

Castells, P., Hopfgartner, F., Said, A., & Lalmas, M.

# Workshop on benchmarking adaptive retrieval and recommender systems

**Conference paper | Accepted manuscript (Postprint)**

This version is available at <https://doi.org/10.14279/depositononce-6800>



© Owner/Author | ACM, 2013. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval - SIGIR '13. ACM Press. <https://doi.org/10.1145/2484028.2484224>.

Castells, P., Hopfgartner, F., Said, A., & Lalmas, M. (2013). Workshop on benchmarking adaptive retrieval and recommender systems. In Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval - SIGIR '13. ACM Press. <https://doi.org/10.1145/2484028.2484224>

## Terms of Use

Copyright applies. A non-exclusive, non-transferable and limited right to use is granted. This document is intended solely for personal, non-commercial use.

**WISSEN IM ZENTRUM**  
**UNIVERSITÄTSBIBLIOTHEK**

Technische  
Universität  
Berlin

# Workshop on Benchmarking Adaptive Retrieval and Recommender Systems – BARS 2013

Pablo Castells  
Univ. Autónoma de Madrid  
Fco. Tomás y Valiente 11  
28049 Madrid, Spain  
pablo.castells@uam.es

Frank Hopfgartner  
Technische Univ. Berlin  
Ernst-Reuter-Platz 7  
10587 Berlin, Germany  
frank.hopfgartner@tu-berlin.de

Alan Said  
Cent. Wiskunde & Informatica  
Science Park 123, 1098 XG  
Amsterdam, Netherlands  
alan.said@cwi.nl

Mounia Lalmas  
Yahoo! Labs Barcelona  
Avinguda Diagonal 177  
08018 Barcelona, Spain  
mounia@acm.org

## ABSTRACT

Evaluating adaptive and personalized information retrieval techniques is known to be a difficult endeavor. The rapid evolution of novel technologies in this scope raises additional challenges that further stress the need for new evaluation approaches and methodologies. The BARS 2013 workshop seeks to provide a specific venue for work on novel, personalization-centric benchmarking approaches to evaluate adaptive retrieval and recommender systems.

## Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: *search process, information filtering.*

## Keywords

Evaluation, adaptive information retrieval, recommender systems, benchmarking, metrics, methodology.

## 1. INTRODUCTION

Great progress has been made in recent years in the development of recommendation, retrieval and personalization techniques. Yet the evaluation of these systems is still based on traditional metrics, e.g. precision, recall or RMSE, often not taking the use-case and situation of the system into consideration, and failing to provide a suitable proxy of user satisfaction and business goals. Moreover, the rapid evolution of novel information retrieval (IR) and recommender systems foster the need for new evaluation paradigms.

New evaluation approaches of adaptive systems should evaluate both functional and non-functional requirements. Functional requirements go beyond traditional relevance metrics and focus on user-centered utility metrics, such as novelty, diversity and serendipity. Non-functional requirements focus on performance and technical aspects, e.g. scalability and reactivity.

The evaluation of adaptive IR systems has been acknowledged to find difficulty in fitting in established evaluation paradigms and methodologies, which can be identified as a hurdle to research and progress in this area. Active research efforts and open discussion

are currently taking place in parallel in the Recommender Systems and Adaptive IR fields, where devising methodologies and metrics suiting the goals and task models of real applications is still a prominent open issue.

The BARS 2013 workshop aimed to serve as a venue for work on novel, personalization-centric benchmarking approaches to evaluate adaptive retrieval and recommender systems. The workshop was set to revise and leverage the latest advances in this area, identify the main issues to be addressed, and share ideas for continued progress. BARS sought, in particular, to join forces and provide a meeting point for researchers working on largely overlapping and connected areas such as adaptive IR and recommender systems, dealing with closely related problems but often from different backgrounds.

## 2. SCOPE

The workshop gathered researchers and practitioners interested in developing better, clearer, and/or more complete evaluation methodologies, and addressing the challenges involved in the evaluation of adaptive retrieval and recommender systems. It provided an informal setting for exchanging and discussing ideas, sharing experiences and viewpoints. The participants worked together in the identification and better understanding of the current gaps in the evaluation methodologies, seeking to lay directions for progress in addressing them, and the consolidation and convergence of experimental methods and practice.

The accepted papers and the discussions held at the workshop addressed, among others, the following topics:

- New metrics and methods for quality estimation of recommender and adaptive IR systems.
- Novel frameworks for the user-centric evaluation of adaptive systems.
- Validation of off-line methods with online studies.
- Comparison of evaluation metrics and methods.
- Comparison of recommender and IR approaches across multiple systems and domains.
- Measuring technical constraints vs. accuracy.
- New datasets for the evaluation of recommender and adaptive IR systems.
- Benchmarking frameworks.
- Multiple-objective benchmarking.

The accepted papers and a summary of discussions are available in the workshop proceedings, which can be reached from the workshop website at <http://www.bars-workshop.org>.