



Fairness and Explanation for Trustworthy AI

Guest Editors:

Dr. Jianlong Zhou

Prof. Dr. Andreas Holzinger

Prof. Dr. Fang Chen

Deadline for manuscript
submissions:

closed (15 December 2023)

Message from the Guest Editors

Artificial intelligence (AI) and machine learning (ML) are increasingly being used to shape our daily lives by making, or at least influencing, decisions with ethical and legal implications in a variety of application areas. However, due to biased input data and/or flawed algorithms, unfair AI-informed decision-making systems may result in reinforcing discrimination, such as racial/gender bias in AI-informed decision-making, or even in high risk environments due to incorrect decisions. Such requirements need to provide re-traceability, explainability, interpretability, and transparency for such AI systems—which is technically challenging. Meanwhile, fairness and explanations are key components in fostering trust and confidence in AI systems. In this Special Issue, we will feature cutting-edge research where fairness and explanations are presented for making trustworthy decisions in AI systems.

This Special Issue invites submissions that feature original research on designing, presenting, and evaluating approaches for fairness and explanations in AI systems.





an Open Access Journal by MDPI

Editor-in-Chief

Prof. Dr. Andreas Holzinger

1. Human-Centered AI Lab,
Institute of Forest Engineering,
Department of Ecosystem
Management, Climate and
Biodiversity, BOKU University,
Vienna, Austria
2. Institute of Human-Centered
Computing, Faculty of Computer
Science and Biomedical
Engineering, Graz University of
Technology, Graz, Austria
3. xAI Lab, Alberta Machine
Intelligence Institute, University
of Alberta, Edmonton, AB,
Canada

Message from the Editor-in-Chief

Machine learning deals with understanding intelligence to design algorithms that can learn from data, gain knowledge from experience and improve their learning behaviour over time. The challenge is to extract relevant structural and/or temporal patterns (“knowledge”) from data, which is often hidden in high dimensional spaces, thus not accessible to humans. Many application domains, e.g., smart health, smart factory, etc. affect our daily life, e.g., recommender systems, speech recognition, autonomous driving, etc. The grand challenge is to understand the context in the real-world under uncertainty. Probabilistic inference can be of great help here as the inverse probability allows to learn from data, to infer unknowns, and to make predictions to support decision making.

Author Benefits

Open Access: free for readers, with [article processing charges \(APC\)](#) paid by authors or their institutions.

High Visibility: indexed within [Scopus](#), [ESCI \(Web of Science\)](#), [dblp](#), and [other databases](#).

Rapid Publication: manuscripts are peer-reviewed and a first decision is provided to authors approximately 27 days after submission; acceptance to publication is undertaken in 4.4 days (median values for papers published in this journal in the second half of 2025).

Contact Us

*Machine Learning and Knowledge
Extraction* Editorial Office
MDPI, Grosspeteranlage 5
4052 Basel, Switzerland

Tel: +41 61 683 77 34
www.mdpi.com

mdpi.com/journal/make
make@mdpi.com
[X@MAKE_MDPI](https://twitter.com/MAKE_MDPI)