

## PENGUNAAN KAMUS SINONIM DAN HIPONIM SEBAGAI SUMBER EKSPANSI KUERI DALAM SISTEM TEMU KEMBALI INFORMASI BERBAHASA INDONESIA

**Husni Thamrin<sup>1)</sup>, Agus Triyono<sup>2)</sup> Umi Fadlillah<sup>3)</sup>**

<sup>1</sup>Program Studi Informatika, Universitas Muhammadiyah Surakarta  
email: husni.thamrin@ums.ac.id

<sup>2</sup>Program Studi Ilmu Komunikasi, Universitas Muhammadiyah Surakarta  
email: agus.triyono@ums.ac.id

<sup>3</sup>Program Studi Informatika, Universitas Muhammadiyah Surakarta  
email: umi.fadlillah@ums.ac.id

### *Abstract*

*In search for information, a user usually inputs a small number of words as a query. A problem of vocabulary mismatch is arisen when the system does not retrieve a document that contains relevant information but does not contain words in the query. Query expansion is used to increase the chance to retrieve potentially relevant documents by adding new words to the original query. In this work, we investigated the use of synonyms, hypernyms and holonyms as the source of query expansion in retrieving scientific papers in bahasa Indonesia. We collected papers in the field of communications science into a database and allowed it accessed through a web portal with a search facility. We used big dictionary of bahasa Indonesia (KBBI) and have appended scientific terms of communications science therein. During retrieval, the documents are sorted based on their similarity with the query words. We measured the number of retrieved documents and the relevance of documents in the top positions. Query expansion is proved to increase the number of retrieved documents by approximately 23 percent compared to retrieval without query expansion. However, the relevance of the documents on the top of the list does not change a lot.*

**Keywords:** *ekspansi kueri, bahasa Indonesia, sinonim, hipernim, sistem temu kembali informasi*

### 1. PENDAHULUAN

Meningkatnya jumlah dokumen dalam sebuah arsip atau basis data memerlukan sistem kearsipan yang dilengkapi fasilitas pencarian informasi dalam dokumen. Fasilitas pencarian (*search*) informasi sering diistilahkan dengan sistem temu kembali informasi (*information retrieval system*). Ketika seseorang mencari informasi dalam dokumen, mungkin saja kata yang diketikkan tidak terdapat dalam dokumen yang ada di arsip karena dokumen yang ada menggunakan istilah yang berbeda dengan kata yang diketikkan. Dokumen yang mengandung kata yang berbeda meskipun mempunyai pengertian yang sama tidak akan ditemukan jika sistem temu kembali hanya mencocokkan kata seperti yang diketikkan pengguna. Kegagalan ini disebut

ketidakcocokan daftar kata atau *vocabulary mismatch* [1].

Ekspansi kueri (*query expansion*) merupakan cara yang dapat digunakan untuk mengatasi ketidakcocokan daftar kata. Ekspansi kueri mencoba mencari dokumen menggunakan kata-kata dengan pengertian sama dengan yang diketikkan pengguna [2]. Kata-kata lain di luar yang diketikkan pengguna diperoleh dari berbagai sumber misalnya kumpulan dokumen (korpus) di dalam arsip atau basis data, jejaring kata, tesaurus, dokumen web, wikipedia dan daftar kueri terdahulu (*query log*) [3].

Tulisan ini membahas penelitian tentang pemanfaatan kamus bahasa Indonesia sebagai sumber ekspansi kueri pada sistem temu kembali informasi berbahasa Indonesia. Kamus tersebut mempunyai daftar sinonim,

hiponim dan meronim dan ekspansi didasarkan pada sinonim, hipernim dan holonim. Penelitian menunjukkan adanya peningkatan pada jumlah dokumen hasil temu kembali namun relevansi dokumen yang mempunyai similaritas tertinggi dengan kueri (berada pada bagian teratas hasil kueri) tidak banyak berubah.

## 2. KAJIAN LITERATUR

Penelitian tentang temu kembali sistem informasi sudah dilakukan sejak lama. Sebagai contoh, tulisan [4] termasuk di antara telaah yang sangat populer yang membahas penilaian relevansi hasil temu kembali oleh manusia. Berbagai cara mengukur relevansi telah diusulkan tetapi berbagai telaah ekspansi kueri masih diupayakan. Beberapa paragraf berikut ini memaparkan sebagian penelitian terkait proses temu kembali informasi dengan ekspansi kueri.

