The Information Retrieval Series 33

Advanced Topics in Information Retrieval

Bearbeitet von
Massimo Melucci, Ricardo Baeza-Yates

Weitere Fachgebiete > EDV, Informatik > Datenbanken, Informationssicherheit, Geschäftssoftware > Data Mining, Information Retrieval

schnell und portofrei erhältlich bei

beck-shop.de
DIE FACHBUCHHANDLUNG

Preface

This book presents a wide-spectrum illustration of what research in Information Retrieval has produced or will produce in the next years. The book content includes chapters on Aggregated Search, Digital Advertising, Digital Libraries, Discovery of Spam and Opinions in the Web, Evaluation, Information Retrieval in Context, Multimedia Resource Discovery, Quantum Mechanics applied to Information Retrieval, Scalability Challenges in Web Search Engines, and Users in Interactive Information Retrieval Evaluation. Every chapter is authored by well-known researchers, with an integrated bibliography and subject index. In the following paragraphs we describe briefly the content of each of these chapters, which are sorted in lexicographical order by first author.

We start with a chapter on Digital Libraries by Maristella Agosti from the University of Padova, Italy. Here she gives some historical background to then introduce the main concepts of present digital library systems. After that, the chapter covers usability, interoperability and evaluation issues.

The second chapter is Scalability Challenges in Web Search Engines by Berkant Barla Cambazoglu and Ricardo Baeza-Yates from Yahoo! Research Barcelona, Spain. This chapters looks at current search engine architectures, from a single search server to search clusters. The chapter ends in a hypothetical geographically distributed multi-site search system, focusing on scalability issues and current open problems.

Chapter three is entitled Spam, Opinions, and other Relationships: Towards a Comprehensive View of the Web by Bettina Berendt from Leuven University, Belgium. This chapter proposes an integrating model of learning cycles involving data, information and knowledge, which includes Information Retrieval and Knowledge Discovery in the Web as particular cases. This is illustrated by applying the model to spam detection, opinion mining and relation mining.

The fourth chapter is The User in Interactive Information Retrieval Evaluation by Peter Ingwersen from the Royal School of Library and Information Science, Denmark. This chapter explores interactive information retrieval by using a laboratory research framework for IR. This involves the definition of request types, ultra-light experiments, interactive-light studies and naturalistic field investigations.

Chapter five is entitled Aggregated Search by Mounia Lalmas from Yahoo! Research Barcelona, Spain. This chapter overviews the state of the art in the aggregation of different information sources to achieve diversity for broad or ambiguous queries, exploring the different challenges that arise in this problem.

The sixth chapter is Quantum Mechanics and Information Retrieval by Massimo Melucci and Keith van Rijsbergen, from University of Padova, Italy, and University of Cambridge, UK, respectively. This chapter provides a survey on quantum mechanics applied to IR, with an emphasis on the notation and probability aspects of this new field, which was started by the second author.

Chapter seven is entitled Multimedia Resource Discovery by Stefan Rüger from the Open University, UK. This chapter examines the challenges and opportunities of Multimedia Information Retrieval and its applications. In the case of image search, current techniques covered included piggy-backing text search, automated annotation of visual components, content-based retrieval and fingerprinting to match near duplicates.

The eighth chapter is Information Retrieval in Context by Ian Ruthven from the University of Strathclyde, UK. Here context refers to a complex set of variables describing the intentions and personal characteristics behind the person searching, the data and systems available, and the physical, social and organizational environments. This chapter studies why differences in context can affect how search systems can operate and ways that contextual information can be used to help search systems behave more intelligently to our changing context.

The last chapter is entitled Digital Advertising: An Information Scientists Perspective by James G. Shanahan and Goutham Kurra, USA. In this chapter, a major Internet industry is exposed, providing a detailed overview of the technologies and business models that are transforming the field of online advertising primarily from statistical machine learning and information science perspectives, including the IR problems behind it.

There are many important areas that are covered by more than one chapter. The most significant is evaluation, which is mentioned in every chapter, but in different ways. Despite the long tradition of evaluation in Information Retrieval, the advent of new problems and technologies requires new evaluation methodologies. Personalization and user interfaces are other two common topics, thus highlighting that the interaction between an Information Retrieval system and its users is more crucial than ever. Machine learning and relevance feedback are two other significant topics across the book and complete the general architecture of a system where the interaction between the user's side and the system's side always needs evaluation.

The book is intended to a wide audience of people interested in Information Retrieval: undergraduate and graduate students, post-doctoral researchers, teachers, scholars, as well as industrial researchers.

The Editors thank the authors for their intellectual generosity in writing the chapters of the book. Massimo Melucci thanks his colleagues of the Information Management Systems research group of the Department of Information Engineering at

the University of Padova, while Ricardo Baeza-Yates would like to acknowledge the support of Yahoo! Research. Finally, we would like to thanks Springer, in particular Ralf Gerstner, for publishing this book.

<div style="text-align: right;">

Massimo Melucci
Ricardo Baeza-Yates

</div>