

'Research Insider': «Making intelligent systems understandable to humans»

Data: xoves, 2 marzo, 2017 - 12:00 - 13:00

Lugar: Assembly Hall

Poñente(s): José María Alonso (Postdoctoral Researcher at CITIUS)

Idioma: Inglés

Streaming: Por confirmar



Among the many relevant challenges that the *Big Data* paradigm has brought to the attention of data scientists and *Artificial Intelligence (AI)* researchers, one of them has been very recently stated by the *USA Defense Advanced Research Projects Agency (DARPA)*: “even though current *AI* systems offer many benefits in many applications, their effectiveness is limited by a lack of explanation ability when interacting with humans”. In this context, it is worthy to note that interpretability issues are deeply rooted in the core of the *Soft Computing* sub-field of *AI* since the first ideas published by Zadeh in 1965, towards the most recent work dealing with human-centric computing, computing with words and perceptions, etc.

This talk will be focused on how to develop *Interpretable Fuzzy Systems (IFS)*, i.e., systems easily understood, trusted on or accounted for human beings; systems which are highly appreciated in applications where the interaction with humans is the main concern, in fields like medicine, agriculture, marketing, robotics, etc. Then, we discuss the role that *IFS* can play with the aim of addressing the last *DARPA* challenge on explainable *AI*. Finally, preliminary results in an empirical study aimed at finding out how end users may understand easier the decision made by a fuzzy system when they were provided with a textual interpretation of the related fuzzy inferences will be presented.