Keith van Rijsbergen Named ACM Fellow

On December 11, 2003 the Association of Computing Machinery (ACM) announced that C. J. (Keith) van Rijsbergen was one of 30 distinguished members named ACM Fellows.

His award summarizes his contributions as follows: *For contributions to the theory and practice of information retrieval, especially probabilistic and logic-based formalisms, and to ACM SIGIR conferences and education.*

We would like to take this opportunity to congratulate Keith and to fill in some additional details about Keith's contributions to SIGIR and the field of information retrieval.

For more than 30 years Keith van Rijsbergen has been dedicated to providing a principled and solid theoretical foundation for information retrieval (IR) systems. His work in the 70's and early 80's focused on automatic document clustering, evaluation, and modeling term co-occurrence relationships. His document clustering research developed efficient and effective techniques for grouping documents, a technique that is again gaining popularity as a way to organize document collections and web search results. His work on evaluation drew from measurement theory to define evaluation metrics for IR systems, including the F measure, which combines recall and precision components and which is widely used in language technology research today.

His work on document clustering led to considering dependencies between terms in the context of probabilistic models of retrieval. His 1977 paper, "A theoretical basis for the use of co-occurrence data in information retrieval", was the first paper to explore alternatives to the assumption of independence of document features (terms) in information retrieval systems. The paper explored a way of removing the independence assumption by using corpus statistics to construct a non-linear weighting function. This paper was influential in developing the widely-used probabilistic model of information retrieval, inference networks, as well as more recent language models. The work on co-occurrence was further developed for use in relevance feedback, a technique that is popular and widely-used today.

During the 70's, Keith also wrote a classic book, *Information Retrieval*. This book is more than a textbook; it contains substantial original material plus original reformulations of the work of others. This book is still widely used today, and has recently been republished as a CD in Belew's book *Finding Out About*. It is arguably the most influential book published on information retrieval theory.

More recently, Keith has worked on the development of non-classical logic as a basis for information retrieval. Although Boolean logic is the basis for most database and early IR systems, its limitations for text retrieval led to exploration of richer models that explicitly represent uncertainty. In 1986 he proposed a new paradigm for probabilistic information retrieval in which retrieval is regarded as a process of uncertain inference. The approach is based on the assumption that queries and documents can be represented as logical statements, and to answer a query the retrieval system must prove the query from the documents. This

logical interpretation of queries and documents emphasizes that retrieval and relevance are an inference process. The use of logic for IR models provides a uniform framework within which to represent features of IR systems such as hypertext links, multimedia data, and a user's knowledge, and to integrate IR with database systems. Uncertainty is critical to this process, because a collection of documents is not a consistent and complete set of statements, and documents may contradict each other. The implication is not implication in the classical sense but rather uses a logical uncertainty principle that characterizes the uncertainty of relevance. Keith explored using possible world semantics for the inference process. This approach has been substantially developed and extended by several research groups, as summarized in a recent book by Crestani, Lalmas and van Rijsbergen, *Information Retrieval: Uncertainty and Logics*.

Keith's main contributions have been theoretical in articulating models of information retrieval that have solid probabilistic and logical foundations. In addition to his theoretical work, Keith has explored a variety of aspects of end-to-end information retrieval systems including automatic query expansion, multi-media retrieval systems, and incorporating user's searching behaviors in relevance feedback.

Keith has also made substantial contributions to the information retrieval community through his teaching and professional commitments. He has contributed to ACM conferences in general, and SIGIR in particular, for more than 25 years. He was the Conference Chair for the 1980 ACM:SIGIR meeting in Cambridge England. This conference was the first time that the SIGIR conference was held outside of North America, and it began a tradition that is a key to SIGIR's success today. Keith also hosted a second SIGIR conference in Cambridge England in 1984, served as Program Chair for six ACM:SIGIR conferences, and has given two invited addresses at the SIGIR conference. He was the founding editor of Information Technology, and the Editor-in-Chief of the Computer Journal from 1993-2000. He is a Fellow of the Institute of Electrical Engineers (IEE) and the British Computer Society (BCS).

Keith's contribution as an educator has been considerable, beginning with his text *Information Retrieval*, published in 1975, revised in 1979 and reissued in 2000. Keith also contributes to the computer science community as a talented and dedicated teacher and mentor. He leads a large and active IR research group at the University Glasgow, and before that at Monash University, Cambridge University and University College Dublin. Over the years Keith has produced an active, and influential group of Ph.D. students that includes Fabio Crestani (Univ. of Strathclyde), W. Bruce Croft (Univ. of Massachusetts), Mark Dunlop (Univ. of Strathclyde), Mounia Lalmas (Queen Mary, Univ of London), Joan Morrissey (Univ. of Windsor), Ian Ruthven (Univ. of Strathclyde), Mark Sanderson (Univ. of Sheffield.), Alan Smeaton (Dublin City Univ.), Tassos Tombros (Queen Mary, Univ. of London), and David Harper (Robert Gordon Univ.).

On behalf of the information retrieval research community, we thank Keith for his many contributions to our field, and we congratulate him on his recognition as an ACM Fellow.

Susan Dumais, SIGIR Past-Chair
Jamie Callan, SIGIR Chair