

effort, especially in transitional sections, to make these connections clear.

It is also the volume's incohesiveness that makes this reviewer question the audience for the volume. Each paper seems directed at a specific audience and it is difficult to find a more general audience that would find the text of interest. One might argue that it could be used as an introduction to all the topics presented, but in this reader's opinion the papers are at a level that requires a background of some sophistication, though more in linguistics than in computing. I am at a loss to suggest any course in which the book might be used as a text. I think there are a number of quality papers in this volume; perhaps it can be best used to reference the included papers on an individual basis.

*Brian D. Monahan*

Department of Computer and Information Sciences  
Iona College  
New Rochelle, NY 10801

---

#### INFORMATION RETRIEVAL EXPERIMENT

**Karen Sparck Jones, Editor**

London: Butterworths, 1981, vii+352 pp.  
ISBN 0-408-10648-4; \$69.95

This book is essential reading for anyone interested in information retrieval. It describes in detail how to design and carry out experiments in information retrieval and how to analyze the results. Why should other workers in computational linguistics read this book – aside from the fact that, like all researchers, we are consumers of information retrieval? Because it includes a fascinating discussion of the philosophical and practical problems of performing experiments, tremendously valuable at this point when we are all supposed to be taking part in the new experimental computer science.

This book also provides a guide to the main stream of research in bibliographic information retrieval over the last twenty-five years. In this paradigm, documents are represented by lists of terms contained in the title, the abstract, or, occasionally, the whole text, along with frequency counts. Queries are represented by similar statistics and then documents are located by matching the word list from the query with the lists from the documents. Many refinements are possible: using phrases instead of single words, adding a thesaurus, controlling the indexing language, using feedback from the user. The experiments described in this book have helped to determine which of these strategies are of value and thus have been crucial in the development of the sophisticated systems of today.

There are fifteen chapters written by thirteen different scholars. The fact that Karen Sparck Jones could command the efforts of the best minds in information

retrieval testifies to the esteem in which she is held by this community. In spite of the multiplicity of authors, the book functions as a whole, with each chapter building on the preceding ones. The editor clearly kept a strong hand on every step of the writing process, through all three parts of the book.

The four chapters in Part I are concerned with fundamental issues in the design of experiments. Robertson sets the scene by discussing general methodology for retrieval system testing. Van Rijsbergen focuses on the issues involved in characterizing and measuring retrieval system effectiveness. Belkin looks at the other side of the problem – how to define human information needs and determine whether a system is satisfying them. Tague's chapter, the longest in the book, describes the practical decisions that have to be made at every stage in designing and carrying out an experiment. It could well serve as a handbook for the would-be experimenter.

The second part discusses the problems of applying this methodology in diverse practical situations. Some experiments are directed at evaluating the performance of systems that are actually in operation. Others are essentially laboratory tests designed to investigate alternative methods of retrieval effectiveness. Lancaster writes about the problems of "Evaluation within the environment of an operating information service". The system he is talking about is a largely manual one. Barraclough is trying to solve the same problems but in an automated, largely interactive system. Keen tackles the question of designing laboratory tests for manual systems, while Oddy looks at the same question for fully automatic systems. Heine defines some situations in which simulation tests can give the best results. Cooper argues for *gedanken* experiments as a substitute for, or at least a preliminary to, expensive tests with actual data.

The third and final part of the book contains descriptions of the major experiments in information retrieval over the last twenty-five years. The first chapter contains a survey by Sparck Jones of the development of retrieval system testing during this period. The second chapter, also by Sparck Jones, describes in detail the influential series of experiments at Cranfield. The final chapter, by Gerard Salton, describes his own experimental system, the SMART system, which has served as a testing ground for the most interesting developments in the United States for the last fifteen years.

The whole book is a fascinating blend of philosophy and practice centering on the problem of how to design and carry out carefully controlled, valid experiments on data that presents all the bewildering peculiarities of human language.

*Martha Evens*

Illinois Institute of Technology  
Chicago, IL 60616