Human-Robot Interaction patent prediction through Long Short-Term Memory network

Jon Borregan-Alvarado (jborregan001@ikasle.ehu.eus), University of the Basque Country UPV/EHU
Izaskun Alvarez-Meaza (izaskun.alvarez@ehu.eus), University of the Basque Country UPV/EHU
Ernesto Cilleruelo-Carrasco (ernesto.cilleruelo@ehu.eus) University of the Basque Country UPV/EHU
Gaizka Garechana-Anacabe (gaizka.garechana@ehu.es) University of the Basque Country UPV/EHU

1 Introduction

The fourth industrial revolution was born in Germany in 2011 with the term Industry 4.0, entailing the connection between humans and things both in the physical and virtual world, thanks to different technologies such as artificial intelligence, virtual reality, robotics, the Internet of Things, among others [1].

Currently, the already existing Industry 5.0 takes as a reference a human-centered environment, highlighting the interaction and collaboration between humans, robots and intelligent machines, leading to an increase in productivity by eliminating non-value-added tasks, increasing quality and safety, and reducing costs and waste, among others [2].

Therefore, the main objective of this research work is to identify and predict the concepts around the term HRI and thus generate a research agenda through a multilayer analysis model, taking patents as a source, since they contain more than 90% of the recent technological information that exists in the world [3]. The initial layer identifies the most HRI related terms in the abstract and claims fields, a key part of the patent application [4], using Natural Language Processing (NLP) techniques. In the second and last layer, a predictive approach using techniques based on Recurrent Neural Networks (RNN) will identify the main issues in the short term, and the prediction is consolidated with the identification of the most related terms.

2 Methodology

The methodology carried out has been divided into the following phases.

In the first phase, a search query based on the term HRI has been developed through the PatSeer patent database.

In the second phase, patent abstracts and claims are refined and analyzed separately. Then, through the application of Natural Language Processing (NLP) using Python, a "Word2Vec" model is obtained referring to the set of feature vectors that represent the words of the corpus (word embeddings), quantitatively obtaining vector distances.

Finally, using a Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM) architecture in Python, a model trained with the corpus has been created, capable of learning the style of the text and hopefully generating new abstracts and claims. The different words that compose the new text, together with the frequency in which they appear, are considered as potential terms to be highlighted in the short term within the HRI subject. These potential terms are also quantitatively observed through the Word2Vec (NLP) model previously executed in Python, in order to observe the vector distance that these potential terms share with others and, therefore, their affinity. In this way, we managed to give greater robustness to the terms observed in the previous phases, and define new research paths.

3 Results

3.1 Layer 1: Identifying topics closely related to HRI

To identify the terms with the highest affinity related to HRI in reference to the abstracts, by applying Word2Vec (NLP) models generated in Python, the vector distances of several terms were quantitatively obtained. The results describe general fields of action, such as robot capa-
2
bilities (intelligent or intelligence; emotional; capable), the characteristics to be taken into account in research and development work (disclosure; modeling; chart), and the result or specialization of this research (knowledge; field).
With respect to claims, the steps to follow are the same as those described above for abstracts. In this case it is observed that the results describe fields of action referring to casuistic or problematic (anomaly; alarm; indicator), and standard (rules; processes; classified).

3.2 Layer 2: Predicting potential short-term terminology regarding HRI

By using Python for training, learning and generating a “Generative” Long Short-Term Memory (LSTM) Recurrent Neural Network (RNN) thanks to the corpus of abstracts and claims previously obtained and cleaned, we have succeeded in automatically generating 6 new claims and an abstract with which to identify the potential terms to be highlighted and studied in the short term by the scientific community with respect to the concept or field of HRI.

It is observed that the words "connected, gesture, information, method, processor, sensor and user" are the terms with the greatest potential, focus and growth for HRI research to be developed in the short term. In addition, the three words vectorially closest to the potential terms have also been quantitatively calculated, highlighting that these are concepts related to methods or actions linked to the potential terms, as well as to characteristics or performance qualities of the potential terms.

4 Conclusions

The results of the study obtained allow the robust identification of future lines of research and applications in the area of Human-Robot Interaction. The terms Connected, Gesture, Information, Method, Processor, Sensor and User are considered as potential terms to appear in the short term in the environment corresponding to the HRI field, taking also into consideration as outstanding concepts or words those referring to the "spectrum" vectorially closer to the potential terms.

Also, it can be concluded that this is a very active and constantly evolving research field, so predicting future trends and needs allows to better focus the technological strategies to be implemented and developed by companies and research groups, and opens doors to new in-depth research related to the HRI concept and collaborative robotics, thus achieving a greater generation and transmission of knowledge for the scientific community.

References

1. Chung, H. Adoption and Development of the Fourth Industrial Revolution Technology: Features and Determinants. *Sustainability* **2021**, *13*, 871, doi:10.3390/su13020871.
2. Martin, L.; Gonzalez-Romo, M.; Sahnoun, M.; Bettayeb, B.; He, N.; Gao, J. Effect of Human-Robot Interaction on the Fleet Size of AIV Transporters in FMS. In Proceedings of the 2021 1st International Conference On Cyber Management And Engineering (CyMaEn); IEEE: Hammamet, Tunisia, May 26 2021; pp. 1–5.

3. Zha, X.; Chen, M. Study on Early Warning of Competitive Technical Intelligence Based on the Patent Map. *JCP* **2010**, *5*, 274–281, doi:10.4304/jcp.5.2.274-281.
4. Fujii, A.; Iwayama, M.; Kando, N. Introduction to the Special Issue on Patent Processing. *Information Processing and Management* **2007**, 5.