SPECIAL ISSUE ON

EXPLAINABLE AND TRUSTWORTHY ARTIFICIAL INTELLIGENCE

Aims and Scope

Computational Intelligence (CI) encompasses the theory, design, application, and development of biologically and linguistically motivated computational paradigms emphasizing fuzzy systems, neural networks, connectionist systems, genetic algorithms, evolutionary programming, and hybrid intelligent systems in which these paradigms are contained. These techniques and their hybridizations work in a cooperative way, taking profit from the main advantages of each individual technique, in order to solve lots of complex real-world problems for which other techniques are not well suited. CI enables Artificial Intelligence (AI) through simulating natural intelligence in all its forms.

In the era of the Internet of Things and Big Data, data scientists are required to extract valuable knowledge from the given data. They first analyze, cure and pre-process data. Then, they apply AI techniques to automatically extract knowledge from data. Actually, AI is identified as a strategic technology and it is already part of our everyday life. The European Commission states that "EU must therefore ensure that AI is developed and applied in an appropriate framework which promotes innovation and respects the Union's values and fundamental rights as well as ethical principles such as accountability and transparency"[1]. It emphasizes the importance of Explainable AI (XAI in short), in order to develop an AI coherent with European values: "to further strengthen trust, people also need to understand how the technology works, hence the importance of research into the explainability of AI systems". Moreover, as remarked in the XAI challenge 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"[2].

[1] European Commission: Artificial Intelligence for Europe. Tech. rep., European Commission, Brussels, Belgium (2018), https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe, Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions (SWD(2018) 137 final)
[2] DARPA: Explainable Artificial Intelligence (XAI), https://www.darpa.mil/program/explainable-artificial-intelligence

Accordingly, users require a new generation of XAI systems. They are expected to naturally interact with humans, thus providing comprehensible explanations of decisions automatically made.

XAI is an endeavor to evolve AI methodologies and technology by focusing on the development of agents capable of both generating decisions that a human could understand in a given context, and explicitly explaining such decisions. This way, it is possible to scrutinize the intelligent models and verify if automated decisions are made on the basis of accepted rules and principles, so that decisions can be trusted and their impact justified.

This Special Issue is supported by the IEEE CIS Task Force on Explainable Fuzzy Systems (TF-EXFS). The mission of the TF-EXFS is to lead the development of a new generation of Explainable Fuzzy Systems, with a holistic view of fundamentals and current research trends in the XAI field, paying special attention to fuzzy-grounded knowledge representation and reasoning but also regarding how to enhance human-machine interaction through multi-modal (e.g., graphical or textual modalities) effective explanations.

The scope of this special issue is not limited to the community of researchers in Fuzzy Logic but it is open to contributions by researchers, from both academy and industry, working in the multidisciplinary field of XAI.

Topics

This special issue is targeted on general readership articles about design and application of XAI technologies. Topics of interest include, but are not limited to:

- Theoretical Aspects of Explainability, Fairness, Accountability and Transparency
- Relations between Explainability and other Quality Criteria (such as Interpretability, Accuracy, Stability, Relevance, etc.)
- Dimensions of Interpretability: Readability versus Understandability
- Explainability Evaluation and Improvements
- Learning Methods and Design Issues for Explainable Systems and Models
- Interpretable Machine Learning
- Explaining Black-box Models
- Hybrid Approaches (e.g., Neuro-Fuzzy systems) for XAI
- Model-specific and Model-agnostic Approaches for XAI
- Models for Explainable Recommendations
- Explainable Conversational Agents
- Self-explanatory Decision-Support Systems
- Factual and Counterfactual Explanations
- Causal Thinking, Reasoning and Modeling

- Cognitive Science and XAI
- Argumentation Theory for XAI
- Natural Language Technology for XAI
- Human-Machine Interaction for XAI
- Ethics and Legal Issues for XAI
- XAI-based Data Analysis and Bias Mitigation
- Safe and Trustworthy AI
- Applications of XAI-based Systems
- Open Source Software for XAI

Submission

The IEEE Computational Intelligence Magazine (CIM) publishes peer-reviewed high-quality articles. All manuscripts must be submitted electronically in PDF format. Manuscripts must be in standard IEEE two-column/single space format and adhere to a length of 10-12 pages (including figures and references) for regular papers. A mandatory page charge is imposed on all papers exceeding 12 pages in length.

More information on manuscript details and submission guidelines can be found at the following websites:

- Special Issue website: https://sites.google.com/view/special-issue-on-xai-ieee-cim
- IEEE CIM website: https://cis.ieee.org/publications/ci-magazine/cim-information-for-authors

Important Dates

- Manuscript Due: **15 February, 2021**
- First Notification: 15 April, 2021
- Revision Due: 15 May, 2021
- Final Notification: 1 July, 2021

Guest Editors

- **José María Alonso,** University of Santiago de Compostela, Spain (josemaria.alonso.moral@usc.es)

- **Corrado Mencar,** University of Bari Aldo Moro, Italy

- **Hisao Ishibuchi,** Southern University of Science and Technology, China