



## **The 5<sup>th</sup> SPECIAL SESSION ON EXPLAINABLE ARTIFICIAL INTELLIGENCE FOR UNVEILING THE BRAIN: FROM THE BLACK-BOX TO THE GLASS-BOX**

**(XAIB2025)**

Half-day

(hybrid: in person and online)

### • Co-chairs:

Prof. **Alessia Sarica**, Neuroscience Research Center, Department of Medical and Surgical Sciences, Magna Graecia University of Catanzaro, Italy, [sarica@unicz.it](mailto:sarica@unicz.it)

Dr. **Vincenzo Dentamaro**, Computer Science Department, University of Bari Aldo Moro, Bari, Italy, [vincenzo.dentamaro@uniba.it](mailto:vincenzo.dentamaro@uniba.it)

### • Short CV of co-chairs:

**Alessia Sarica** is Associate Professor of Advanced Medical and Surgical Technology and Methodology at the Neuroscience Research Center, Magna Graecia University of Catanzaro, Italy. She received the Ph.D. degree in Biomedical and Computer Science Engineering in 2015, and her main research field is Machine Learning on Neuroimaging and Neuropsychological data for the early and differential diagnosis of neurodegenerative diseases. She conceived and organized the XAIB2021, XAIB2022, XAIB2023 and XAIB2024 for the Brain Informatics Conference.

**Vincenzo Dentamaro** is Assistant Professor (RTD/A) at the Department of Computer Science, University of Bari Aldo Moro, Italy. He received his Ph.D. in Computer Science and Mathematics cum laude in 2022. His main research fields include Advanced

Machine Learning Methodologies, Computational Neuroscience, and AI applications in Clinical Settings, with a focus on early detection of neurodegenerative diseases using multimodal approaches. He has founded multiple startups, including Nextome S.r.L., and holds several international patents in positioning technologies. He conceived and organized multiple workshops including AI4Alzheimer (WCCI 2022), BMVC2023 Workshop on AI for Neurodegenerative Diseases, and NeroPRAI 2024, among others.

## Introduction

**Abstract:** Artificial Intelligence (AI) and Machine Learning (ML) have become indispensable tools in brain research, with applications ranging from neuroimage processing and analysis to the automated diagnosis of neurodegenerative diseases. However, despite their high accuracy, the black-box nature of many ML models hinders trust and interpretability, making their integration into clinical practice challenging. Without a clear explanation of their findings, AI-driven medical and clinical decisions remain difficult to validate and adopt.

In response to this challenge, there has been a growing demand for interpretable and explainable AI (XAI), particularly in healthcare. Researchers are increasingly focusing on glass-box models that balance intelligibility and optimal performance, ensuring that AI-driven insights are both accurate and transparent. This Special Session aims to advance the field by showcasing novel methods that enhance the explainability, interpretability, and uncertainty quantification of AI models applied to neuroscience, neuroimaging, and neuropsychological data.

**Scope and topics:** This Special Session seeks to bring together cutting-edge research on Explainable Artificial Intelligence (XAI) in Neuroscience, Neuroimaging, and Neuropsychological data analysis to unveil the complexities of the brain. Researchers are invited to submit high-quality papers or abstracts on novel and state-of-the-art intelligible, interpretable, and understandable AI approaches, including but not limited to:

- Post-hoc explainability techniques (both model-agnostic, e.g., LIME, SHAP, and model-specific, e.g., CNN, SVM, Random Forests)
- Transparent and inherently interpretable models (e.g., linear/logistic regression, decision trees, Generalized Additive Models (GAMs))
- Uncertainty quantification in brain exploration, including Bayesian

inference, conformal prediction, probabilistic graphical models, and ensemble methods for improving reliability in neuroimaging and clinical decision-making

- Robustness and trustworthiness of AI models, addressing issues such as bias, calibration, and out-of-distribution generalization in neuroscience applications
- Global and local explainability approaches to better understand AI-based predictions in neurological disorders
- Systematic reviews and meta-analyses on XAI in neuroscience

### **Motivation and Rationale**

In its past four editions, XAIB has established itself as a leading forum for experts in interpretable machine learning, featuring distinguished scientists such as Dr. Rich Caruana, Dr. Michele Ferrante, Dr. Dimistris Pinotsis, Prof. Monica Hernandez, Dr. Bojan Bogdanovic, Dr. Parul Verma, and Dr. Antonio Parziale. These speakers have shared insights on the state-of-the-art in XAI, their ongoing research, and the transformative potential of explainable AI in uncovering new knowledge about the brain and its underlying mechanisms.

The 2021 and 2022 editions featured comprehensive discussions on cutting-edge explainability techniques, while the 2023 session at Stevens Institute of Technology saw an unprecedented level of participation and engagement. The 2024 edition hosted Christoph Molnar, a widely recognized author of machine learning books, as the invited speaker, further enriching the discourse on XAI.

With the growing importance of trustworthy AI in healthcare, XAIB2025 will continue this tradition by exploring the next frontier: integrating uncertainty quantification into explainable models, ensuring both interpretability and reliability in brain research and clinical applications.

### **Special Session Details**

• Program committee:

- **Vera Gramigna**, Neuroscience Research Center, Department of Medical and Surgical Sciences, Magna Graecia University of Catanzaro, Italy, [gramigna@unicz.it](mailto:gramigna@unicz.it)
- **Chiara Camastra**, Neuroscience Research Center, Department of Medical and Surgical Sciences, Magna Graecia University of Catanzaro, Italy, [chiara.camastra@unicz.it](mailto:chiara.camastra@unicz.it)

- **Assunta Pelagi**, Neuroscience Research Center, Department of Medical and Surgical Sciences, Magna Graecia University of Catanzaro, Italy, [assunta.pelagi@studenti.unicz.it](mailto:assunta.pelagi@studenti.unicz.it)
- **Felice Franchini**, Department of Precision and Regenerative Medicine, University of Bari Aldo Moro, Bari, Italy, [f.franchini9@phd.uniba.it](mailto:f.franchini9@phd.uniba.it)
- **Stefano Galantucci**, Department of Computer Science, University of Bari Aldo Moro, Bari, Italy, [stefano.galantucci@uniba.it](mailto:stefano.galantucci@uniba.it)

**Call for Papers and Abstracts:** We are pleased to invite submissions for the 5th Special Session on Explainable Artificial Intelligence for Unveiling the Brain: From the Black-Box to the Glass-Box (XAIB2025). This half-day hybrid session (in-person and online) will be held as part of the Brain Informatics Conference 2025. Manuscripts should be written in English and conform to the Brain Informatics Conference 2025 formatting guidelines. All submissions will undergo peer review and will be evaluated based on originality, technical quality, and relevance to the session themes. Accepted papers and abstracts will be presented at the workshop, with accepted papers included in the conference proceedings.

Expected number of participants and the expected number of submissions:  
5-8 submissions.

**Tentative Internal and External Schedule (Except adjustments by the Chairs to align all workshop/special session schedules)**

- 20 April 2025: Workshop/Special session Proposal Deadline
- 20 June 2025: Full Paper Submission Deadline
- 20 July 2025: Abstract Presentation Submission Deadline
- 10 August 2025: Final Paper and Abstract Acceptance Notification
- 30 August 2025: Accepted Paper and Abstract Registration Deadline
- 11 Nov 2025: Workshops and Special Sessions
- 12-13 Nov 2025: The Brain Informatics Conference