Pada penelitian [5], Salton dan Buckley menyatakan bahwa kinerja sistem temu kembali informasi dapat ditingkatkan melalui penentuan relevansi oleh pengguna. Kedua peneliti ini menyimpulkan bahwa umpan balik relevansi merupakan cara yang murah untuk membentuk ulang kueri berdasarkan dokumen terdahulu yang relevan maupun tidak relevan.

Proses pada sistem temu kembali dokumen ilmiah mempunyai satu perbedaan penting dibanding dokumen teks biasa. Dokumen ilmiah atau dokumen yang disimpan secara terstruktur mempunyai perbedaan bobot relevansi tergantung dari posisi kata [6]. Kata yang muncul dalam judul dokumen ilmiah mempunyai bobot yang lebih tinggi dibanding kata yang sama yang muncul pada abstrak atau isi dokumen. Perhitungan relevansi disarankan agar memperhatikan faktor jenis data yang tersimpan dan menyimpulkan bahwa penggunaan bobot yang berbeda pada tiap jenis data dapat meningkatkan efektivitas dari sistem temu kembali.

Penelitian lain dilakukan terkait penggunaan kata-kata dalam beberapa kueri dalam satu sesi web (*web session*) dalam proses ekspansi kueri [7]. Pengguna dalam satu sesi web sering mencoba beberapa kueri ketika mencari informasi atau dokumen yang

diinginkan. Kueri-kueri yang diinputkan sebelumnya oleh pengguna yang sama dijadikan data dalam upaya mengekspansi kueri. Berdasarkan pengamatannya, para peneliti mendapati bahwa kumpulan kata dalam beberapa kueri pada satu sesi sangat efektif digunakan dalam ekspansi kueri untuk mendapatkan dokumen yang relevan. Metode ini menaikkan keakuratan ditemukannya dokumen yang relevan sebesar lebih dari 25 persen.

Pengguna sistem temu kembali kebanyakan mengetikkan kueri pendek [8]. Oleh karena itu diupayakan cara untuk memperbaiki keefektifan dari kueri pendek. Ekspansi kueri dilakukan dengan menambah operasi Boolean secara manual dan kemudian menambahkan kata baru berdasarkan korelasi dengan kata pada kueri semula. Cara ini dipandang dapat memperbaiki efektivitas kueri dan mengurangi pergeseran kueri (*query drift*) seperti yang terjadi pada ekspansi buta. Efektivitas kueri bisa ditingkatkan dilihat dari angka rata-rata total maupun dilihat pada 20 dokumen paling atas yang ada pada daftar hasil temu kembali.

Penelitian terbaru tentang ekspansi kueri dilakukan pada [3] oleh Pal, dkk. yang menggunakan distribusi istilah (*term distribution*) dan asosiasi istilah (*term association*). Kedua metode ini menggunakan korpus target sebagai sumber kata yang ditambahkan pada kueri asli. Metode distribusi istilah membandingkan distribusi sebuah istilah pada dokumen yang relevan dengan distribusi istilah pada keseluruhan korpus. Metode asosiasi istilah menggunakan informasi tentang cara sebuah istilah tertentu muncul bersama istilah dari kueri. Pal dkk. mencoba mengkombinasikan kedua metode itu, yaitu menggunakan distribusi istilah, dan dilanjutkan dengan metode asosiasi istilah. Hasil penelitian mereka menunjukkan bahwa kombinasi kedua metode memberikan daftar pencarian yang lebih baik daripada penggunaan salah satu metode.

## 3. METODE PENELITIAN

Penelitian ini mencermati proses temu kembali informasi berupa dokumen ilmiah (*paper*). Untuk itu telah dibuat sebuah portal atau basis data terdiri atas 373 paper

berbahasa Indonesia pada bidang ilmu komunikasi. Portal yang berbasis web tersebut mempunyai fasilitas pencarian dokumen. Kueri pada proses pencarian dokumen diekspansi dengan data sinonim, hipernim dan holonim dari kata kunci pencarian.

