

Report

Minsait's technology reduces the average time spent classifying legal documentation by 95%

Minsait's product, Onesait Language, leads a European innovation project that takes Natural Language Processing towards the automated processing of large quantities of documentation

Nowadays, in companies and organizations, there are few processes as tedious as the classification of documents and the extraction of their content. Transforming unstructured information into structured information has become one of the great challenges to make use of the innumerable synergies and opportunities that repositories and archives offer us and, in this way, generate a significant impact on the business.

Minsait, an Indra company, has developed a product that reduces the average time used for the classification of documents by 95% and the average time for the automated extraction of information by up to 90%. Its Onesait Language solution allows unstructured documents written in natural language to be classified automatically and extract the relevant information, through the use of Natural Language Processing (NLP) technologies and neural networks. A leap in quality for the intensive business processes regarding mass document treatment, which will improve productivity and efficiency.

"NLP technology allows our solution to automatically understand written human language. In addition, we combine the most *traditional* Natural Language Processing with the latest techniques based on deep learning using neural networks, which gives the product the capability to constantly learn from its own mistakes," explains Hitoshi Yano, expert in Digital Solutions Development in Minsait. As he explains, the solution "works with an Intelligent Workspace and allows users to manage their workload, review classified documents and monitor the data extracted automatically."

In fact, among its main attributes, Onesait Language makes it possible to extract information in several languages, and has a scalable, flexible architecture that is perfectly integrated with other client systems. From a technical point of view, the system has been trained with legal and real estate documents, and applies an operating model based on SaaS with flexible billing. It also offers an accessible REST API, for the creation of web services, and can be integrated by modules with any of the preexisting systems.

The solution applies to all sectors where massive amounts of unstructured documents written in natural language are managed, that is, physical or digital information that is not of the same type, such as printed documents, videos, photographs, emails, social networks, etc., data that until now was a challenge, and that, with this solution, not only are selected but can add value and constitute a starting point for the execution of new operations in an increasingly digital world.

Leading innovation in Europe

The value of Minsait's solution will now be measured in a European project that has been financed by the European Institute of Innovation and Technology (EIT), an organization of the European Union that, in turn, receives the support of the EU's Horizon 2020 research and innovation program. The initiative, which will bring Natural Language Processing to the management of legal documents and has the collaboration of knowledge societies and other companies in the EU as potential users, aims to make progress in the research into this type of technology.

Specifically, Onesait Language will handle the classifying of various tenders automatically, extracting the most relevant details (bidders, amounts, nature, geographical location, term, risks, etc.) in order to streamline

Report

decision-making in the business and redirect each process to the right people, thus managing a greater number of tenders and reducing their cost. This is a preliminary step for the development of an automated business pipeline, which will greatly facilitate operations and data analysis in this sector, providing greater clarity and agility for decision-making.