KeyQ-solr

KeyQ is what you were searching for. Now let us search for you.





Contact information

Address: ETSI Informáticos (antigua Facultad de Informática) Ontology Engineering Group (OEG) Departamento de Inteligencia Artificial Universidad Politécnica de Madrid Campus de Montegancedo, sn Boadilla del Monte, 28660 Madrid Phone number: 910672915 Website: demo.inno.oeg-upm.net Email: mariano.rico@upm.es

Technological Offers type

Technological solutions

Research and innovation areas

• Digital Technologies, Artificial Intelligence, Cybersecurity, 5G, Robotics



Available from: 2021

Where?

Artificial Intelligence R&I Centre (Al.nnovation Space) Ontology Engineering Group

Keywords: | Key Q | NLP | PLN | recover of information | seeker | terminology

Smart searching across millions of documents

Linear scalability and consistent response times

Our tests on corpora of millions of documents show a consistent response time. It can be used by means of a web application or a REST API.

As a web application, it offers a sophisticated but intuitive means of writing search terms.

Demo here.

Description of the technological basis

For storage and recovery, it uses the open-source Apache solr technology standard and KeyQ, our terminology extractor.

The web interface can be adapted to any kind of device (mobile phone, PC, tablet, etc.), using Bootstrap.

Business needs / application

Companies with large volumes of documents, in all kinds of formats (PDF, Word, etc.). Searching for information in their document repositories is a need shared by all kinds of companies.

Competitive advantages

Compared to the basic version, KeyQ (terminology extraction and online command searching, with the corpus stored in the memory), KeyQ-solr distributes the corpus documents to the hard disks of one or several machines using Apache solr technology. Its other features include:

- Filters: partial searches, upper/lower-case sensitivity, file metadata.
- Granularity selector for the information retrieved: page, paragraph or sentence.
- Detailed statistics of the corpus texts.
- Advanced visual analysis to identify thematic groupings.
- Handling of multiple corpora in different languages.
- Terminology management. Collaborative terminology evaluator (consensus among experts).
- Searches with logical operators.

Past performance references

We have handled corpora in fields such as:

- Biology: tens of thousands of scientific articles on Covid-19.
- Law: large documents with hundreds of pages, with instant access to the most relevant page/paragraph/sentence for a particular search.
- Energy and the environment: thousands of European Parliament documents in different formats.

KeyQ was developed in the Artificial Intelligence R&I Centre (Al.nnovation Space) (joint UPM-Accenture centre) between 2020 and 2021.

Protection

• Software registration

Stage of development



KeyQ-solr is what you were searching for. Now let us search for you.