Penambahan sinonim, hipernim dan holonim membutuhkan kamus sebagai basis data pengetahuan. Data diperoleh dari Tesaurus Bahasa Indonesia [9] yang mengandung entri tentang sinonim, hiponim dan meronim. Pada kamus telah ditambahkan pula frase istilah bidang ilmu komunikasi termasuk sinonim, hiponim dan meronimnya. Beberapa entri kamus istilah komunikasi diperoleh dari kamus tentang media dan komunikasi [10]. Jumlah frase dasar pada kamus adalah 69.782 frase, dengan 109.530 frase sinonim dan 4.579 hiponim/meronim sedangkan jumlah frase istilah bidang komunikasi sebesar 103 dengan 637 kata terkait.

Untuk mencermati pengaruh ekspansi kueri terhadap hasil temu kembali, diujikan 13 macam kueri ke dalam sistem temu kembali (Tabel 1). Kueri bervariasi dalam jumlah kata yang digunakan. Ketiga belas kueri diberikan pada sistem tanpa ekspansi kueri (TE), dengan ekspansi menggunakan sinonim (Sin), dengan ekspansi menggunakan hipernim/holonim (Hip), dan dengan ekspansi menggunakan sinonim/hipernim/holonim (SH). Jumlah dokumen yang ditemukan oleh masing-masing metode kemudian dibandingkan.

Selain jumlah dokumen yang dihasilkan, dicermati pula relevansi hasil temu kembali. Kesesuaian dokumen dengan kueri dinilai oleh 2 orang yang merupakan anggota peneliti dengan memberi pernyataan relevan atau tidak relevan. Jika kedua orang itu berbeda keputusan, orang ketiga diundang untuk memberi penilaian. Penilaian relevansi oleh manusia merupakan tolok ukur (*benchmark*) terhadap hasil temu kembali yang dilakukan oleh program komputer (algoritma).

Hasil temu kembali yang dicermati biasanya adalah hasil yang terdapat pada bagian teratas dalam daftar (misalnya pada

halaman pertama web). Oleh karena itu penelitian difokuskan pada pengamatan relevansi terhadap hasil temu kembali yang ada pada daftar teratas, yaitu 5 dokumen teratas, kemudian 10, 15 dan 20 dokumen teratas.

Tabel 1. Kueri yang digunakan dalam pengujian

Nomor Kueri	
1	Studi deskriptif kualitatif komunikasi antar pribadi orang tua terhadap anak
2	Pengaruh program corporate social responsibility terhadap peningkatan kesejahteraan karyawan
3	Studi analisis isi adegan kekerasan dalam serial televisi
4	Studi tingkat kepuasan penggunaan media dan gratifikasi
5	Studi deskriptif kualitatif aktivitas media relations
6	Hubungan komunikasi organisasi dengan kinerja karyawan
7	Analisis semiotik representasi perempuan dalam film
8	Pengaruh tayangan televisi terhadap perilaku
9	Strategi komunikasi pemasaran perusahaan
10	Framing pemberitaan koran
11	Iklan di media massa
12	Komunikasi organisasi
13	Media

Penentuan posisi dokumen pada daftar teratas dilakukan dengan menghitung similaritas kueri dan dokumen seperti yang ditunjukkan pada persamaan (1).

$$\text{sim} = \sum_{i=1}^T W_{qi} \cdot W_{di} \quad (1)$$

Pada persamaan (1),  $T$  adalah jumlah istilah dalam kamus,  $W_{qi}$  adalah bobot istilah  $i$  dalam kueri, dan  $W_{di}$  adalah bobot istilah  $i$  dalam dokumen. Bobot istilah ditentukan dengan *raw term frequency* (tf) yaitu dengan menghitung jumlah kemunculan istilah atau

kata dalam kueri maupun dalam dokumen. Bobot untuk kata dalam judul dibedakan dari bobot untuk kata dalam abstrak yaitu sebesar 3:1. Metode *inverse document frequency* tidak digunakan dalam perhitungan bobot dengan asumsi bahwa ukuran abstrak pada tiap paper relatif sama (200 – 300 kata).

