QUARE: 2nd Workshop on Measuring the Quality of Explanations in Recommender Systems

Oana Inel
University of Zurich
Zurich, Switzerland
inel@ifi.uzh.ch

Nicolas Mattis
Vrije Universiteit Amsterdam
Amsterdam, The Netherlands
n.m.mattis@vu.nl

Milda Norkute
Thomson Reuters
Zurich, Switzerland
Milda.Norkute@thomsonreuters.com

Alessandro Piscopo
BBC, Product Group
London, United Kingdom
alessandro.piscopo@bbc.co.uk

Timothée Schmude
University of Vienna
Vienna, Austria
timothee.schmude@univie.ac.at

Sanne Vrijenhoek
University of Amsterdam
Amsterdam, the Netherlands
s.vrijenhoek@uva.nl

Krisztian Balog
Google
Stavanger, Norway
krisztianb@google.com

ABSTRACT

QUARE[1]—measuring the QUality of explAnations in REcommender systems—is the second workshop which focuses on evaluation methodologies for explanations in recommender systems. We bring together researchers and practitioners from academia and industry to facilitate discussions about the main issues and best practices in the respective areas, identify possible synergies, and outline priorities regarding future research directions. Additionally, we want to stimulate reflections around methods to systematically and holistically assess explanation approaches, impact, and goals, at the interplay between organisational and human values. To that end, this workshop aims to co-create a research agenda for evaluating the quality of explanations for recommender systems.

CCS CONCEPTS

• **Information systems** → **Recommender systems**; • **Human-centered computing** → **HCI design and evaluation methods**;

KEYWORDS

Recommender systems, explanation quality, explanation goals, explanation evaluation

[1]In Latin, *why, from what cause.*

1 INTRODUCTION AND MOTIVATION

Recommendations are ubiquitous in many contexts and domains due to a continuously growing adoption of decision-support systems. However, they often face challenges in preserving the users' privacy and autonomy, as these systems might "shape user preferences and guide choices, both individually and socially" [6], while remaining opaque to users. Explanations may be provided along with recommendations to mitigate these issues and to show the reasoning behind the suggestion of a particular item [3], with the design, modality, and purpose of the explanation depending on stakeholders' needs and interests [5]. Thus, explanations can increase the acceptance of recommendations and improve their transparency, effectiveness, trust, and satisfaction among users [12, 13]. However, despite a vast amount of research [14], explanations evaluation is still an area where significant gaps remain. For example, whereas there is no one-size-fits-all explanation [10]—as explainability needs depend on the user goals—there is also no established way to evaluate the quality of "personalised" explanations as of yet. Furthermore, although explanations may significantly affect a user's decision-making process, serving several different goals [12], their effects on different groups of stakeholders have not been investigated in depth yet [2, 4]. This makes it difficult to compare different explanations and adopt this practice more widely within industry and academia.

Recommender systems in industry settings are typically deployed in *multi-stakeholder* environments, which take into consideration different categories of agents [7], with various, if not contrasting, interests [1]. For instance, engagement maximisation purposes may drive a commercial media platform to focus more on persuasive and efficient explanations, thus increasing the probability of users to select the recommendations. However, end-users of a recommender system may be bearers of different values, and explanations can affect them differently [8]. If end-users value transparency and trust, persuasive explanations may discourage them from consuming the recommendations. Therefore, organisations willing to adopt explanations might have insufficient resources to

run tests for every use case, which might lead to the development of *ad hoc* explanation approaches. Consequently, the lack of established, generally applicable, and actionable methodologies available to evaluate explanations of recommendations may result in i) complication of cross-comparison between explanation approaches and ii) delayed adoption of explanation practices in industry settings.

Insights on explanations from the social sciences suggest that explanations in human interactions possess particular qualities [9]. More precisely, explanations may lead to knowledge transfer and interactions, which resemble a conversation (i.e., social). Furthermore, people may find relevant or sufficient a smaller number of reasons for interactions, even when a greater number of reasons is available (i.e., selected). These two points are an additional suggestion that explanations are highly individual. For example, explanations for a decision-maker may differ from those for domain experts or laypeople. Therefore, given the multitude of personal characteristics that can inform the decision regarding the effectiveness of an explanation, it is crucial to adopt user-centred evaluation approaches. Nevertheless, such approaches still pose challenges when faced with the heterogeneity of the potential stakeholders of recommender systems. Indeed, as Vultureanu-Albişi and Bădică [13] note, explainability is "domain-dependent and cannot be defined independently of a specific domain", goal, or target group. However, the specific requirements remain to be established.

