

Pengembangan sistem temu-kembali informasi terdistribusi berbahasa Indonesia

Achmad Nizar Hidayanto, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=126657&lokasi=lokal>

Abstrak

Ide dasar dilakukannya penelitian ini adalah untuk mengembangkan dan mengimplementasikan sistem temu kembali informasi dalam lingkungan yang sifatnya terdistribusi (Distributed Information Retrieval System). Pengertian terdistribusi di sini adalah penggunaan satu prosesor untuk mengakses beberapa koleksi database dokumen.

Pengimplementasian sistem temu-kembali informasi terdistribusi menjadi begitu penting karena sifat alamiah dari koleksi dokumen yang letak penyimpanannya tersebar secara fisik (misalnya, koleksi dokumen yang berada di berbagai instansi). Di samping itu, setiap pengelola dokumen tentunya menginginkan adanya kebebasan dalam mengatur koleksi dokumen yang dimilikinya. Sifat dokumen yang seperti inilah yang menyebabkan diperlukannya suatu mekanisme khusus, agar seseorang yang membutuhkan informasi tidak mengalami kesulitan ketika mencari informasi tersebut.

Untuk mengatasi hal tersebut di atas, diperlukan suatu sistem temu kembali yang bisa digunakan dalam lingkungan yang terdistribusi. Ada tiga bagian yang harus diperhatikan dalam mengembangkan sistem temu kembali informasi terdistribusi, yaitu (1) antar-muka ke pemakai (user interface), (2) mesin pencari (search engine) dengan kemampuan melakukan pencarian terhadap sumber-sumber dokumen yang ada di dalam berbagai LAN/WAN, (3) basis data dokumen.

Pengembangan sistem temu-kembali informasi terdistribusi ini (khususnya bagian mesin pencari) dibagi dalam 3 bagian yaitu: bagian pemeringkatan koleksi, mesin pencari di database lokal, bagian yang menggabungkan hasil pencarian yang diperoleh dari suatu koleksi. Teknik pengindeksan yang dipergunakan adalah pengindeksan frekuensi dan pengindeksan berbasis rumus Savoy. Teknik pemeringkatan yang dipergunakan untuk pemilihan koleksi adalah teknik Lexicon Inspection, D-WISE, dan modifikasi dari rumus Savoy. Sedangkan teknik temu kembali informasi yang dipergunakan untuk mencari koleksi dokumen dalam suatu server adalah teknik Extended Boolean P-Norm Model. Untuk menggabungkan hasil pencarian di berbagai koleksi dokumen, teknik yang dipergunakan adalah raw score dan weighted score.

Hasil penelitian menunjukkan bahwa teknik Lexicon Inspection memberikan hasil pemeringkatan yang lebih baik dibandingkan dengan teknik D-WISE dan Savoy Extension. Penelitian juga menunjukkan bahwa penggunaan meta-index berbasis indeks Savoy memberikan hasil yang lebih baik dibandingkan dengan indeks frekuensi. Selain itu penelitian juga menunjukkan bahwa pencarian tidak harus dilakukan di seluruh koleksi yang terambil. Dengan mengambil 90% koleksi teratas dari koleksi yang terambil, dokumen relevan yang terambil mendekati hasil yang diperoleh apabila dilakukan pencarian terhadap seluruh koleksi yang terambil.