

## Expression of Interest



### Contact Person/Scientist in Charge

- **Name and surname:** Elena Lloret Pastor
- **Email:** [elloret@dlsi.ua.es](mailto:elloret@dlsi.ua.es)

### Universidad de Alicante

### Department / Institute / Centre

- **Name:** Natural Language Processing and Information System Research Group (GPLSI), Department of Software and Computing Systems
- **Address:** Universidad de Alicante, Carretera San Vicente del Raspeig, s/n
- **Province:** Alicante

### Research Area

- Information Science and Engineering (ENG)

### Brief description of the institution:

The University of Alicante (UA) was created in 1979. Today it educates and trains more than 25.000 students -3.000 of them are international students - and offers more than 43 undergraduate and 80 postgraduate programmes: consequently it is proportionally one of the fastest growing universities in Spain. The UA houses 227 research groups in Social and Legal Sciences, Experimental Sciences, Technological Sciences, Human Sciences, Education and Health Sciences and 15 Research Institutes (Water & Environment, Materials, Electrochemistry, Biodiversity, Chemical Processes and Organic Synthesis, and Modern Languages, among others). Thus, the UA employs over 3.500 researchers/ professors and has a complex management /administration structure of 1.300 people, which involves an annual budget of 200 million Euros.

UA is a young and dynamic university with vast experience in implementing EU funded projects in different programmes and areas, with presence in more than 60 countries worldwide. In the last 10 years, UA has successfully acted as coordinator of many Tempus, Alfa, Edulink projects involving Third Countries and

Lifelong Learning and Framework Programme (FP, DG Research) Projects. Moreover, the participation in FP has been increasing in the last years, taking part in 25 5th FP, 26 6th FP, 45 7th FP projects (13 of them coordinated by UA), and 11 in H2020.

It is worth underline the big effort performed by UA in order to meet the commitment with the principles set out in the European Charter for Researchers and in the Code of Conduct for the Recruitment of Researcher

**Brief description of the Centre/Research Group (including URL if applicable):**

The Natural Language Processing and Information System Research Group (GPLSI) belongs to the Department of Software and Computing Systems at the University of Alicante (Spain). Created more than 20 years ago, it has consolidated its research experience in Human Language Technologies, Natural Language Processing and Text Mining, developing different techniques for the analysis and conversion of the documentary information into knowledge.

The main research lines are:

- Information retrieval
- Information extraction
- Text classification
- Semantic Analysis
- Text Summarization
- Sentiment analysis
- Text Normalization
- Social media monitoring and analysis, including the detection and classification of the informality appearing on the Web 2.0
- Resources and corpus development
- Natural Language Generation
- Human Language Technologies for Social Purposes

More information about the GPLSI research group and department can be found at:

<https://gplsi.dlsi.ua.es/gplsi13/en>

**Project description:**

The over-abundance of information and its heterogeneity requires new ways to access, process and generate knowledge according to the user needs. In this regard, Human Language Technologies (HLT) play a key role in information analysis.

Despite the great advances in artificial intelligence, there are yet no systems capable of directly generating natural language texts that target diverse topics and communication purposes within the same approach, e.g., to report on the latest fashion trends, or to recommend/criticise a product. Such a system will significantly enhance human-computer communication and interaction. Besides contributing to the development and improvement of "robot writers" and "storytellers", myriad writing tasks will be supported, including reporting, information/feedback generation and storytelling in a number of sectors, e.g., education, marketing, business or medicine.

Current natural language generation approaches only provide partial solutions to this challenge, where the communicative goal is already pre-established or assumed in the process, and their inputs are mainly limited to either data or text. Consequently, ad-hoc systems are created, failing to dynamically adapt to several contexts, as well as to exploit complex multimodal and heterogeneous information. We aim to investigate and obtain a single NLG method to automatically produce different text types that are guided by a communicative goal. To do this, deep and hierarchical learning architectures will lay the groundwork for the automatic development of new semantically-meaningful written language models. Reliable knowledge will be derived from raw information available on the Internet using up-to-date advanced natural language understanding and machine learning techniques. In addition, specific characteristics of a communicative goal to target the text to a desired purpose will be determined. These key elements - the extracted knowledge and the characteristics of a communicative goal - will then be integrated and considered as major features of the learning process for the entire generation pipeline, being this project pioneer in this respect.

To meet the aforementioned overall objective of this research project, the following specific objectives have been proposed:

(OBJ1) Analyze and define what a communicative objective is and its involvement in the production of language.

(OBJ2) Characterize communicative objectives and analyze their relationship with textual genres.

(OBJ3) Research, propose and develop novel approaches to generate communicative language models using NLP techniques and advanced automatic learning algorithms, such as deep learning.

(OBJ4) Research, propose and develop a holistic, flexible and dynamically adaptive NLG method that uses the knowledge obtained from communicative language models.

(OBJ5) Evaluate intrinsically and extrinsically each of the proposed techniques and approaches, as well as in a real scenario to demonstrate the validity and application of the NLG method.

(OBJ6) Promote and disseminate the research results obtained from the project through different national and international media – including well indexed journals, conferences, seminars, etc.-, as well as exploit the potential for transferring this technology to society.

## Applications

- Curriculum Vitae.
- A personal statement/application letter.
- At least 1 reference letter.

Deadline - June 30, 2019