

Linguamatics I2E: Solution Overview

Turning Information To Competitive Advantage

Linguamatics helps you to find business-critical facts and relationships hidden in unstructured text to effectively support information-driven decision-making. Using I2E, organizations can do much more than simply cope with overwhelming quantities of information. Our unique combination of literature search and mining in context helps you turn this information into a competitive advantage.

For example, in drug discovery and development, effective use of existing and newly published research is critical for optimizing return on R&D investment. It has been suggested that for half of drug projects that undergo attrition due to safety concerns, information existed in the literature that could have enabled prediction of failure earlier, thus saving valuable time and resources.

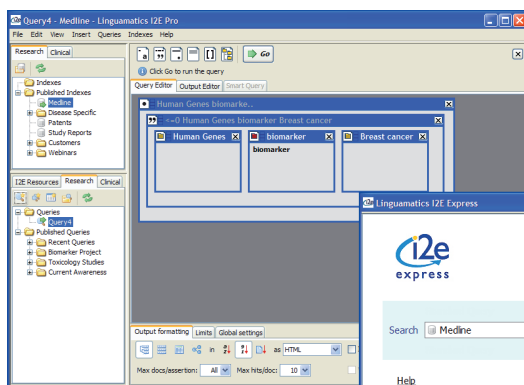
Similarly, in the business intelligence domain, key opinion leader identification and sentiment analysis can be critical in marketing strategy development for brand and product life cycle management.

Around 80% of information of interest to a typical organization is unstructured and written in natural language. Mining such information is an increasing challenge. For example, with around 17 million abstracts in the MEDLINE® biomedical database and over 1.5 million patents registered in the US alone, how can organizations mine such resources for critical information and find hidden connections?

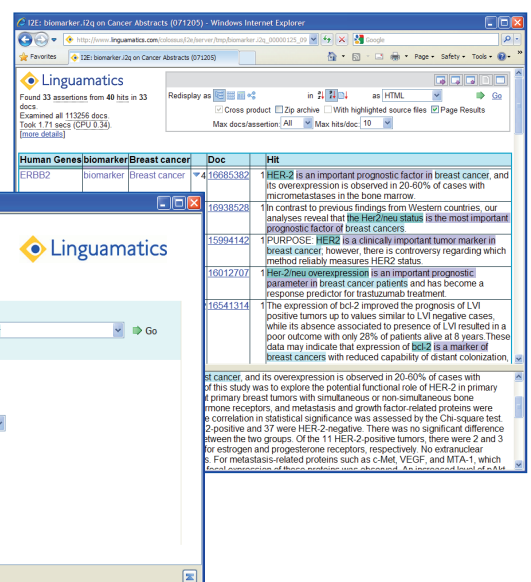
The answer is I2E, the agile interactive information extraction software from Linguamatics that lets you answer questions and discover knowledge that would be intractable or impossible to find using conventional methods. I2E's powerful natural language processing (NLP) capabilities go way beyond keyword searching to mine large and often complex collections of documents, extracting relevant facts, relationships and entities from unstructured and semi-structured content, and delivering results in the right format to support your organization's decision-making needs. Users can select the appropriate literature sources for querying, including scientific journal articles, internal project reports, news feeds, patent documents, or emails.

Linguamatics enables rapid discovery of high value insights from text through:

- ◆ On-demand querying
- ◆ Semantic information extraction
- ◆ Systematic and reproducible analysis
- ◆ Powerful and rapidly-deployable text search and mining



Linguamatics I2E is a flexible, scalable knowledge discovery platform that delivers high quality and structured results, fast



Why Choose I2E?

Developed by a team with specific expertise in information extraction for information-rich and context-sensitive environments, from life science research to business intelligence, I2E has particular strengths for answering a wide range of questions, whether open queries like “what treatments are being used for breast cancer?” or precisely defined queries like “what positive comments have been made by a group of opinion leaders about product X?” By providing occasional and expert user interfaces, I2E can address the needs of users across the enterprise, leveraging expert knowledge and providing accurate and timely results to support effective decision-making.

Searching in Context

I2E makes it easy to plug-in domain-specific knowledge using thesauri, taxonomies or ontologies, which describe the concepts in a given field and their relationships. I2E is used with commercial, public domain and in-house knowledge bases. The system is equally able to answer questions that are natural to a life scientist who may be interested in biological targets, biomarkers, pathways and potential new therapeutics, as for the business analyst who wants to know the key trends and influencers in a particular market.

Versatility

I2E is not a point solution but offers a spectrum of capabilities that can be applied flexibly to address text search and mining challenges. Strategies range from simple document retrieval using keyword search to advanced information extraction approaches, including retrieval of facts, relationships and entities using linguistic structure. Thus users match the right approach to their specific task and combine complementary techniques to get the answers they need.

Getting Directly to the Answers

I2E gets straight to high value, relevant, re-usable knowledge. Results are highly structured, enabling rapid analysis and export directly to spreadsheets like Microsoft Excel or network visualizers like Cytoscape. Researchers and information professionals get the flexibility and analytical rigour that they need. Tabular reports summarize the extracted information in a compact, clustered format with easy drill-down to supporting evidence, including direct links to source documents.

For more information or a demonstration, please call us today

☎ +44 1223 421360

✉ info@linguamatics.com

User Focus

I2E is highly user-centric. The I2E Express interface is easy to use and similar in appearance to a conventional search engine. For more advanced use, I2E Pro lets the user view, construct, and manage sophisticated queries using an intuitive drag and drop interface. Optimized queries can then be saved and published for use by less experienced users. So validated queries can be applied in a simple and repeatable way, for example to run regular queries on a growing database of literature.

Quality and Accuracy

I2E's powerful linguistic queries enable users to extract high value information. Searching with entities and concepts (e.g. gene, protein) is highly flexible. Linguistic wildcard features allow users to ask open questions and search for entities, verbs and unknown relationships. Searching using positional constraints can identify stronger relationships by finding terms that often appear together, for example within a proximity of n words, or in the same sentence, or in a particular region of a document. Unlike other systems, I2E is not limited to searching inside a single sentence, but can also filter based on the broader context of the document. More distant and indirect relationship can be detected by combining result sets or visualizing information networks to find relationships across multiple documents.

For more information or a demonstration, please call us today

+44 1223 421360

info@linguamatics.com

Performance and Robustness

I2E is scalable to millions of documents, such as the whole of MEDLINE, and works with large documents such as patents. In addition I2E is able to combine large numbers of search terms. For example, classes of genes, proteins, diseases or compounds that may contain up to tens of thousands of terms, can be encapsulated within a single "class" search term. Search performance is fast and customers report that discovering relevant information using I2E has been shown to be at least ten times faster than conventional keyword search-based approaches.

Control and Interactivity

The flexibility, transparency, speed, and structured results delivered by I2E mean that users can refine queries on-the-fly. I2E makes NLP based text search and mining interactive and intuitive, putting power into the hands of the user themselves.

Integration

I2E can play a key role embedding information extraction into more complex workflows. This can be achieved by integrating with analytic workflow software or by exporting results in data formats compatible with standard enterprise software systems.