

| About the Conference

Please consider to contribute to and/or forward to the appropriate groups the following opportunity to submit and publish original scientific results to:

- EXPLAINABILITY 2025, The Second International Conference on Systems Explainability

EXPLAINABILITY 2025 is scheduled to be October 26 - 30, 2025 in Barcelona, Spain under the TechWorld 2025 umbrella.

The submission deadline is July 8, 2025.

Authors of selected papers will be invited to submit extended article versions to one of the IARIA Journals:

<https://www.ariajournals.org>

All events will be held in a hybrid mode: on site, online, prerecorded videos, voiced presentation slides, pdf slides.

=====

===== EXPLAINABILITY 2025 | Call for Papers


=====


CALL FOR PAPERS, TUTORIALS, PANELS

EXPLAINABILITY 2025, The Second International Conference on Systems Explainability

General page:

<https://www.aria.org/conferences2025/EXPLAINABILITY25.html>

 **Important Dates**

 **CONFERENCE DATE**
October 26-30, 2025

Submission page:

<https://www.iaia.org/conferences2025/SubmitEXPLAINABILITY25.html>

Event schedule: October 26 - 30, 2025

Contributions:

- regular papers [in the proceedings, digital library]
- short papers (work in progress) [in the proceedings, digital library]
- ideas: two pages [in the proceedings, digital library]
- extended abstracts: two pages [in the proceedings, digital library]
- posters: two pages [in the proceedings, digital library]
- posters: slide only [slide-deck posted at www.iaia.org]
- presentations: slide only [slide-deck posted at www.iaia.org]
- demos: two pages [posted at www.iaia.org]

Submission deadline: July 8, 2025

Extended versions of selected papers will be published in IARIA Journals: <https://www.iaiajournals.org>

Print proceedings will be available via Curran Associates, Inc.: <https://www.proceedings.com/9769.html>

Articles will be archived in the free access ThinkMind Digital Library: <https://www.thinkmind.org>

The topics suggested by the conference can be discussed in

term of concepts, state of the art, research, standards, implementations, running experiments, applications, and industrial case studies. Authors are invited to submit complete unpublished papers, which are not under review in any other conference or journal in the following, but not limited to, topic areas.

All tracks are open to both research and industry contributions. Before submission, please check and comply with the editorial rules: <https://www.iaria.org/editorialrules.html>

EXPLAINABILITY 2025 Topics (for topics and submission details: see CfP on the site)

Call for Papers:

<https://www.iaria.org/conferences2025/CfPEXPLAINABILITY25.html>

=====

EXPLAINABILITY 2025 Tracks (topics and submission details: see CfP on the site)

Concepts for the foundation of explainability

- Explainability requirements
- Explainability for a diverse audience
- Standards to support a device-agnostic cooperation
- Explainability via inclusivity, empathy, and emotion adoption
- Post hoc explainability
- Design guidelines for explainable interfaces
- Causality and explainability
- Interpretability and understandability
- Procedural vs distributive fairness

- Fairness, accountability, and transparency
- Interpretability methods (predictive accuracy, descriptive accuracy, and relevancy)
- Relation: prediction, accuracy, explainability, and trust

Explainability Models

- Transparent models for practitioners and users
- Unifying approach for interpreting model predictions
- Design guidelines for explainable models
- Explainable levels vs prediction accuracy of results
- Local explanations to global understanding
- Intrinsic explainable models
- Trustfulness and acceptability models
- Model interpretability
- Black-box machine learning models (LIME, SHAP)

Classical Explainability Revisited

- Improve Product "User's Manual"
- Essentials in Drug explanation side effects
- Directory of FAQ (Frequently Asked Questions)
- Explanatory buyer's contacts
- Adverse analytics of laws and governmental decisions
- Observability and in-context interpretability

- Explainability via social networks
- Explainability via validated reputation metrics

Explainability Classical Tools

- Interpretation model of product/software predictions
- Key Performance Indicators (KPIs)
- Repository of data models
- Interpretability models
- Explainability for human-in-the-middle systems
- Cultural context-sensitive social explainability guidelines

Explainable (personalized) Interfaces

- Explainable models for personality
- Explainability and social norms
- Explainability in personality design
- Explainability on emotional interaction
- Explainability for tactile and haptic interactions
- Explainability for linguistics of personality needs
- Explainability for conversational user interfaces (CUIs) (e.g., text-based chatbots and voice-based assistants)
- Observable personality
- Explainability for impaired users

Explainable Software

- Explainability by-design (designer/programmer comments)
- Challenges for tracking requirements thru the deployment process
- Transparency levels (interface, component, the entire model, learning algorithms)
- Screening methods for deviation and bias (data and algorithms)
- Black box vs Explainable box
- Insights on model failures/performance
- Explainability feature for evaluation of software analytics models
- Design for approachability
- IF-THEN understanding vs scalability
- Metrics and metrology for compliance validation with the requirements

Explainability of Data Processing Algorithms

- Classification Prediction accuracy vs Explainability
- Deep Learning (Neural Networks)
- Support Vector Machines
- Ensemble Methods (e.g., Random Forests)
- Graphical Models (e.g., Bayesian Networks)
- Decision Trees, Classification Rules
- Convolutional Neural Networks (for images)

