Timo Breuer

breuert.github.io Google Scholar 0000-0002-1765-2449 timobreuer@acm.org

Monographs

Reproducibility in Information Retrieval, *Maria Maistro, Timo Breuer, Nicola Ferro*, FNTIR 2026

Reproducible Information Retrieval Research: From Principled System-Oriented Evaluations Towards User-Oriented Experimentation, *Timo Breuer*, DuePublico 2023

Journal articles

Now that your System has been Reproduced, What does this Mean for the Users?, *Angelo Geninatti Cossatin, Timo Breuer, Noemi Mauro, Maria Maistro*, TOIS 2026

Large Language Models for Information Retrieval: Challenges and Chances, *Timo Breuer, Sameh Frihat, Norbert Fuhr, Dirk Lewandowski, Philipp Schaer, Ralf Schenkel*, Datenbank-Spektrum 2025

Validating Synthetic Usage Data in Living Lab Environments, *Timo Breuer, Norbert Fuhr, Philipp Schaer*, JDIQ 2024

An in-depth investigation on the behavior of measures to quantify reproducibility, *Maria Maistro, Timo Breuer, Philipp Schaer, Nicola Ferro*, IPM 2023

Relevance assessments, bibliometrics, and altmetrics: a quantitative study on PubMed and arXiv, *Timo Breuer, Dirk Tunger, Philipp Schaer*, Scientometrics 2022

Evaluation Infrastructures for Academic Shared Tasks, *Johann Schaible, Timo Breuer, Narges Tavakolpoursaleh, Bernd Müller, Benjamin Wolff, Philipp Schaer*, Datenbank-Spektrum 2020

Conference proceedings

Formalized Information Needs Improve Large-Language-Model Relevance Judgments, *Jüri Keller, Maik Fröbe, Björn Engelmann, Fabian Haak, Timo Breuer, Birger Larsen, Philipp Schaer*, SIGIR 2026

Sim4IA-Bench: A User Simulation Benchmark Suite for Next Query and Utterance Prediction, *Andreas Konstantin Kruff, Christin Katharina Kreutz, Timo Breuer, Philipp Schaer, Krisztian Balog*, ECIR 2026

"We Share Our Code Online": Why This Is Not Enough to Ensure Reproducibility and Progress in Recommender Systems Research, *Faisal Shehzad, Timo Breuer, Maria Maistro, Dietmar Jannach*, RecSys 2025

Evaluating Contrastive Feedback for Effective User Simulations, *Andreas Konstantin Kruff, Timo Breuer, Philipp Schaer*, SIGIR 2025

Data Fusion of Synthetic Query Variants With Generative Large Language Models, *Timo Breuer*, SIGIR-AP 2024

SimIIR 3: A Framework for the Simulation of Interactive and Conversational Information Retrieval, *Leif Azzopardi, Timo Breuer, Björn Engelmann, Christin Kreutz, Sean MacAvaney, David Maxwell, Andrew Parry, Adam Roegiest, Xi Wang, Saber Zerhoudi*, SIGIR-AP 2024

Browsing and Searching Metadata of TREC, *Timo Breuer, Ellen M. Voorhees, Ian Soboroff*, SIGIR 2024

Toward Evaluating the Reproducibility of Information Retrieval Systems with Simulated Users, *Timo Breuer and Maria Maistro*, ACM REP 2024

Evaluation of Temporal Change in IR Test Collections, *Jüri Keller, Timo Breuer, Philipp Schaer*, ICTIR 2024

Replicability Measures for Longitudinal Information Retrieval Evaluation, *Jüri Keller, Timo Breuer, Philipp Schaer*, CLEF 2024,

Context-Driven Interactive Query Simulations Based on Generative Large Language Models, *Björn Engelmann, Timo Breuer, Jana Isabelle Friese, Philipp Schaer, Norbert Fuhr*, ECIR 2024

Simulating Users in Interactive Web Table Retrieval, *Björn Engelmann, Timo Breuer, Philipp Schaer*, CIKM 2023

Bibliometric Data Fusion for Biomedical Information Retrieval, *Timo Breuer, Christin Katharina Kreutz, Philipp Schaer, Dirk Tunger*, JCDL 2023

`ir_metadata`**: An Extensible Metadata Schema for Information Retrieval Experiments**, *Timo Breuer, Jüri Keller, Philipp Schaer*, SIGIR 2022

Validating Simulations of User Query Variants, *Timo Breuer, Norbert Fuhr, Philipp Schaer*, ECIR 2022

Evaluating Elements of Web-Based Data Enrichment for Pseudo-relevance Feedback Retrieval, *Timo Breuer, Melanie Pest, Philipp Schaer*, CLEF 2021

A Living Lab Architecture for Reproducible Shared Task Experimentation, *Timo Breuer and Philipp Schaer*, ISI 2021