Kinerja algoritma diukur menggunakan parameter *precision*, *recall* dan *F-measure* dengan perhitungan sebagaimana disajikan pada persamaan (2).

$$\begin{aligned} \text{Precision} &= P = TP/(TP + FP) \\ \text{Recall} &= R = TP/(TP + FN) \\ \text{F-measure} &= F = (1 + \beta) P R / (\beta P + R) \end{aligned} \quad (2)$$

Pada persamaan (2), TP adalah jumlah dokumen yang dimasukkan dalam kelompok ranking teratas oleh algoritma dan termasuk dokumen yang relevan, FP adalah jumlah dokumen yang dimasukkan dalam kelompok ranking teratas oleh algoritma tapi tidak termasuk yang relevan, FN adalah jumlah dokumen yang tidak dimasukkan dalam kelompok ranking teratas oleh algoritma tapi merupakan dokumen yang relevan, dan  $\beta$  adalah bobot yang menunjukkan mana yang

lebih penting antara *precision*  $P$  atau *recall*  $R$ . Pada penelitian ini dipilih  $\beta = 1$ .

#### 4. HASIL DAN PEMBAHASAN

Hasil temu kembali informasi dipengaruhi oleh kueri yang dimasukkan. Sebagaimana terlihat pada Tabel 2, tiga belas kueri menghasilkan jumlah hasil temu kembali yang bervariasi dari yang hanya dua puluhan paper hingga ratusan paper (di antara total 373 paper yang ada dalam basis data).

Jumlah hasil temu kembali dapat ditingkatkan jika proses temu kembali melibatkan ekspansi kueri, yaitu penggunaan kata-kata tambahan selain kata yang ada pada kueri semula. Pada Tabel 2 tersaji data jumlah hasil temu kembali untuk kueri yang tidak diekspansi, kueri yang diekspansi dengan sinonim, kueri yang diekspansi dengan hipernim dan holonim, dan kueri yang diekspansi dengan sinonim, hipernim dan holonim. Sinonim merupakan kata yang memiliki makna setara, hipernim adalah kata yang memiliki makna lebih luas dari sebuah kata benda, holonim merupakan kata yang memiliki makna lebih luas dari sebuah kata kerja atau kata sifat.

Tabel 2. Jumlah dan kenaikan paper yang ditemukan dengan proses yang menggunakan beberapa cara ekspansi; TE: tanpa ekspansi, Sin: ekspansi dengan sinonim, Hip: ekspansi dengan hipernim/holonim, SH: ekspansi dengan sinonim dan hipernim/holonim

Nomor kueri	TE	Sin		Hip		SH	
		Jumlah	Kenaikan (%)	Jumlah	Kenaikan (%)	Jumlah	Kenaikan (%)
1	104	144	38,5	108	3,8	146	40,4
2	55	70	27,3	55	0,0	70	27,3
3	188	188	0,0	188	0,0	188	0,0
4	75	101	34,7	75	0,0	101	34,7
5	68	105	54,4	68	0,0	105	54,4
6	114	124	8,8	117	2,6	127	11,4
7	218	228	4,6	221	1,4	230	5,5
8	92	129	40,2	101	9,8	135	46,7
9	96	133	38,5	99	3,1	136	41,7
10	23	33	43,5	26	13,0	35	52,2
11	133	145	9,0	138	3,8	148	11,3
12	114	157	37,7	119	4,4	162	42,1
13	151	166	9,9	152	0,7	167	10,6
Seluruh kueri	1431	1723	20,4	1467	2,5	1750	22,3

Jumlah hasil temu kembali meningkat signifikan pada proses ekspansi menggunakan sinonim. Secara total peningkatannya

mencapai 20,4 persen dari jumlah hasil kueri tanpa ekspansi. Peningkatan hasil temu kembali juga terjadi jika ekspansi dilakukan

menggunakan hipernim atau holonim, yaitu sebesar 2,5 persen. Jika ekspansi dilakukan menggunakan sinonim, hipernim dan holonim, kenaikan yang diperoleh menjadi maksimal yaitu sebesar 22,3 persen.

