HUMAN LANGUAGE TECHNOLOGIES FOR THE RETRIEVAL, CLASSIFICATION, EXTRACTION, ANALYSIS AND GENERATION OF DOCUMENTARY INFORMATION AND SOCIAL WEB

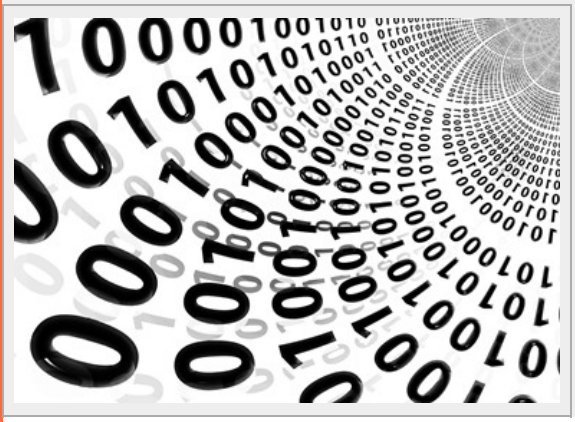
DADES DE CONTACTE:

Research Results Transfer Office-OTRI
University of Alicante
Tel.: +34 96 590 99 59
Email: areaempresas@ua.es
http://innoua.ua.es

RESUM

The University of Alicante has developed different techniques for the conversion of the documentary information into precise knowledge. The key lies on detecting the relevant aspects of the data dismissing its useless parts, for then analysing, classifying and discovering the hidden knowledge in the documentation and thus generate structured and specialized databases. This will feed new vigilance or decision-making processes but also generates new documents with specific purposes.

Sentiment Analysis and Opinion Mining are also employed in this process due to the exponential growth of the Social Web and the Web 2.0 data. By means of the abovementioned technology, researchers have developed four products that could be customized to specific requirements of any company or institution.



DESCRIPCIÓ TÈCNICA

Nowadays the access to information plays an essential role for any type of organization, business but also for people's daily life. Every day more often, the information generated has to be processed to be converted into knowledge, useful to take decisions in the future.

What are our clients asking for? Which are the new market trends? What are doing our competitors? Which are the opinions of

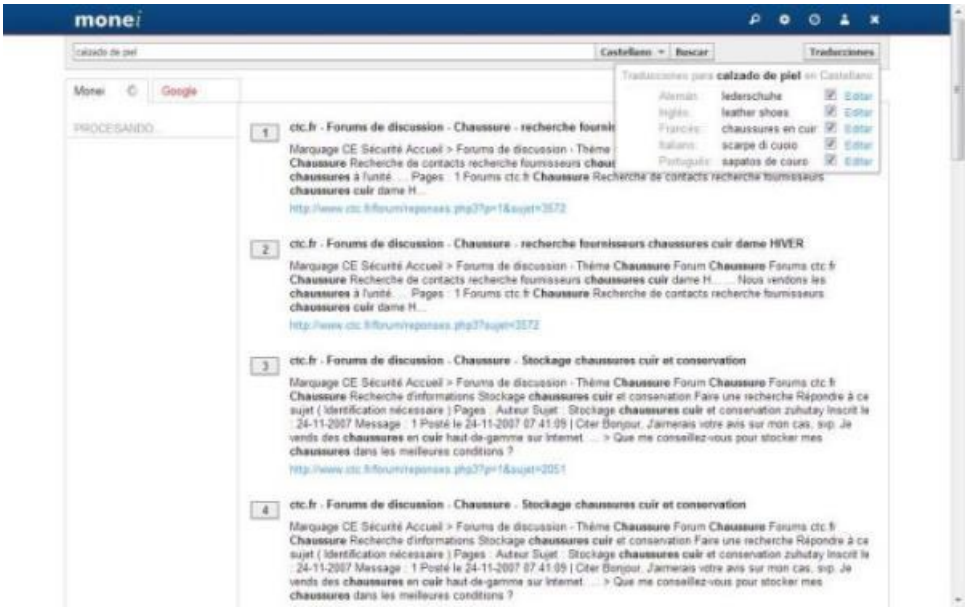
previous visitors of hotels, places, products that I am going to book? Who and how have solved this problem that I also have? Who is offering this product I need? These are the questions that feed the actual business intelligence systems that will support us in the decision making process. However, in order for these systems to be really effective, precise answers are needed and thus we need specialised information in terms of domain, classified according to the business type needs, but also automatized to guarantee a continuous and updated information flow.

