PONENCIAS INVITADAS

Retrieval of XML documents

Jaap Kamps, University of Amsterdam

Correo electrónico: kamps@uva.nl

Standard evaluation benchmarks in the Cranfield/TREC paradigm have served our field well by quantifying system effectiveness in a meaningful way and by aligning the research agendas of many groups around the world. But this is currently under challenge by rapid recent developments. Increasingly IR research is about improving a particular search engine for a specific task and having a unique content and user community. How does this content differ, and how do the tasks and searchers differ? Can we make the evaluation tailored to their unique characteristics? Also any action on the Web leaves its trails and rich contextual information is available, especially in a mobile setting. How do such changes factor into the old search problem? Within this frame, some of the initial ideas at INEX, CLEF and the other evaluation fora will be discussed.

Experimental Evaluation in Visual Information Retrieval

Paul Clough, University of Sheffield

Correo electrónico: p.d.clough@sheffield.ac.uk

Evaluation is critical for designing, developing and evolving effective IR systems and ensuring high levels of user satisfaction. Much of the work on IR evaluation has focused on measuring system effectiveness using the Cranfield methodology. An example of this style of evaluation for visual IR systems has been the ImageCLEF series of evaluations, which has been run as part of the Cross Language Evaluation Forum (CLEF) since 2003. This talk will describe approaches to evaluating IR systems and in particular focus on the evolution of the ImageCLEF track, including tasks addressed, resources created, challenges faced and some of the lessons learned. I will also describe other evaluation activities, such as the TREC Session Track and the Evaluating Information Access Systems (ELIAS) ESF Research Networking Programme, that seek to address some of the limitations of existing evaluation approaches and further our understanding of IR evaluation. I will conclude with some thoughts on possible future directions for IR evaluation.