UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

DEPARTEMENT D'INFORMATIQUE
DEPARTEMENT FÜR INFORMATIK

INFORMATICS COLLOQUIUM

Speaker:

Dr. Jose Maria Alonso Moral, Research Centre on Intelligent Technologies (CiTIUS), University of Santiago de Compostela, Spain

Paving the way from Explainable Fuzzy Systems to Trustworthy AI

Abstract:

In the era of the Internet of Things and Big Data, data scientists are aimed for finding out valuable knowledge from data. They first analyze, cure and pre-process data. Then, they apply Artificial Intelligence techniques to automatically extract knowledge from data and afterwards translating knowledge into products and services for economic growth and social benefit. Trustworthy AI is an endeavor to evolve AI methodologies and technology by focusing on generating a new generation of intelligent devices which respect non-discrimination and fairness with the ability to explain their decisions to humans. Furthermore, when it comes to consider the impact of AI systems on human behavior, several dimensions have to be addressed, being the most well-known among them the ethical, legal, social, economic, and cultural issues.

Explainable Fuzzy Systems born with the aim of paving the way from interpretable machine learning to Explainable AI. Such systems deal naturally with uncertainty and approximate reasoning (as humans do) through computing with words and perceptions. This way, they facilitate humans to scrutinize the underlying intelligent models and verify if automated decisions are made based on accepted rules and principles, so that decisions can be trusted and their impact justified. Notice that, EXFS automatically generate factual and counterfactual verbal and non-verbal explanations. Such explanations include linguistic pieces of information which embrace vague concepts representing naturally the uncertainty inherent to most activities in the everyday life.

In this colloquium, we will begin with a non-technical introduction to the field of TAI (i.e., revisiting definitions and fundamentals, reviewing the state of the art and enumerating open challenges). Then, we will briefly review the history of fuzzy systems from the pioneer works of L.A. Zadeh to the most recent developments on EXFS, with special emphasis on fuzzy-grounded knowledge representation and reasoning, as well as how to use EXFS to deal with imprecision and uncertainty in the contexts of XAI and TAI. Finally, we will see some practical software tools as well as critical issues related to psycholinguistic human evaluation.

Bio:

Jose M. Alonso-Moral received his degrees in Telecommunication Engineering, from the Technical University of Madrid, Spain. He is currently "Ramón y Cajal" researcher funded by the Spanish Government, affiliated to CiTIUS-USC, President of the Executive Board and Deputy Coordinator of the project "Interactive Natural Language Technology for Explainable Artificial Intelligence", Chair of the IEEE-CIS TF on Explainable Fuzzy Systems, member of the IEEE-CIS TF on Explainable Machine Learning, member of the IEEE-CIS WG on eXplainable AI, member of the IEEE-CIS TF on Fuzzy Systems Software, board member of the ACL SIG on Natural Language Generation. He has published more than 160 papers in international journals, book chapters and conferences and is co-author of the book "Explainable Fuzzy Systems".

Date and time: *Thursday, October 27nd, 2022, 11.00 pm*
Location: *Pérolles 21, room E230, Bd de Pérolles 90, Fribourg*
Contact person: *Prof. Edy Portmann*

The colloquium is free and open to the public.