Most of the informative needs have already been developed and are now available and waiting to be found. Documentary records, web pages, online databases, social networks, blogs, online press are updated daily and they provide this information we need. In addition, thanks to the new communication channels and to the digitalization this task should be easy to achieve. However, the huge amount of information accessible by search engines makes the process of specialized data access still a challenge.

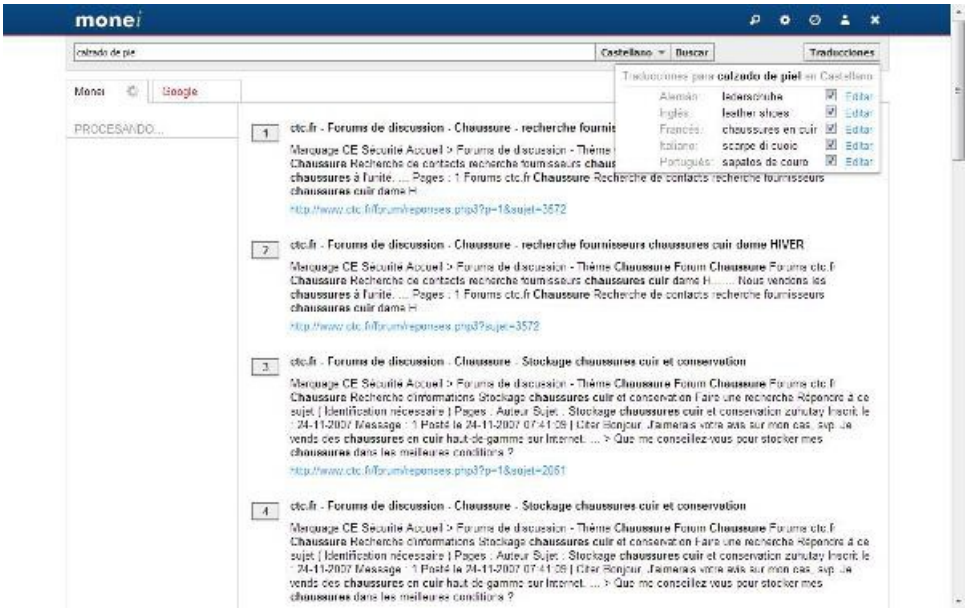
The Research Group of the University of Alicante has developed different techniques for the conversion of the documentary information into precise knowledge. The key lies on detecting the relevant aspects of the data dismissing its useless parts (in webs as well as in digitalized texts databases) for then analysing, classifying and discovering the hidden knowledge in the documentation and thus generate structured and specialized databases. This will feed new vigilance or decision-making processes but also generates new documents with specific purposes. Sentiment Analysis and Opinion Mining are also employed in this process due to the exponential growth of the Social Web and the Web 2.0 data.

By means of the abovementioned technologies GPLSI has developed the following products:

GPLSI MONEI is a meta search and intelligent engine that is adaptable to the needs of the organization. Monei allows activating multilingual simultaneous searches, restricting the search field to our domains, access hidden pages and databases, classifying, grouping information before its visualization and defining a ranking of results depending on its own business needs.



