

## BAB 5 PENUTUP

Bab ini akan membahas kesimpulan dari penelitian perolehan informasi lintas bahasa Inggris-Indonesia serta saran-saran yang dapat diimplementasikan untuk kelanjutan dari penelitian di topik yang sama.

### 5.1 Kesimpulan

Pada penelitian ini dipelajari pengaruh teknik-teknik perolehan informasi lintas bahasa yang terdiri dari teknik perbaikan kueri sebelum penerjemahan, perbaikan terjemahan oleh kamus, dan pengelompokan dokumen, serta teknik perbaikan kueri dengan melakukan perluasan. Teknik perbaikan kueri sebelum penerjemahan menggunakan teknik pengelompokan kata. Teknik perbaikan terjemahan kamus terdiri dari penerjemahan langsung, penerjemahan dengan menggunakan terjemahan yang muncul berulang kali pada setiap kamus, dan teknik pengelompokan dokumen. Teknik perluasan kueri menggunakan umpan balik relevan semu. Teknik-teknik tersebut berbasiskan teknik perolehan informasi lintas bahasa Inggris-Indonesia menggunakan kamus dwi-bahasa.

Teknik-teknik yang diterapkan untuk perolehan informasi lintas bahasa Inggris-Indonesia tidak dapat mendekati hasil perolehan informasi satu bahasa (Indonesia). Secara keseluruhan, hasil yang dicapai pada teknik untuk mengurangi keambiguan terjemahan mencapai hanya 63.84% terhadap perolehan informasi satu bahasa. Sedangkan penerapan teknik perluasan kueri hanya mencapai 47.56% terhadap perolehan informasi satu bahasa.

Penerapan teknik perbaikan terjemahan oleh kamus dwi-bahasa Inggris-Indonesia (*Post-Translation*) hanya dapat menghasilkan nilai hasil perolehan informasi tertinggi sebesar 60.80% terhadap hasil perolehan informasi satu bahasa (Indonesia).

Pada penerapan teknik pengelompokan kata pada kueri sebelum penerjemahan yang dikombinasikan dengan teknik perbaikan terjemahan oleh kamus (Kombinasi *Post-Translation* – *Pre-Translation*) dapat ditemukan peningkatan nilai hasil perolehan informasi antara 2.63% - 2.84% bila

dibandingkan hasil perolehan informasi untuk teknik perbaikan terjemahan oleh kamus.

Untuk penerapan umpan balik relevan semu dimana teknik ini melakukan perluasan kueri pada hasil penerjemahan dengan teknik-teknik lain, dapat ditemukan bahwa teknik ini tidak dapat menaikkan hasil perolehan informasi. Hasil terbaik yang diperoleh pada perluasan kueri adalah pada teknik yang menggunakan kombinasi perluasan kueri pada perbaikan kueri sebelum penerjemahan dan perluasan kueri pada perbaikan hasil penerjemahan yang mencapai hasil perolehan informasi sebesar 47.56% terhadap perolehan informasi satu bahasa.

## 5.2 Saran

Timbulnya permasalahan pada penelitian perolehan informasi Inggris-Indonesia karena kurang lengkapnya kamus sehingga penerjemahan hanya dapat dilakukan secara kata per kata. Untuk itu, pendeteksian frase diperlukan agar dapat dikenali dalam proses penerjemahan. Selain itu, untuk memperbaiki kualitas kamus dwi-bahasa Inggris-Indonesia, sebaiknya perbendaharaan kata-kata pada kamus dilengkapi dengan menambahkan juga terjemahan dari sumber yang lain.

Kurangnya koleksi dokumen berbahasa Indonesia dan berbahasa Inggris juga merupakan masalah pada penelitian ini. Dengan adanya koleksi dokumen yang besar maka topik dan perbendaharaan kata akan semakin bervariasi. Hal ini akan memudahkan dalam proses pengelompokan dokumen berdasarkan topik serta memungkinkan perluasan kueri memberikan hasil yang lebih baik. Untuk itu penambahan koleksi dokumen sebaiknya dilakukan untuk penelitian berikutnya.