In planning this workshop, we build on the lessons learned from the first edition, co-located with SIGIR 2022 (https://sites.google.com/view/quare-2022/) [11]. During the first edition, we identified the need for measuring the quality of explanations for recommender systems from multiple perspectives, including 1. user-centric explanation evaluation; 2. design requirements for explanations from actual recommender systems target users; 3. diverse, interdisciplinary, and multi-stakeholder evaluation. Accordingly, this discussion-centric workshop edition aims to *shape and outline a research agenda* for evaluating the quality of explanations for recommender systems. We plan to engage a diverse, multidisciplinary community of academic and industry researchers and practitioners and facilitate the exchange of perspectives and solutions to the challenges and issues aforementioned. In particular, we aim to outline a research agenda for evaluating the quality of explanations in recommender systems by focusing on the following three topics:

- *Generating explanations*: How to balance user needs and organisational objectives when explaining recommendations?
- *Evaluating explanations*: How to evaluate the quality of explanations through both subjective and objective measures?
- *Operationalising explanations*: What could be a typical MVP (Minimum Viable Product) for explanations of recommendations? What are the most basic features we may need?

2 WORKSHOP CONTRIBUTIONS

As part of pre-workshop activities, the organisers will summarise the current state of research on all three topics of interest to the workshop, namely the generation, evaluation, and operationalisation of explanations. At the same time, the workshop organisers will approach potential workshop panelists and invite them to contribute to the background work by preparing lightning talks

on specific aspects of each topic. We ensure diverse and interdisciplinary attendance through targeted calls for participation to specific communities such as HCI researchers, research scientists working on recommender systems, and social science researchers. Besides those communities, we aim to reach out to industry practitioners by sharing calls through the industry networks some of the organisers belong to (e.g., the European Broadcasting Union).

As part of workshop outcomes dissemination activities, we plan to write a series of blog posts summarising the main findings from the lightning and the interactive discussion sessions. The blog posts will be advertised on the workshop's website and shared on social media. More formally, we plan to summarise the discussions and develop the *research agenda for evaluating the quality of explanations for recommender systems* in a technical report (for SIGIR Forum) or white paper. We plan to write this research agenda with the help of all workshop organisers and workshop panelists (see lightning sessions in Section 3) interested in joining the efforts to set the research agenda for evaluating the quality of explanations for recommender systems. This outcome is facilitated by the structure and organisation of our workshop, which foster real-time discussions and collaborative note-taking. Furthermore, we plan to make available for the community as much of the workshop's recordings and discussions as possible, given the consent of the participants and attendees.

3 WORKSHOP ACTIVITIES AND FORMAT

We plan for an interactive, hands-on workshop, which should engage participants not only during the event co-located with RecSys 2023, but also pre- and post-workshop. By doing so, we aim to bring together a community of researchers and practitioners working on explanations in the context of recommender systems and to foster exchanges among them that could help shape and co-create a research agenda for evaluating the quality of explanations for recommender systems.

To facilitate the discussions that form the basis of shaping and co-creating a research agenda for the generation, evaluation, and operationalisation of explanations, the organising team will play an active role pre-workshop. In summary, the following activities are planned:

- We will provide a blueprint for workshop participants that summarises the relevant literature for the three threads and proposes relevant questions to guide the discussion.
- For each of the topics of interest, we will actively invite expert academic and industry researchers and practitioners, who will provide a brief introduction to the topic during the lightning rounds and contribute to keeping the discussion flow during the interactive discussion sessions.

These pre-workshop activities, together with the synergetic format of the workshop, will allow us to collaboratively work on a document that outlines a research agenda.

The workshop will consist of several sessions:

- *Welcome and Introduction* (10 minutes): The workshop organisers welcome the participants and outline the sessions of the day and the planned activities. We also introduce all communication channels, such as the website of the workshop, the Slack channel, and the online collaborative whiteboard

platform Miro. We will encourage workshop participants to introduce themselves on the Slack channel, if they have not done so already.

- *Lightning Talks* (15 minutes x 3): we plan for three short plenary lightning talks from leading researchers and practitioners in each of the three areas we outlined for shaping the research agenda. The lightning talks are intended to stimulate interaction in the discussion sessions that will be the core sessions of the day.
- *Interactive Discussion Sessions* (60 minutes): each lightning talk session is paired with an interactive discussion session. Each such session will discuss one of the three topics that we identified for shaping the research agenda for evaluating the quality of explanations for recommender systems. The discussions will be facilitated by at least one workshop organiser and by the three presenters in each of the lightning rounds. The notes taken during these sessions will be made available online, on the workshop's website. We plan to use an online collaborative whiteboard platform (Miro) to encourage all participants to take notes collaboratively.
- *Final Discussion* (60 minutes): A plenary session with all workshop attendees, summarising the main points discussed in the breakout groups. We will invite representatives of all lightning talks and associated interactive discussion sessions to summarise the main points discussed and findings in a short and concise presentation. We will use this time to discuss emerging ideas and reactions to points raised during the interactive discussion sessions.
- *Closing & Wrap* (5 minutes): Concluding thoughts and future plans.
- *Breaks and social breakout rooms* (10 minutes): We plan for several breaks in between the sessions of the workshop. The breaks are aimed at informal networking. A dedicated Slack channel where all participants can socialise during breaks will be made available.