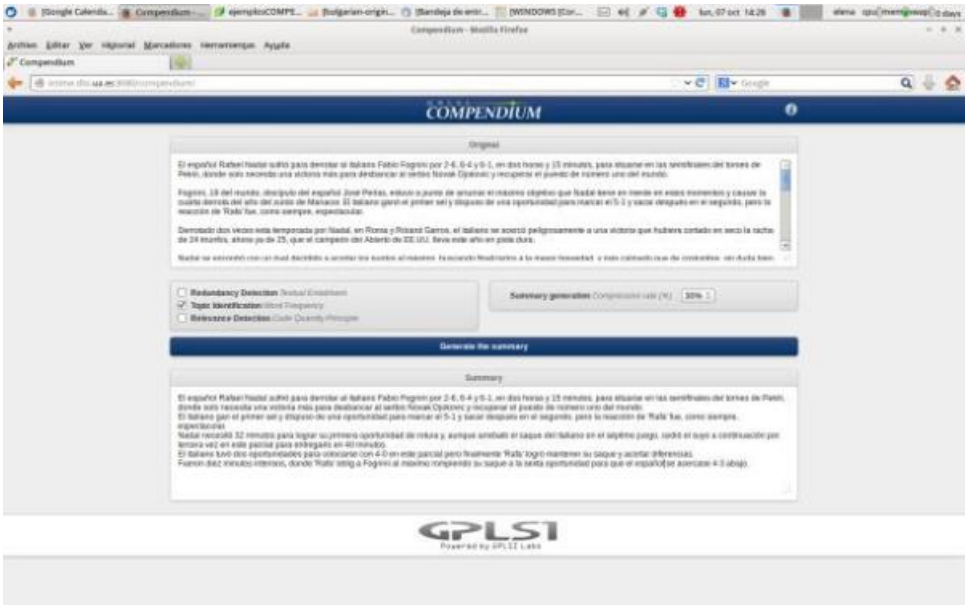
GPLSI DOSSIER is an application that allows the automatic classification of information coming from the press webpages, official bulletins, but also social networks. This is possible by means of extracting the information that is relevant to our business. It stores the information in structured databases for its future processing using business intelligent systems.



GPLSI SOCIAL OBSERVER is an application that retrieves tweets of the social network Twitter on a concrete topic, a brand, product, and person and automatically evaluates the expressed opinions by users. After that it carries out a behavioural and trend analysis by managing comparative analysis, time frames or geographical areas.



GPLSI COMPENDIUM is an automatic summarization system able to generate both generic as well as informative summaries. Compendium allows the inclusion of advanced natural language generation techniques to enrich the summary with additional information such as multimedia information, terms replacement with their synonyms or the replacement of formal with informal register or vice versa.



AVANTATGES I ASPECTES INNOVADORS

- Adaptation of developed applications to any sector or area of activity.
- Automatisation of search need and requirements for business information analysis.
- Possibility to use the different applications separately or have them integrated in a overall application.
- Built up of tools for decision-making and technological surveillance.

ESTAT ACTUAL

Products are already developed and could be adapted to the needs of companies in short time.

APLICACIONS DE L'OFERTA

Applications can be employed in any sector, which deals with big amounts of data that need to be processed and where there is the need for information extraction and data upon request.

Among others:

- Management of administrative data, databases and other kind of text: retrieval, follow up and classification

- Monitoring and follow up of Web published information
- Public administration
- Journalism and news management
- Marketing and drafting of socio economic studies
- Business development, Business intelligence

COL·LABORACIÓ BUSCADA

GPLSI is looking for companies or entities interested in its know how to develop concrete applications the client needs depending on his requirements and context. We also look for partners to develop other applications and or research project jointly.

The proposed technological agreements are:

- License agreements: Know how license
- Commercial agreements with technical assistance
- Joint Research and Development projects

DRETS DE PROPIETAT INTEL·LECTUAL

Products are protecte with IPR and author rights.

SECTORS D'APLICACIÓ (3)

Pharmacology, Cosmetics and Ophthalmology
Computer Science, Language and Communication
Medicine and Health