`repro_eval`**: A Python Interface to Reproducibility Measures of System-Oriented IR Experiments**, *Timo Breuer, Nicola Ferro, Maria Maistro, Philipp Schaer*, ECIR 2021

How to Measure the Reproducibility of System-oriented IR Experiments, *Timo Breuer, Nicola Ferro, Norbert Fuhr, Maria Maistro, Tetsuya Sakai, Philipp Schaer, Ian Soboroff*, SIGIR 2020

Reproducible Online Search Experiments, *Timo Breuer*, ECIR 2020

Workshop articles and activities

Evaluating Information Retrieval Models Along Time: The LongEval Lab at CLEF 2026, *Timo Breuer, Matteo Cancellieri, Alaa El-Ebshihy, Maik Fröbe, Petra Galuscáková, Lorraine Goeuriot, Gabriel Iturra-Bocaz, Jüri Keller, Petr Knoth, Andreas Konstantin Kruff, Philippe Mulhem, Florina Piroi, David Pride, Philipp Schaer, Didier Schwab*, ECIR 2026

Conversational Agents: A Framework for Evaluation (CAFE), *Christine Bauer, Li Chen, Nicola Ferro, Norbert Fuhr, Avishek Anand, Timo Breuer, Guglielmo Faggioli, Ophir Frieder, Hideo Joho, Jussi Karlgren, Johannes Kiesel, Bart P. Knijnenburg, Aldo Lipani, Lien Michiels, Andrea Papenmeier, Maria Soledad Pera, Mark Sanderson, Scott Sanner, Benno Stein, Johanne R. Trippas, Karin Verspoor, Martijn C. Willemsen*, Dagstuhl Perspectives Workshop 24352

User Simulation in Practice: Lessons Learned from Three Shared Tasks, *Marcel Gohsen, Zahra Abbasiantaeb, Mohammad Aliannejadi, Krisztian Balog, Timo Breuer, Jeffrey Dalton, Maik Fröbe, Christin Katharina Kreutz, Andreas Kruff, Simon Lupart, Nailia Mirzakhmedova, Harrisen Scells, Philipp Schaer, Benno Stein, Johannes Kiesel*, SIGIR Forum 2025

Second SIGIR Workshop on Simulations for Information Access (Sim4IA 2025), *Philipp Schaer, Christin Katharina Kreutz, Krisztian Balog, Timo Breuer, Andreas Konstantin Kruff*, SIGIR 2025

Report on the Workshop on Simulations for Information Access (Sim4IA 2024) at SIGIR 2024, *Timo Breuer, Christin Katharina Kreutz, Norbert Fuhr, Krisztian Balog, Philipp Schaer et al.*, Sim4IA, SIGIR Forum

SIGIR 2024 Workshop on Simulations for Information Access (Sim4IA 2024), *Philipp Schaer, Christin Katharina Kreutz, Krisztian Balog, Timo Breuer, Norbert Fuhr*, Sim4IA, SIGIR 2024

The Two Sides of the Coin: Hallucination Generation and Detection with LLMs as Evaluators for LLMs, *Anh Thu Maria Bui, Saskia Felizitas Brech, Natalie Hußfeldt, Tobias Jennert, Melanie Ullrich, Timo Breuer, Narjes Nikzad Khasmakhi, Philipp Schaer*, ELOQUENT, CLEF 2024

Leveraging Prior Relevance Signals in Web Search, *Jüri Keller, Timo Breuer, Philipp Schaer*, LongEval, CLEF 2024

Evaluating Temporal Persistence Using Replicability Measures, *Jüri Keller, Timo Breuer, Philipp Schaer*, LongEval, CLEF 2023

Overview of LiLAS 2021 - Living Labs for Academic Search, *Philipp Schaer, Timo Breuer, Leyla Jael Castro, Benjamin Wolff, Johann Schaible, Narges Tavakolpoursaleh*, LiLAS, CLEF 2021

Relations Between Relevance Assessments, Bibliometrics and Altmetrics, *Timo Breuer, Dirk Tunger, Philipp Schaer*, BIR, ECIR 2020

- **STELLA: Towards a Framework for the Reproducibility of Online Search Experiments**, *Timo Breuer, Philipp Schaer, Narges Tavakolpoursaleh, Johann Schaible, Benjamin Wolff, Bernd Müller*, OSIRRC, SIGIR 2019
- **Dockerizing Automatic Routing Runs for The Open-Source IR Replicability Challenge (OSIRRC 2019)**, *Timo Breuer and Philipp Schaer*, OSIRRC, SIGIR 2019
- **Replicability and Reproducibility of Automatic Routing Runs**, *Timo Breuer and Philipp Schaer*, CENTRE, CLEF 2019