Datasets Explainability

- Training datasets vs validation datasets selection explainability
- Poor explainability from huge data patterns
- Methods for pattern explanation
- Explainability on validation algorithms and thresholds selection
- Explainability on computation power vs performance trade-off
- Post hoc on a dataset (in biostatistics data analytics)
- Explaining type-specific topic profiles of datasets
- Transformers datasets (for natural language processing model)
- Explainability of heterogeneous dataset collections

Personalized Datasets (DS) Explainability

- Universal vs. cultural personalized datasets
- Sensitive social cues to the cultural context
- Ramifications of personality
- Observable personality
- Explainability for impaired users

Explainability in Small Datasets

- Explainability between small data and big data

- Statistics on small data
- Handling small datasets
- Predictive modeling methods for small datasets
- Small and incomplete datasets
- Normality in small datasets
- Confidence intervals of small datasets
- Causal discovery from small datasets
- Dynamic domain-oriented small datasets (health, sentiment, personal behavior, vital metrics, mobility)

Machine Learning (ML) Explainability

- Taxonomy for ML Interpretability
- ML Interpretability (ML model accuracy for a valid 'from cause to effect')
- ML vs machine personality
- Explainability of opacity and non-intuitiveness models
- Explainability for ML models (supervised, unsupervised, reinforcement, constrained, etc.);
- Explainability for generative modeling (Gaussian, HMM, GAN, Bayesian networks, autoencoders, etc.)
- Explainability of prediction uncertainty (approximation learning, similarity, quasi-similarity)
- Training of models (hyperparameter optimization, regularization, optimizers)
- Explainability of data types (no data, small data, big data,

graph data, time series, sparse data, etc.)

- Explainability of hardware-efficient machine learning methods

- Methods to enhance fairness in ML models

Deep Learning (DL) Explainability

- Explainability for Sentiment Analysis

- Active learning (partially labels datasets, faulty labels, semi-supervised)

- Details on model training and inference

- Data Inference for Small/Big Data

- Theoretical models for Small/Big Data

- (Integrated) Gradients explanation technique

- Deep LIFT (deep neural predictions)

- Guided BackPropagation, Deconvolution (Convolution Networks)

- Class Activation Maps (CAMs), GradCAM, Layer-wise Relevance Propagation (LRP)

- RISE algorithm (prediction of Deep Neural Networks for images)

Explainable AI

- Large Language Models (LLM)

- Autoregressive language models

- Limitation of AI-based analytics agents

- Visibility into the AI decision-making process
- Explainable AI (feature importance, LIME, SHAP, etc.)
- Local Interpretable Model-agnostic Explanations (LIME)
- Shapley additive explanations (SHAP) (multiple explanations for different kinds of models)
- User role-based and system target-based AI explainability

Explainability at work

- Lessons learned for deploying explainable models
- Limitation self-awareness
- Limitation by design (critical missions)
- Controlled machine personality
- Setting wrong expectations
- Wrong (misleading) explainability models
- Pitfalls of explainable ML
- Missing needs for various stakeholders

AI/ML/DS/DL Explainability tools

- Open-source experimental environments
- Matching observability perception vs official explainability
- Precision model-agnostic explanations
- Criticism for interpretability
- Fairness-aware ranking

- Conflicting explanations
- Additive explanations
- Counterfactual explanations
- Datasets-based tools (e.g., collection faces reacting to robots making mistakes)
- Explainability for emerging artificial intelligent partners (robots, chatbots, driverless car transportation systems, etc.)
- Bias detection for diversity and inclusion
- Small datasets for benchmarking and testing
- Small data toolkits
- Data summarization

Explainability case studies

- Lessons learned with existing generative-AI tools (ChatGPT, Bard AI, ChatSonic, etc.)
- Sentiment analysis:
 - - Explainability DL for sentiment analysis (detection: bias, hate speech, emotions; models)
 - - Word-embedding and embedding representations
 - - Lexicon-based explainability for sentiment analysis
- Industry AI explainability
 - - Predictive maintenance
 - - Robot-based production lines
 - - Pre-scheduled renewals of machinery

- - Pharmaceutical

- Output explainability for other case studies

- - Social networks

- - Educational environments

- - Healthcare systems

- - Scholarly discussions (e.g., peer review process discussions, mailing lists, etc.)

- - Mental health systems

- - Human fatigue estimation

- - Hazard prevention

EXPLAINABILITY 2025 Committee:

<https://www.iaria.org/conferences2025/ComEXPLAINABILITY25.html>

Open Access Special Advertising and Publicity Board

Lorena Parra Boronat, Universitat Politècnica de Barcelona,
Spain

Laura Garcia, Universidad Politécnica de Cartagena, Spain

José Miguel Jimenez, Universitat Politècnica de Barcelona,
Spain

Sandra Viciano Tudela, Universitat Politècnica de Barcelona,
Spain

Francisco Javier Díaz Blasco, Universitat Politècnica de
València, Spain

Ali Ahmad, Universitat Politècnica de València, Spain

 **TOPICS OF INTEREST**

3 topics

Research papers are invited in, but not limited to, the following areas:

Multidisciplinary & General

Uncategorized

Artificial Intelligence & Machine Learning