

# The Text REtrieval Conferences (TRECs)

*Donna Harman*

**National Institute of Standards and Technology  
Gaithersburg, MD 20899**

There have been four Text REtrieval Conferences (TRECs); TREC-1 in November 1992, TREC-2 in August 1993, TREC-3 in November 1994 and TREC-4 in November 1995. The number of participating systems has grown from 25 in TREC-1 to 36 in TREC-4, including most of the major text retrieval software companies and most of the universities doing research in text retrieval (see table for some of the participants). The diversity of the participating groups has ensured that TREC represents many different approaches to text retrieval, while the emphasis on individual experiments evaluated in a common setting has proven to be a major strength of TREC.

The test design and test collection used for document detection in TIPSTER was also used in TREC. The participants ran the various tasks, sent results into NIST for evaluation, presented the results at the TREC conferences, and submitted papers for a proceedings. The test collection consists of over 1 million documents from diverse full-text sources, 250 topics, and the set of relevant documents or "right answers" to those topics. A Spanish collection has been built and used during TREC-3 and TREC-4, with a total of 50 topics.

TREC-1 required significant system rebuilding by most groups due to the huge increase in the size of the document collection (from a traditional test collection of several megabytes in size to the 2 gigabyte TIPSTER collection). The results from TREC-2 showed significant improvements over the TREC-1 results, and should be viewed as the appropriate baseline representing state-of-the-art retrieval techniques as scaled up to handling a 2 gigabyte collection.

TREC-3 therefore provided the first opportunity for more complex experimentation. The major experiments in TREC-3 included the development of automatic query expansion techniques, the use of passages or sub-documents to increase the precision of retrieval results, and the use of the training information to select only the best terms for routing queries. Some groups explored hybrid approaches (such as the use of the Rocchio methodology in systems not using a vector space model), and others tried approaches that were radically different from their original approaches.

TREC-4 allowed a continuation of many of these complex experiments. The topics were made much

shorter and this change triggered extensive investigations in automatic query expansion. There were also five new tasks, called tracks. These were added to help focus research on certain known problem areas, and included such issues as investigating searching as an interactive task by examining the process as well as the outcome, investigating techniques for merging results from the various TREC subcollections, examining the effects of corrupted data, and evaluating routing systems using a specific effectiveness measure. Additionally more groups participated in a track for Spanish retrieval.

The TREC conferences have proven to be very successful, allowing broad participation in the overall DARPA TIPSTER effort, and causing widespread use of a very large test collection. All conferences have had very open, honest discussions of technical issues, and there have been large amounts of "cross-fertilization" of ideas. This will be a continuing effort, with a TREC-5 conference scheduled in November of 1996.

## A Sample of the TREC-4 Participants

CLARITECH/Carnegie Melon University
CITRI, Australia
City University, London
Cornell University
Department of Defense
Excalibur Technologies, Inc.
GE Corporate R & D/New York University
George Mason University
HNC, Inc.
Lexis-Nexis
Logicon Operating Systems
NEC Corporation
New Mexico State University
Queens College, CUNY
Rutgers University (two groups)
Siemens Corporate Research Inc.
Swiss Federal Institute of Technology (ETH)
University of California - Berkeley
University of Massachusetts at Amherst
University of Waterloo
Xerox Palo Alto Research Center