System Summary and Timing
Organization Name: Dortmund

- Construction of Indices, Knowledge Bases, and other Data Structures
  - Methods Used to build Data Structures
    - Length (in words) of the stopword list:577 words
    - Stemming Algorithm: Standard SMART stemmer
    - Term Weighting: Yes
    - Phrase Discovery? :
      - Kind of Phrase: Any pair of adjacent non-stopwords that occurred 25 times in D1
      - Other Techniques for building Data Structures:we incorporated the manually assigned index-terms appearing in ziff material (the SGML-tag DESCRIPT) in the indexing process, because we considered it part of the documents. The SGML-tags MS and NS in the wsj material are not included.
  - Statistics on Data Structures built from TREC Text
    - Inverted index
      - Total Storage (in MB): 460 MB for D3 (test set for routing)
      - 900 MB for D12 (learning set for routing and test set for ad-hoc)
      - Total Computer Time to Build (in hours): 5 CPU hours for D3
      - 10 CPU hours for D12
      - Use of Term Positions? : No
      - Only Single Terms Used? :No
- Query construction
  - Automatically Built Queries (Ad-Hoc)
    - Topic Fields Used:Everything except Definitions
    - Average Computer Time to Build Query (in cpu seconds): 30/60 CPU seconds for each query
    - Method used in Query Construction
      - Term Weighting (weights based on terms in topics)? :Yes
      - Phrase Extraction from Topics? :Yes
  - Machine Information
    - Machine Type for TREC Experiment: Sparc 10
    - Amount of Hard Disk Storage (in MB): 6300 MB
    - Amount of RAM (in MB): 157 MB
  - System Comparisons
    - Amount of "Software Engineering" which went into the Development of the System: Several years