We plan for extensive asynchronous communication and interaction avenues, using tools such as 1.) the website of the workshop as the main communication channel in terms of activities and organisation; 2.) a Slack channel for community building, sharing of relevant information among participants and organisers; 3.) an online platform such as Zoom for organising the workshop. Furthermore, we will use online collaborative whiteboard platforms, such as Miro, to facilitate collaboration and discussion among participants, as well as collaborative note taking. We plan to include accessibility requirements such as live transcripts and be mindful of the various constraints our prospective participants may have.

4 WORKSHOP ORGANISERS

- Oana Inel - Postdoctoral Researcher at the University of Zurich.
- Nicolas Mattis - PhD candidate at the Vrije Universiteit Amsterdam, Department of Communication Science.
- Milda Norkute - Lead Designer at Thomson Reuters Labs.
- Alessandro Piscopo - Lead Data Scientist at the BBC.
- Timothée Schmude - PhD Candidate at the University of Vienna.

- Sanne Vrijenhoek - PhD Candidate at the University of Amsterdam's Institute of Information Law.
- Krisztian Balog - Staff Research Scientist at Google and a Full Professor at the University of Stavanger.

REFERENCES

[1] Himan Abdollahpouri, Gediminas Adomavicius, Robin Burke, Ido Guy, Dietmar Jannach, Toshihiro Kamishima, Jan Krasnodebski, and Luiz Augusto Pizzato. 2020. Multistakeholder recommendation: Survey and research directions. *User Model. User Adapt. Interact.* 30, 1 (2020), 127–158.

[2] Krisztian Balog and Filip Radlinski. 2020. Measuring Recommendation Explanation Quality: The Conflicting Goals of Explanations. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20)*. ACM, 329–338.

[3] Deepesh V. Hada, Vijaikumar M, and Shirish K. Shevade. 2021. ReXPlug: Explainable Recommendation using Plug-and-Play Language Model. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '21)*. ACM, 81–91.

[4] Chen He, Denis Parra, and Katrien Verbert. 2016. Interactive Recommender Systems: A survey of the State of the Art and Future Research Challenges and Opportunities. *Expert Syst. Appl.* 56 (2016), 9–27.

[5] Markus Langer, Daniel Oster, Timo Speith, Holger Hermanns, Lena Kästner, Eva Schmidt, Andreas Sesing, and Kevin Baum. 2021. What do we want from Explainable Artificial Intelligence (XAI)? – A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. *Artificial Intelligence* 296 (July 2021), 103473.

[6] Silvia Milano, Mariarosaria Taddeo, and Luciano Floridi. 2020. Recommender systems and their ethical challenges. *AI & SOCIETY* 35, 4 (01 Dec 2020), 957–967.

[7] Silvia Milano, Mariarosaria Taddeo, and Luciano Floridi. 2021. Ethical aspects of multi-stakeholder recommendation systems. *Inf. Soc.* 37, 1 (2021), 35–45.

[8] Martijn Millecamp, Cristina Conati, and Katrien Verbert. 2022. "Knowing Me, Knowing You": Personalized Explanations for a Music Recommender System. *User Model. User-adapt. Interact.* 32 (2022), 215–252.

[9] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence* 267 (Feb. 2019), 1–38.

[10] Ingrid Nunes and Dietmar Jannach. 2017. A systematic review and taxonomy of explanations in decision support and recommender systems. *User Model. User-adapt. Interact.* 27, 3-5 (2017), 393–444.

[11] Alessandro Piscopo, Oana Inel, Sanne Vrijenhoek, Martijn Millecamp, and Krisztian Balog. 2022. Report on the 1st Workshop on Measuring the Quality of Explanations in Recommender Systems (QUARE 2022) at SIGIR 2022. *SIGIR Forum* 56, 2, Article 11 (dec 2022), 16 pages.

[12] Nava Tintarev and Judith Masthoff. 2015. Explaining Recommendations: Design and Evaluation. In *Recommender Systems Handbook*. Springer, 353–382.

[13] Alexandra Vultureanu-Albişi and Costin Bădică. 2022. A survey on effects of adding explanations to recommender systems. *Concurrency and Computation: Practice and Experience* 34, 20 (2022), e6834.

[14] Yongfeng Zhang and Xu Chen. 2020. Explainable Recommendation: A Survey and New Perspectives. *Found. Trends Inf. Retr.* 14, 1 (2020), 1–101.