Tingkat kenaikan jumlah hasil temu kembali tampaknya dipengaruhi oleh jumlah basis pengetahuan yang dimanfaatkan dalam ekspansi kueri. Penelitian ini menggunakan kamus sebagai basis data pengetahuan dan dalam kamus tersebut terdapat 109.530 data sinonim dan sebanyak 4.579 data hiponim atau meronim. Sebagai akibatnya, penggunaan sinonim dalam ekspansi kueri meningkatkan jumlah temu kembali secara signifikan sedangkan penggunaan hipernim dan holonim hanya mengakibatkan sedikit peningkatan.

Jumlah kata dalam kueri tidak mempunyai korelasi yang jelas dengan jumlah hasil temu kembali. Kueri nomor 4, misalnya, terdiri atas 6 kata dan menghasilkan 75 paper hasil temu kembali tanpa ekspansi (kolom TE pada Tabel 2). Akan tetapi, kueri nomor 5 yang terdiri atas 6 kata memberikan 68 hasil temu kembali dan kueri nomor 6 yang terdiri atas 5 kata memberikan 114 hasil temu kembali. Lebih kontras lagi adalah kueri nomor 7 yang terdiri atas 5 kata dan memberikan 218 paper hasil temu kembali.

Hasil penentuan relevansi paper terhadap kueri menunjukkan bahwa 171 dari 515 paper yang ditemukan (dan ditelaah) merupakan paper yang relevan dengan kueri yang dimasukkan oleh pengguna. Jumlah ini setara dengan 33 persen. Jika dicermati lebih lanjut, relevansi hasil temu kembali untuk setiap kueri bervariasi antara kueri yang satu dengan kueri yang lain dan tidak ditemukan pola hubungan relevansi dengan jumlah hasil temu kembali. Ada kueri yang hanya menghasilkan 1 paper relevan (2,5% dari paper yang ditelaah) dari 179 paper yang ditemukan, dan ada pula yang menghasilkan 27 paper relevan (67,5% dari paper yang ditelaah) dari 119 paper yang ditemukan.

Perhitungan nilai *precision*, *recall* dan *F-measure* telah dilakukan untuk seluruh kueri dan metode ekspansi kueri, yaitu ekspansi dengan sinonim, ekspansi dengan hipernim/holonim, dan ekspansi dengan sinonim/hipernim/holonim. Hasil perhitungan

disajikan pada Tabel 3. Perhitungan parameter kinerja dilakukan terhadap sejumlah dokumen yang diletakkan pada posisi teratas dalam daftar hasil temu kembali, yaitu posisi 5 teratas (Top 5), 10, 15 dan 20 dokumen teratas (masing-masing diberi label Top 10, Top 15 dan Top 20 pada Tabel 3). Nilai rata-rata parameter disajikan di bagian paling bawah tabel.

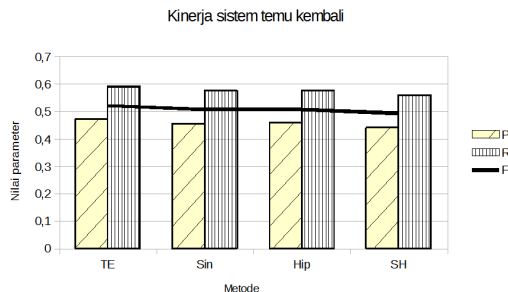
Tabel 3. Kinerja algoritma ekspansi kueri untuk beberapa jumlah data yang dianalisis

Posisi data	Metode ekspansi	Precision	Recall	F-measure
Top 5	TE	0,533	0,533	0,533
	Sin	0,517	0,517	0,517
	Hip	0,500	0,500	0,500
	SH	0,467	0,467	0,467
Top 10	TE	0,517	0,569	0,541
	Sin	0,500	0,550	0,524
	Hip	0,500	0,550	0,524
	SH	0,500	0,550	0,524
Top 15	TE	0,489	0,601	0,539
	Sin	0,475	0,587	0,525
	Hip	0,483	0,594	0,533
	SH	0,469	0,580	0,519
Top 20	TE	0,474	0,683	0,560
	Sin	0,466	0,671	0,550
	Hip	0,472	0,683	0,558
	SH	0,461	0,665	0,545

Hasil perhitungan menunjukkan bahwa sistem temu kembali dengan ekspansi kueri memberikan nilai kinerja yang hampir sama dibanding sistem tanpa ekspansi kueri. Ada sedikit penurunan angka kinerja untuk seluruh parameter yang diamati yaitu *precision*, *recall* maupun *F-measure*, namun perbedaan kinerja tersebut tidak signifikan yaitu antara 0,01 hingga 0,03 poin. Fenomena ini terjadi secara konsisten untuk berbagai variasi jumlah data yang dicermati, baik untuk data 5 dokumen teratas hingga data 20 dokumen teratas.

