7th Workshop on Explainable User Models and Personalised Systems (ExUM 2025)

Cataldo Musto
Department of Computer Science
Dipartimento di Informatica - University of Bari
Bari, Italy
cataldo.musto@uniba.it

Marco Polignano
Computer Science
Università degli Studi di Bari Aldo Moro
Bari, Italy
marco.polignano@uniba.it

Amon Rapp
Computer Science Department
University of Turin
Torino, Italy
amon.rapp@unito.it

Giovanni Semeraro
University of Bari Aldo Moro
Bari, Italy
giovanni.semeraro@uniba.it

Jürgen Ziegler
University of Duisburg - Essen
Duisburg, Germany
juergen.ziegler@uni-due.de

Abstract

In recent years, adaptive and personalized systems, underpinned by cutting-edge technologies such as Large Language Models (LLMs), have emerged as pivotal forces in reshaping the digital landscape. These systems, seamlessly woven into the fabric of everyday life, manifest in diverse forms—from conversational agents that emulate human-like dialogue to recommendation algorithms that tailor content like music, films, and consumer products to individual preferences. Their pervasive integration into digital platforms has fundamentally altered the ways in which users engage with information, make decisions, and interact with technology, positioning them as indispensable tools in modern society. The transformative potential of these technologies lies in their ability to enhance user engagement, streamline content delivery, and support decision-making processes with unprecedented precision. However, as their influence continues to expand, so too does the urgency to confront critical challenges surrounding transparency, fairness, and user trust. ExUM workshop seeks to explore these pressing issues, advocating for a balanced approach that prioritizes not only technological efficacy but also ethical integrity and user empowerment.

CCS Concepts

• **Information systems** → **Personalization**; **Recommender systems**; • **Human-centered computing** → **User models**; • **Computing methodologies** → **Natural language processing**.

Keywords

Explainability, Explanations, Transparency, Interpretability, User Models, Personalization, Recommender Systems

UMAP Adjunct '25, New York City, NY, USA

ACM ISBN 979-8-4007-1399-6/25/06
https://doi.org/10.1145/3708319.3727557

ACM Reference Format:
Cataldo Musto, Marco Polignano, Amon Rapp, Giovanni Semeraro, and Jürgen Ziegler. 2025. 7th Workshop on Explainable User Models and Personalised Systems (ExUM 2025). In *Adjunct Proceedings of the 33rd ACM Conference on User Modeling, Adaptation and Personalization (UMAP Adjunct '25), June 16–19, 2025, New York City, NY, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3708319.3727557

1 Introduction

Transparency is essential for building trust in adaptive and personalized systems, especially as they increasingly impact critical areas like employment and healthcare. Regulations such as the EU GDPR and AI Act now legally require that individuals receive clear explanations of how intelligent systems process data and make decisions. While research has traditionally prioritized maximizing system performance, this focus has often led to "black box" technologies that lack transparency, potentially undermining user trust and failing to meet ethical and regulatory demands. There's a clear tension between achieving high-performance personalization and providing explainable, transparent results. Users may distrust even the most accurate systems if they do not understand how decisions are made, which can raise concerns about fairness, bias, and misuse of personal data. These challenges highlight the need for explainability and transparency to be treated as core design principles rather than afterthoughts, given that opaque systems can reinforce social inequalities and prompt scrutiny from regulators and researchers alike. Addressing such issues requires more than technical improvements: developers should adopt user-centric design approaches that embed transparency from the outset, involve users in system development, and provide accessible explanations and privacy controls. Collaboration among technologists, ethicists, policymakers, and end-users is essential to ensure responsible system development that meets societal and regulatory expectations. Moving forward, adaptive technologies must balance innovation with ethical responsibility by developing scalable, sector-spanning frameworks for explainability. Sustained collaboration and ongoing evaluation are key to keeping adaptive systems aligned with evolving user needs, ethical norms, and regulatory standards—ultimately aiming to deliver technologies that are innovative, transparent, and respectful of societal values. The ExUM workshop aims to provide a

forum to discuss such problems, challenges, and innovative research approaches in the area, by investigating the role of transparency and explainability in the recent methodologies for building user models or developing personalized and adaptive systems.
Topics of interest include but are not limited to:

- **Transparent and Explainable Personalization Strategies** (1) Scrutable User Models, (2) Transparent User Profiling and Personal Data Extraction, (3) Explainable Personalization and Adaptation Methodologies, (4) Novel strategies (e.g., conversational recommender systems) for building transparent algorithms, or (5) Transparent Personalization and Adaptation to Groups of users
- **Designing Explanation Algorithms** (1) Explanation algorithms based on item description and item properties, (2) Explanation algorithms based on user-generated content (e.g., reviews), (3) Explanation algorithms based on collaborative information, (4) Building explanation algorithms for opaque personalization techniques, (e.g., neural networks, matrix factorization), or (5) Explanation algorithms based on methods to build group models
- **Designing Transparent and Explainable User Interfaces** (1) Transparent User Interfaces, (2) Designing Transparent Interaction methodologies, or (3) Novel paradigms (e.g. chatbots) for building transparent models
- **Evaluating Transparency and Explainability** (1) Evaluating Transparency in interaction or personalization, (2) Evaluating Explainability of the algorithms, (3) Designing User Studies for evaluating transparency and explainability, or (4) Novel metrics and experimental protocols
- **Open Issues in Transparent and Explainable User Models and Personalized Systems** (1) Ethical issues (Fairness and Biases) in User / Group Models and Personalized Systems, (2) Privacy management of Personal and Social data

2 Accepted Papers

We accepted 10 papers (out of 11 papers submitted).

(1) Milad Sabouri, Masoud Mansoury, Kun Lin and Bamshad Mobasher in "***Towards Explainable Temporal User Profiling with LLMs***" The paper describe a framework that leverages large language models (LLMs) to generate natural language summaries of users' interaction histories, explicitly distinguishing between recent (short-term) and persistent (long-term) preferences. By encoding these summaries and dynamically fusing them with an attention mechanism, the approach creates comprehensive user profiles that adapt to evolving interests. This method not only improves recommendation performance over traditional baselines but also enhances transparency by providing interpretable textual explanations and attention-based insights into recommendation decisions.

(2) Muhammad Raees, Vassilis-Javed Khan and Konstantinos Papangelis in "***Towards Understanding Persuasive and Personalized Engagement for Human-AI Reliance***" tackle the topic of a novel framework that uses large language models (LLMs) to generate natural language summaries of users' interaction histories, effectively distinguishing between short-term and long-term preferences. By encoding these summaries and combining them through an attention mechanism, the approach forms a dynamic and comprehensive user profile that better captures evolving interests. Experimental results on real-world datasets demonstrate both improved performance and enhanced explainability for content-based recommender systems.

(3) Sofie Goethals, Joao Sedoc and Foster Provost in "***What If the Prompt Were Different? Counterfactual Explanations for the Characteristics of Generative Outputs***" describe a framework for applying counterfactual explanations to nondeterministic generative AI systems, particularly when their outputs are used as inputs for downstream classifiers. By addressing the challenge of assessing whether generative AI outputs display attributes like toxicity, sentiment, or bias, the approach enables more effective analysis and interpretation of generative system behaviors in real-world applications. This method enhances interpretability by providing actionable insights into how changes in generated content could affect subsequent classifications.

(4) Mohammed Muheeb Ghori, Arman Dehpanah, Jonathan Gemmell and Bamshad Mobasher in "***They Only Offer the Illusion of Choice: Exploring User Perceptions of Control and Agency on YouTubein***" describe a study that investigates how users perceive and exercise control over YouTube's recommender system by analyzing 200 Reddit discussion threads through thematic analysis. It explores user experiences and strategies for exerting agency, highlighting the prevalent opacity in how recommendation systems function and how they can be personalized. The findings reveal insights into users' awareness of available control mechanisms and their effectiveness in tailoring recommendations, shedding light on the relationship between perceived control, trust, and the perceived usefulness of recommender systems.

(5) Jasmina Gajcin, Jovan Jeromela and Ivana Dusparic in "***Towards Personalised and User-Friendly Counterfactual Sequences for Failure Correction***" claim that a human-centered approach to counterfactual sequence generation for sequential decision-making tasks, with a focus on improving failure correction, user trust, and personalized policy development. While previous work has emphasized algorithmic aspects, this study highlights the value of counterfactuals as user-friendly explanations. Drawing on insights from psychology and cognitive science, the authors propose key requirements for making counterfactual sequences more understandable and useful. They then present an algorithm that generates diverse, personalized counterfactual sequences, empowering users to better understand and correct failures according to their own preferences.

(6) Cataldo Musto, Amon Rapp and Ludovico Boratto in "***PhASE Project - Promoting Healthy and Sustainable Eating through Interactive and Explainable AI Methods***" describe a project, funded by the Italian Ministry of University and Research under the PRIN program, promotes a healthier and more sustainable diet through the integration of recommender systems and conversational agents. The project aims

to foster users' awareness and perception of responsible eating, enabling people to explore recipe healthiness, identify more sustainable substitute ingredients, and receive tailored advice.

