Philipp Hager

p.k.hager@uva.nl
[web] [github] [linkedin] [scholar]

Last updated on July 9th, 2024

Research interests: I am interested in leveraging complex user feedback to learn and evaluate new AI systems, particularly for search and recommendation.
Keywords: Off-policy learning, offline evaluation, unbiased learning-to-rank, click modeling.

Education

01/2022 - 01/2026 (planned graduation)	**PhD Candidate, University of Amsterdam** Supervisors: Prof. Dr. Maarten de Rijke and Dr. Onno Zoeter Member of the Mercury Machine Learning Lab with Booking.com and TU Delft
04/2017 - 09/2020	**M.Sc. IT-Systems Engineering, HPI Unviersity of Potsdam** Thesis: Multi-faceted domain-specific document embeddings (1.2, cum laude)
08/2013 - 04/2017	**B.Sc. Media Informatics, University of Applied Sciences Düsseldorf** Thesis: App-based detection and analysis of security and privacy concerns introduced into Android apps by third-party libraries (1.3)

Professional Experience

01/2021 - 01/2022	**Research Assistant, University of Southern Denmark - Odense** Part-time research work on mainstream bias in recommender systems. Supervisor: Dr. Pantelis P. Analytis
07/2020 - 12/2021	**Data Scientist L2, Blinkist - Berlin** Built production systems for recommending multilingual audio and textual content for over 18M users in real time. Algorithms included transformer-based dense retrieval and autoencoders implemented using a.o. Tensorflow, Serverless, DynamoDB, Docker, and FastAPI **[reference letter]**.
08/2019 - 06/2020	**Working Student Data Science, Blinkist - Berlin** Multi-lingual content-based book recommendation using transformer models.
09/2018 - 01/2019	**Software Development Engineering Intern, Amazon - Madrid** Data analytics for the EU leadership of Amazon Fashion.
09/2015 - 09/2018	**Working Student Android Development, Blinkist - Berlin** Early development of the Blinkist Android app using reactive programming and Kotlin.

Publications

2024	**P. Hager***, R. Deffayet*, JM. Renders, O. Zoeter, M. de Rijke. **Unbiased Learning to Rank Meets Reality: Lessons from Baidu's Large-Scale Search Dataset.** Accepted at the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval.
2023	P. Analytis*, **P. Hager*.** **Collaborative filtering algorithms are prone to mainstream-taste bias.** In In Proceedings of the 17th ACM Conference on Recommender Systems.

R. Deffayet*, **P. Hager***, JM. Renders, M. de Rijke. **An Offline Metric for the Debiasedness of Click Models.** In Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval.

P. Hager, M. de Rijke, and O. Zoeter. **Contrasting Neural Click Models and Pointwise IPS Rankers.** In Advances in Information Retrieval: 45th European Conference on Information Retrieval.

2022 **P. Hager,** M. de Rijke, and O. Zoeter. **Are Neural Click Models Pointwise IPS Rankers?** The CONSEQUENCES+REVEAL Workshop co-located with the 16th ACM Conference on Recommender Systems.

2021 **J. Risch**, P. Hager, R. Krestel. **Multifaceted Domain-Specific Document Embeddings.** In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Demonstrations.

Tutorials & Talks

2024 [Lecture] **P. Hager. Improvements That Add Up: An opinionated rant on reproducibility and progress in IR.** Guest lecture at the 15th European Summer School on Information Retrieval.

[Tutorial] S. Gupta, **P. Hager**, J. Huang, A. Vardasbi, H. Oosterhuis. **Unbiased Learning to Rank: On Recent Advances and Practical Applications.** Tutorial at the 17th ACM International Conference on Web Search and Data Mining.

2023 [Tutorial] S. Gupta, **P. Hager**, H. Oosterhuis. **Recent Advancements in Unbiased Learning to Rank.** Tutorial at the Forum for Information Retrieval Evaluation 2023.

[Talk] **P. Hager,** R. Deffayet, JM. Renders, M. de Rijke. **An Offline Metric for the Debiasedness of Click Models.** Talk at the 21st Dutch-Belgian Information Retrieval Workshop.

[Talk] **P. Hager. When Metrics Break Down - On Evaluating User Models from Clicks.** Invited talk at ICAI: The Labs - Machine Learning in the service industry.

[Lecture] **P. Hager. Learning to Rank.** Search Engines Course, University of Amsterdam.

[Tutorial] S. Gupta, **P. Hager**, J. Huang, A. Vardasbi, H. Oosterhuis. **Recent Advances in the Foundations and Applications of Unbiased Learning to Rank.** In Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval.

[Talk] **P. Hager**, M. De Rijke. **A Brief Tutorial on Supervised Learning to Rank.** Booking.com.

2021 [Lecture] **P. Hager. NLP in Production - A Content-based Recommender System Case Study.** Data Science Course, University of Southern Denmark.

Teaching & Supervison

2024 **Morris De Haan**, B.Sc. Computer Science, University of Amsterdam
Thesis: Understanding the Effects of the Logging Policy on Two Tower Models.

2023 **Search Engines Course**, B.Sc. Artificial Intelligence, University of Amsterdam
Cedrik Blommestijn, B.Sc. Computer Science, University of Amsterdam
Thesis: Bridging the gap between large language models and traditional learning-to-rank.

2022 **Search Engines Course**, B.Sc. Artificial Intelligence, University of Amsterdam

2014 - 2015 **Database Systems I & II**, B.Sc. Media Informatics, University of Applied Sciences Düsseldorf

Activities

2024	**Search Engines Amsterdam, University of Amsterdam** Organizer of the monthly SEA meetup of the IRLab Amsterdam.
2023 - now	**PhD Council Member, University of Amsterdam** Student representative in the PhD council of the informatics institute, which tackles academic, social, and cultural issues affecting PhD students.
2023	**Research Meeting Chair, University of Amsterdam** Organizer of the weekly group meetings of the IRLab Amsterdam.
2022 - now	**Mentor at Inclusive AI, University of Amsterdam** IAI is an inclusive space for students to get non-academic help from senior peers in the field and connect with people of a similar background.
2014 - 2020	**Scholar of the German Academic Scholarship Foundation** Merit-based scholarship for outstanding academic achievements.

Reviewing

2024	**SIGIR**, **RecSys, CIKM**
2023	**CIKM**, **ML Reproducibility Challenge** (outstanding reviewer award), **ICTIR** (student PC)
2022	**SIGIR** (subreviewer), **CIKM** (subreviewer)

Skills

Programming	Python, Java, Kotlin, SQL
Tools	PyTorch, NumPy, Jax, Scikit-Learn, Hydra, PySpark, AWS
Languages	English, German

Open Source

2021 **Faceted domain encoder**
Learn multifaceted document embeddings for domain-specific text.

2019 **Kafka SALSA**
Implemtation of Twitter's GraphJet recommender system using Kafka Streams.

2016 **Disclosure Android app**
Android app for detecting privacy-infringing third-party code in other apps.