Invited Talk:
Principles for Robust Evaluation Infrastructure

Mark Sanderson

RMIT University, Melbourne, Australia

William Webber

The University of Maryland, USA

Abstract

The standard "Cranfield" approach to the evaluation of information retrieval systems has been used and refined for nearly fifty years, and has been a key element in the development of large-scale retrieval systems. The resources created by such systematic evaluations have enabled thorough retrospective investigation of the strengths and limitations of particular variants of this evaluation approach; over the last few years, such investigation has for example led to identification of serious flaws in some experiments. Knowledge of these flaws can prevent their perpetuation into future work and informs the design of new experiments and infrastructures.

In this position statement we briefly review some aspects of evaluation and, based on our research and observations over the last decade, outline some principles on which we believe new infrastructure should rest.

Bio

Mark Sanderson is a Professor in Information Retrieval at RMIT University, Melbourne, Australia.
William Webber is a researcher at University of Maryland, USA.