Preface to the Proceedings of BEWARE-26

The fourth edition of the BEWARE workshop (https://sites.google.com/view/beware2026/) was held in Perugia, Italy, on October 8, 2026, as part of the programme of the 24th International Conference of the Italian Association for Artificial Intelligence (AIxIA 2026).

BEWARE is a Workshop Series focusing on bias, risk and explainability in Artificial Intelligence, and on the contribution of logic and logic programming to understanding and addressing these issues. The workshop provides a forum for connecting philosophical and ethical analysis with formal models and computational methods. Its scope ranges from the epistemic and normative analysis of principles for AI to their representation in formal frameworks and their implementation in the design of AI systems. This dialogue is particularly relevant when assessing how automated systems can support human judgement, how their limitations should be communicated, and how their behaviour can be evaluated.

This year the workshop received 16 submissions, of which 10 papers were accepted and presented (62.5%). Submissions underwent single-blind peer review, with two completed reviews for each paper. The contributions address the workshop's themes through conceptual analysis, formal approaches and applications, reflecting the range of perspectives that motivates the BEWARE series.

BEWARE 2026 brought together philosophical, ethical and technical perspectives on understanding and assessing AI. The morning session explored automation and attribution bias (Aldo Pisano, "The Bodiless Oracle: A Critical Literacy against the Automation and Attribution Bias"); human and AI assessments of stereotypes (Elisa Famulari, Silvana Badaloni and Antonio Rodà, "Comparing human and Large Language Model judgments of gender stereotypes in Italian texts"); and the challenges of evaluating AI explanations (Nicolò Cristante, "Faithfulness All the Way Down: A Wittgensteinian Diagnosis of the XAI Evaluation Regress"). It also examined predictive models in healthcare (Martina Baltuzzi, "The Patient as a Prediction: Four Epistemological Challenges of Digital Twins in Health"), and ethics of care in assistive robotics (Gabriele Nino and Francesca Alessandra Lisi, "Human-AI Symbiosis and the Ethics of Care: Towards Care-Sensitive Design in Assistive Robotics"), concluding with questions about AI understanding and responsibility (Emanuele Damiano and Francesco Orciuoli, "Contradiction Awareness in Ontology-grounded LLMs: An Epistemological Perspective"). The afternoon shifted towards tools and formal approaches: formal ontologies, Logical Neural Networks and proof theoretical methods with application to fairness in language and other AI technologies for audit purposes (Annapaola Ginammi, "Formal ontologies for fair language technologies: A conceptual exploration"; Simone Giovannini, Fabio Coppini and Simone Marinai, "Language agnostic document key value extraction using Logical Neural Networks"; Melissa Antonelli, Alessandro G. Buda, Leonardo Ceragioli and Giuseppe Primiero, "Towards a proof theory for reasoning about frequencies on series"; and Alessio Donvito "A General Quantitative Methodology for the Ethical Auditing of AI Systems"). Together, these contributions connected broad questions about trust, bias and responsibility with concrete approaches to designing and auditing AI.

The extended submission deadline was July 22, 2026. The call for papers advertised August 30, 2026, as the acceptance-notification deadline. The final camera-ready deadline communicated to authors was October 15, 2026.

We thank the members of the Program Committee for their time, expertise and careful assessment of the submissions:

- Melissa Antonelli, University of Tübingen, CFvW
- Paolo Baldi, University of Salento
- Alessandro Giuseppe Buda, School of Advanced Studies IUSS Pavia and University of Milan

BEWARE-26 (Bias, Risk, Explainability and the role of Logic and Logic Programming), co-located with AIxIA 2026, October 8, 2026, Perugia, Italy



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

published 2026-11-17

CEUR-WS.org/Vol-4332/preface.pdf

- Stefania Costantini, University of L'Aquila
- Fabio Aurelio D'Asaro, University of Verona
- Giovanni De Gasperis, University of L'Aquila
- Roberta Ferrario, Institute of Cognitive Sciences and Technologies, CNR
- Chiara Gallese, Tilburg University
- Francesca Alessandra Lisi, University of Bari Aldo Moro
- Ludovica Marinucci, Consiglio Nazionale delle Ricerche
- Corrado Mencar, University of Bari Aldo Moro
- Luca Pasetto, University of Luxembourg
- Aldo Pisano, SIpEIA
- Valentina Pitoni, University of L'Aquila
- Giuseppe Primiero, University of Milan
- Teresa Scantamburlo, University of Trieste
- Viola Schiaffonati, Politecnico di Milano
- Alberto Termine, SUPSI

and the additional reviewers for their last-minute help

- Alessio Donvito, University of Bari Aldo Moro
- Piero Marra, University of Bari Aldo Moro
- Gabriele Nino, University of Bari Aldo Moro
- Lucia Siciliani, University of Bari Aldo Moro

We are grateful to all authors for contributing their work, to all the reviewers for their constructive feedback, and to the participants for the discussions that form the basis of the workshop. We also thank the AIxIA 2026 organizers for hosting BEWARE-26. Microsoft CMT was used to manage submissions and reviewing.

We hope that these proceedings will encourage further exchange between researchers working on the conceptual foundations of trustworthy AI and those developing the formal and computational tools needed to put these foundations into practice.

October 2026
The Organizers,

- **Alessandro Giuseppe Buda**, School of Advanced Studies IUSS Pavia and University of Milan, Italy
- **Fabio Aurelio D'Asaro**, University of Verona, Italy
- **Francesca Alessandra Lisi**, University of Bari Aldo Moro, Italy
- **Giuseppe Primiero**, University of Milan, Italy