistency news classification

Solution

COGITO[®] Categorizer

Benefits

Automatic categorization of all incoming news based on European standard (IPTC)

Increased consistency in categorization process
Reduction in operating costs

Customer

Telecom Italia (ISP Division)

Objectives

- Reduce Response Time of Customers' inquiries via email
- Reduce Operating Costs of managing Customers emails
- Improve quality and consistency of Customer Care answers
- Increase revenue by redirecting customer to pay per use phone service in case further assistance is required

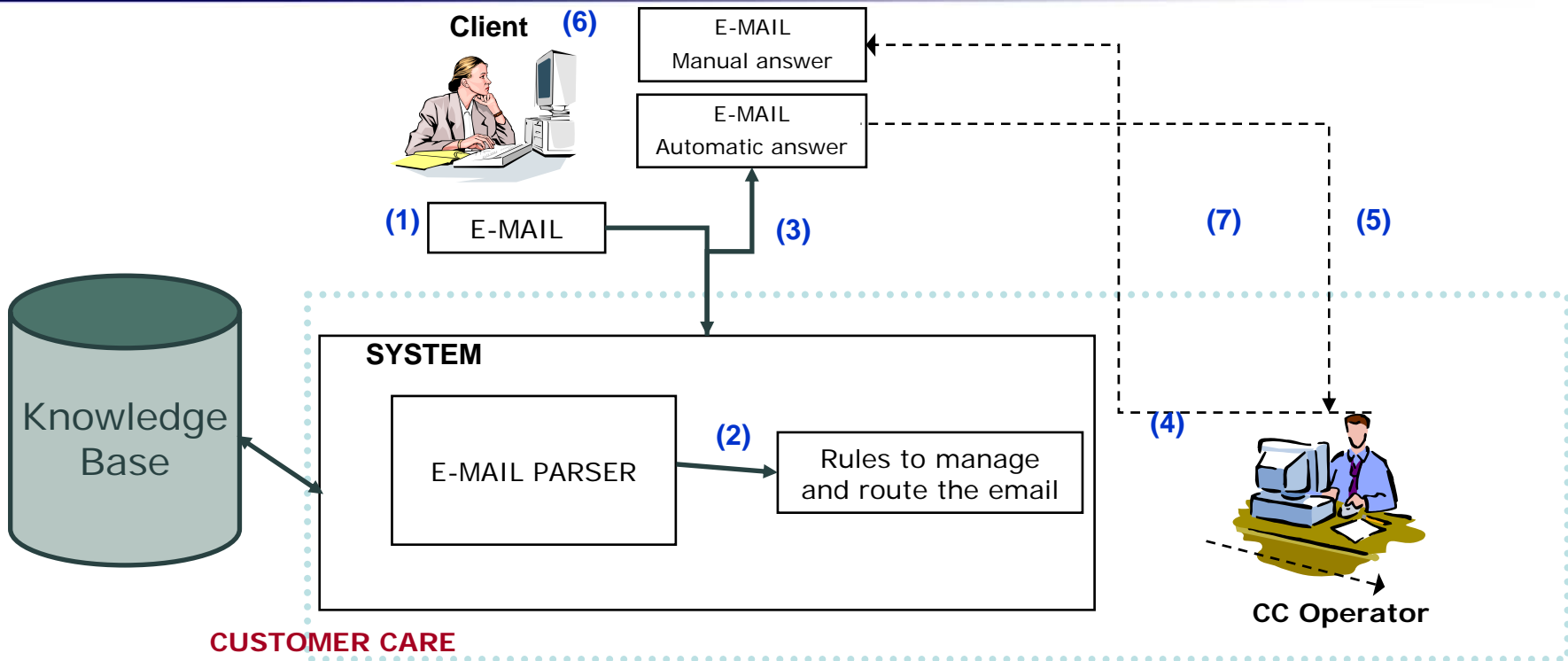
Solution

COGITO® Contact Mail

Benefits

- Automatic management (answer, routing) of customer emails
- Streamlined interaction with customers
- Regular and accurate statistics on customer emails (proactive problem management, input for marketing activities etc.)
- Automating refinement through voting
- Fast implementation (4 weeks)
- User friendly programming interface

Autom@il: How it works



- (1) Customer sends email (structured or unstructured)
- (2) The system processes the email interpreting the natural language
- (3) Automatic reply (+ link to the sections of the site when the topic is described)
- (4) If the system doesn't understand the question, it sends a standard reply acknowledging the reception and routes the email to the contact center operator who
- (5) Will answer manually
- (6) The customer reads the reply and if needed clicks on the links
- (7) If the automatic answer doesn't completely address the need of the customer, he/she can click on the link to inform that the answer is not sufficient and routes the message to the operator

Before Autom@il

Average Response Time = 30 Minutes

100% Answers Through Operators

0% of calls routed to pay per use phone service (unacceptable for the customer not to resolve issue working directly with Customer Care Engineer)

After Autom@il

Average Response Time = 60 Seconds

In 4 weeks, 71% email answered automatically

In 8 weeks, 85% of email answered automatically

15% growth in pay per use phone support

Total Operating Costs for email management reduced by 94%

72% of customers say that Autom@il provides the same or better quality of service