Perbaikan teknik pengelompokan dokumen juga dapat dilakukan dengan menggunakan algoritma lain selain *centroid*. Pengelompokan dokumen yang sesuai dengan topik akan meningkatkan hasil teknik perbaikan kueri sebelum penerjemahan serta teknik perbaikan hasil penerjemahan kamus karena akan memperbaiki kata-kata kunci yang dibuat untuk setiap kelompok dokumen.

Selain itu, penerapan teknik lain seperti perluasan kueri menggunakan analisis konteks lokal (*local context analysis*) dapat dilakukan untuk menemukan teknik yang tepat pada perolehan informasi lintas bahasa Inggris-Indonesia.

## DAFTAR PUSTAKA

- Baeza-Yates, R. dan B. Ribeiro-Neto. (1999). *Modern Information Retrieval*. New York: Addison-Wesley.
- Ballesteros, L. dan W. B. Croft. (1996) "Dictionary Methods for Cross-Lingual Information Retrieval". *Proceedings of the 7th International DEXA Conference on Database and Expert Systems Applications*, pp. 791-801.
- Ballesteros, L. dan W. B. Croft. (1997). "Phrasal translation and query expansion techniques for cross-language information." *In Proc. of SIGIR'97*, pp. 84-91.
- Ballesteros, L. dan W. B. Croft. (1998). "Resolving ambiguity for cross-language retrieval." *In Proceedings of the 21st Annual International ACM SIGIR Conference*, pp. 64-71.
- Buckley, C., *et. al.* (2005). "The trec eval Evaluation Package." Trec Retrieval Conference. [http://trec.nist.gov/trec\\_eval/trec\\_eval.7.3.tar.gz](http://trec.nist.gov/trec_eval/trec_eval.7.3.tar.gz) [15 Mei 2008]
- Gao, Wei, Cheng Niu, Jian-Yun Nie, MingZhou, Jian Hu, Kam-Fai Wong, Hsiao Wuen Hon. (2007). "Cross-lingual query suggestion using query logs of different languages." *In Proceedings of SIGIR'2007*, pp. 463-470.
- Grossman, D., dan O. Frieder (2003). *Information Retrieval: Algorithms and Heuristics*. second Edition, Kluwer Academic Publishers. Norwell, Mass. 2003.
- Hasnah, A. M., dan Jihad, J. M. (2002). "Thesaurus-based Query Disambiguation Method for Cross-Language Information Retrieval." *IJICIS vol.2, no.2* , pp. 58-68.
- Hayurani, H., S. Sari, dan M. Adriani. (2006). "Evaluating Language Resources for English-Indonesian CLIR." CLEF 2006 Workshop. [http://www.clef-campaign.org/2006/working\\_notes/workingnotes2006/hayuraniCLEF2006.pdf](http://www.clef-campaign.org/2006/working_notes/workingnotes2006/hayuraniCLEF2006.pdf) (30 Mei 2008)

- Hedlund, T. (2002). "Compounds in dictionary-based cross-language information retrieval". *Information Research*, [Online], 7(2). Dari: <http://InformationR.net/ir/7-2/paper128.html> [27 Juni 2008]
- Manning Christopher D., Prabhakar Raghavan dan Hinrich Schütze. (2008). *Introduction to Information Retrieval*. Cambridge University Press.
- Nazief, B. A. A. dan M. Adriani. (1996). *Pendekatan Algoritma Pemotong Imbuhan Bahasa Indonesia*. Publikasi internal. Depok: Fakultas Ilmu Komputer Universitas Indonesia.
- Steinbach, M., G. Karypis, dan V. Kumar. *A Comparison of Document Clustering Techniques*. Laporan teknis, Department of Computer Science and Engineering, University of Minnesota.
- Talvensaari, T., J. Laurikkala, K. Järvelin dan Juhola M. (2006). "A study on automatic creation of a comparable document collection in cross-language information retrieval". *Journal of Documentation* 62(3), pp. 372-387.
- Wen, J. R., J. Y. Nie, dan H. J. Zhang. (2002) "Query Clustering Using User Logs." *ACM Trans. Information Systems*, 20(1):59-81.