

# Bibliometric-enhanced Information Retrieval (BIR) 2014



## Workshop at ECIR 2014

13 April 2014

You are invited to participate in the upcoming workshop on **Bibliometric-enhanced Information Retrieval (BIR)**, to be held as part of the [36th European Conference on Information Retrieval \(ECIR\)](http://ceur-ws.org/Vol-1143/). All workshop papers have been published in the CEUR Workshop Proceedings. <http://ceur-ws.org/Vol-1143/>

## Workshop program

Half day workshop from 9:00 am - 12:30 pm (3,5 h)

- Introduction (10 min) [Presentation](#)
- Block A
  - Marc Bertin and Iana Atanassova: **A Study of Lexical Distribution in Citation Contexts through the IMRaD Standard** (10 min) [Paper](#) | [Presentation](#)
  - Nees Jan van Eck and Ludo Waltman: **Systematic retrieval of scientific literature based on citation relations: Introducing the CitNetExplorer tool** (10 min) [Paper](#) | [Presentation](#)
  - Group discussion (30 min)
- Block B
  - Muhammad Kamran Abbasi and Ingo Frommholz: **Exploiting Information Needs and Bibliographics for Polyrepresentative Document Clustering** (10 min) [Paper](#) | [Presentation](#)
  - Haozhen Zhao and Xiaohua Hu: **Language Model Document Priors based on Citation and Co-citation Analysis** (10 min) [Paper](#) | [Presentation](#)
  - Group discussion (30 min)
- Break (30 min)

- Block C
  - Zeljko Carevic and Philipp Schaer: **On the Connection Between Citation-based and Topical Relevance Ranking: Results of a Pretest using iSearch** (10 min) [Paper](#) | [Presentation](#)
  - Kris Jack, Pablo López-García, Maya Hristakeva and Roman Kern: **{{citation needed}}**: **Filling in Wikipedia's Citation Shaped Holes** (10 min) [Paper](#) | [Presentation](#)
  - Group discussion (30 min)
- Conclusions, feedback and further steps (15 min)

## Important Dates

- Submissions: 14 February 2014
- Notification: 28 February 2014
- Camera Ready Contributions: 14 March 2014
- Workshop: 13 April 2014 in Amsterdam (NL)

## Introduction

Bibliometric techniques are not yet widely used to enhance retrieval processes in digital libraries, although they offer value-added effects for users. In this workshop we will explore how statistical modelling of scholarship, such as Bradfordizing or network analysis of coauthorship network, can improve retrieval services for specific communities, as well as for large, cross-domain collections. This workshop aims to raise awareness of the missing link between information retrieval (IR) and bibliometrics/scientometrics and to create a common ground for the incorporation of bibliometric-enhanced services into retrieval at the digital library interface.

## Aim of the Workshop

In this workshop we aim to engage with the IR community about possible links to bibliometrics and complex network theory which also explores networks of scholarly communication. Bibliometric techniques are not yet widely used to enhance retrieval processes in digital libraries, yet they offer value-added effects for users. Our interests include information retrieval, information seeking, science modelling, network analysis, and digital libraries. The goal is to apply insights from bibliometrics, scientometrics, and informetrics to concrete practical problems of information retrieval and browsing.

Retrieval evaluations have shown that simple text-based retrieval methods scale up well but do not progress. Traditional retrieval has reached a high level in terms of measures like precision and recall, but scientists and scholars still face challenges present since the early days of digital libraries: mismatches between search terms and indexing terms, overload from result sets that are too large and complex, and the drawbacks of text-based relevance rankings. Therefore we will focus on statistical modelling and corresponding visualizations of the evolving science system. Such analyses have revealed not only the fundamental laws of Bradford and Lotka, but also network structures and dynamic mechanisms in scientific production. Statistical models of scholarly activities are increasingly used to evaluate specialties, to forecast and discover research trends, and to shape science policy. Their use as tools in navigating scientific information in public digital libraries is a promising but still

relatively new development. We will explore how statistical modelling of scholarship can improve retrieval services for specific communities, as well as for large, cross-domain collections. Some of these techniques are already used in working systems but not well integrated in larger scholarly IR environments.

The availability of new IR test collections that contain citation and bibliographic information like the iSearch collection or the ACL collection could deliver enough ground to interest (again) the IR community in these kind of bibliographic systems. The long-term research goal is to develop and evaluate new approaches based on informetrics and bibliometrics.

The aim of this workshop is to bring together researchers from different domains, such as information retrieval, information seeking, science modelling, bibliometrics, scientometrics, network analysis, and digital libraries to move toward a deeper understanding of this research challenge.

## **Workshop Topics**

To support the previously described goals the workshop topics include (but are not limited to) the following:

- IR for digital libraries and scientific information portals
- IR for scientific domains, e.g. social sciences, life sciences etc.
- Information Seeking Behaviour
- Bibliometrics, citation analysis and network analysis for IR
- Query expansion and relevance feedback approaches
- Science Modelling (both formal & empirical)
- Task based user modelling, interaction, and personalisation
- (Long-term) Evaluation methods and test collection design
- Collaborative information handling and information sharing
- Classification, categorisation and clustering approaches
- Information extraction (including topic detection, entity and relation extraction)
- Recommendations based on explicit and implicit user feedback

We especially invite descriptions of running projects and ongoing work. Papers that investigate multiple themes directly are especially welcome.

## **Submission Details**

All submissions must be written in English following Springer LNCS author guidelines (4 to 8 pages) and should be submitted as PDF files to EasyChair. All submissions will be reviewed by at least two independent reviewers. Please be aware of the fact that at least one author per paper needs to register for the workshop and attend the workshop to present the work. In case of no-show the paper (even if accepted) will be deleted from the proceedings AND from the program.

Springer LNCS: [www.springer.com/lncs](http://www.springer.com/lncs)

EasyChair: <http://www.easychair.org/conferences/?conf=bir2014>

Workshop proceedings will be deposited online in the CEUR workshop proceedings publication service (ISSN 1613-0073) - This way the proceedings will be permanently available and citable (digital persistent identifiers and long term preservation).

## **Programm Committee**

- Cornelia Caragea, University of North Texas (USA)
- Ingo Frommholz, University of Bedfordshire (UK)
- Norbert Fuhr, University of Duisburg-Essen (Germany)
- Claus-Peter Klas, University of Hagen (Germany)
- Stasa Milojevic, Indiana University (USA)
- Lynda Tamine-Lechani, University Paul Sabatier (France)
- Howard D. White, Drexel University (USA)
- Ed A. Fox, Virginia Tech (USA, to be confirmed)
- C. Lee Giles, Pennsylvania State University (USA, to be confirmed)
- Stephen Robertson, University College London (UK, to be confirmed)
- Henry Small, SciTech Strategies (USA, to be confirmed)
- Simone Teufel, University of Cambridge (UK, to be confirmed)
- Keith van Rijsbergen, University of Glasgow (UK, to be confirmed)

This workshop is also informed by an ongoing COST Action TD1210 [KnowEscape](#).

## **Organizers**

- [Philipp Mayr](#), GESIS - Leibniz Institute for the Social Sciences, Germany
- [Andrea Scharnhorst](#), DANS, Royal Netherlands Academy of Arts and Sciences in Amsterdam, Netherlands
- [Birger Larsen](#), Aalborg University, Denmark
- [Philipp Schaer](#), GESIS - Leibniz Institute for the Social Sciences, Germany
- [Peter Mutschke](#), GESIS - Leibniz Institute for the Social Sciences, Germany