(7) Miguel Portaz, Angeles Manjarrés, Olga C. Santos, Raúl Cabestrero, Pilar Quirós, Mar Hermosilla, David Puertas-Ramirez, Jesus G. Boticario, Gadea Lucas Pérez, Ana Serrano-Mamolar, Álvar Arnaiz-González, Miguel Arevalillo-Herráez, David Arnau, Pablo Arnau-González, Raúl Fernández-Matellán and David Martin Gomez in "***Developing Human-Centered Intelligent Learning Systems: the application of CARAIX framework***" present the use of the CARAIX framework—a Collaborative, Adaptive, and Responsible AI assisted by eXplainability—for developing intelligent educational systems that prioritize adaptability, explainability, ethical compliance, and user trust. CARAIX is grounded in the principles of hybrid intelligence and incorporates explainable AI techniques to support real-time feedback and bias-aware decision-making within educational contexts. The study validates CARAIX by applying a peer-reviewed rubric to analyze its integration across four learning scenarios: intelligent tutoring, psychomotor skill development, autonomous driving training, and occupational safety education. Designed for scalability and reusability, the framework demonstrates its potential to create sustainable, human-centered, and ethically robust AI-driven learning environments.

(8) Simone Prete, Giovanni Maria Biancofiore, Fedelucio Narducci, Eugenio Di Sciascio and Tommaso Di Noia in "***Explainable Sentiment Analysis through Counterfactual Reasoning***" present an approach to enhancing interpretability in sentiment analysis by integrating eXplainable AI (XAI) techniques, specifically SHAP (SHapley Additive exPlanations) and counterfactual generation, into sentiment classification of movie reviews. By combining a neural classifier with these methods, the study identifies the key words influencing model predictions and generates counterfactual examples illustrating how small changes in text can alter sentiment outcomes. Experimental results demonstrate that SHAP-based attribution and counterfactual analysis offer valuable insights into the linguistic features guiding sentiment classification decisions, making the model's inner workings more transparent.

(9) Lilian Marey, Charlotte Laclau, Bruno Sguerra, Tiphaine Viard and Manuel Moussallam in "***Modeling Musical Genre Trajectories through Pathlet Learning***" describe a framework that applies the dictionary learning paradigm to model how musical tastes evolve on music streaming platforms by analyzing user trajectories across genres. Central to this approach is the introduction of "pathlets," recurring patterns in users' genre trajectories, which allow for the creation of interpretable trajectory embeddings. The framework demonstrates that pathlet learning can uncover meaningful and diverse listening patterns, offering both qualitative and quantitative insights into music consumption behavior. By releasing a dataset of 2000 user histories tagged by genre over 17 months, the work advances our understanding of dynamic user preferences and provides new opportunities for researching user behavior and promoting diversity in recommender systems.

(10) Md Ashaduzzaman and Chun-Hua Tsai in "***Leveraging Generative AI to Improve Comprehensibility in Social Recommender Systems***" describe a study investigating how users comprehend AI-generated explanations provided by LLM-powered chatbots in recommender systems, which play a pivotal role in enhancing explainability, trust, and user engagement. While users often perceive these personalized explanations as helpful, actual comprehension can be limited. Through a three-phase user study involving 17 participants, the research analyzes comprehension levels, identifies gaps in understanding, and explores behavioral differences between users who effectively grasp AI explanations and those who do not. The findings reveal that the clarity and usefulness of chatbot explanations depend on users' willingness to ask follow-up questions.

3 Workshop Organization

The ExUM workshop is organized as a half-day event, consisting of two or three blocks of 1.5 hours. We welcome different types of submissions. We encourage the submission of original contributions, position papers, and demo papers, investigating or showcasing novel methodologies to exploit heterogeneous personal data and approaches to build transparent and scrutable user models. We will host a great Invited Talk provided by Mauro Martino, a Principal Scientist at the MIT-IBM Watson AI Lab. He founded the Visual Artificial Intelligence Lab, and he is Professor of Practice at Northeastern University in Boston. His previous roles include Research Scientist at the Senseable City Lab, collaborating with Carlo Ratti, and Assistant Research Professor at Northeastern University, working with Albert-László Barabási at the Center for Complex Network Research and David Lazer at the Institute for Quantitative Social Science (IQSS) at Harvard University. Throughout his career, Martino has received numerous accolades, including the Gold Medal at the 2017 Vizzies Visualization Challenge by the National Science Foundation, multiple Webby Awards, the Innovation by Design Award by Fast Company, and the Information is Beautiful Award. His viral art project, AIPortraits (2018/2019), was the first generative GAN-based application to achieve global recognition, attracting over 8 million unique users per day. His inspiring talk is titled: *"Illuminating AI Mechanics: A Visual Overview of LLM Structures"*.

Acknowledgments

The research is partially funded by the European Union - Next Generation EU: NRRP Initiative, Mission 4, Component 2, Investment 1.3 – Partnerships extended to universities, research centers, companies, and research D.D. MUR n. 341 del 15.03.2022 – Next Generation EU (PE0000013 – "Future Artificial Intelligence Research – FAIR" - CUP: H97G22000210007).