Gambar 1 memperlihatkan secara grafis kinerja rata-rata sistem temu kembali. Sekilas data untuk sistem tanpa ekspansi dan dengan ekspansi tampak sama. Namun jika dihitung lebih akurat, nilai *precision* secara rata-rata turun 0,029 poin, nilai *recall* turun 0,031 poin.

Rata-rata angka F-measure adalah 0,543 untuk proses temu kembali tanpa ekspansi kueri, sedangkan proses temu kembali dengan ekspansi kueri dengan sinonim, hipernim dan holonim mempunyai F-measure sebesar 0,513, jadi turun sebesar 0,03 poin.



Gambar 1. Kinerja rata-rata sistem temu kembali tanpa dan dengan ekspansi kueri;  $P$ =precision,  $R$ =recall,  $F$ =F-measure.

Pendekatan dalam penelitian ini adalah upaya penambahan kueri dengan menggunakan kamus sebagai basis pengetahuan (*knowledge based*) dan tidak melihat dari sudut pandang yang lain. Sebenarnya peningkatan relevansi dapat dilakukan dari berbagai sisi. Pengujian selama pengamatan menunjukkan bahwa relevansi berubah karena pemberian bobot yang berbeda ketika istilah ditemukan pada judul atau pada abstrak. Relevansi juga berubah ketika perhitungan bobot istilah dikombinasikan antara bobot tiap kata dan bobot kumpulan kata (*n-gram*).

Beberapa peneliti lain berupaya meningkatkan relevansi dengan sudut pandang lain. Pada [11], misalnya, ekspansi kueri tidak dilakukan secara buta tetapi didasarkan pada penentuan topik terlebih dahulu. Kumpulan kata pada kueri diolah untuk melihat topik yang dicari oleh pengguna dan kata-kata baru dalam ekspansi ditambahkan jika termasuk dalam topik hasil telaah. Pada [12], topik ditentukan melalui catatan kueri sebelumnya dan ekspansi tidak harus dilakukan terhadap semua istilah, namun dilakukan terhadap istilah yang paling terkait dengan kueri.

Zhou dan kawan-kawan [13] melihat dari sudut pandang pengguna yang menginputkan kueri dan melakukan ekspansi berdasarkan data yang ada pada akun media sosial

pengguna. Data yang dimaksud adalah profil, *tag* dan komentar yang diberikan oleh pengguna tersebut pada media sosial yang dimilikinya. Oleh karena itu, kata-kata ekspansi berasal dari dunia dari pengguna itu sendiri. Cara ini sepenuhnya sesuai dengan yang telah dilakukan dalam penelitian ini, yaitu melakukan ekspansi berdasarkan bidang dari informasi yang tersimpan, dalam hal ini bidang ilmu komunikasi. Hal yang mirip telah pula dilakukan yaitu pada [14] dengan menggunakan istilah-istilah bidang hukum dalam upaya menemukan dokumen legal.

## 5. SIMPULAN

Penelitian ini menghasilkan portal dokumen ilmiah bidang ilmu komunikasi dengan fasilitas pencarian. Pengujian telah dilakukan untuk mencermati pengaruh ekspansi kueri dengan kamus sinonim dan hiponim terhadap hasil temu kembali informasi. Ekspansi kueri meningkatkan jumlah hasil temu kembali. Ekspansi kueri dengan sinonim meningkatkan jumlah hasil temu kembali sebesar 20,4 persen dan ekspansi dengan hipernim meningkatkan hasil temu kembali sebesar 2,5 persen. Penggabungan sinonimi dan hipernimi untuk ekspansi kueri menghasilkan peningkatan jumlah hasil temu kembali yang paling besar yaitu sebesar 22,3 persen.

Meskipun ekspansi kueri meningkatkan jumlah hasil temu kembali secara signifikan, relevansi dari dokumen atau paper yang ditemukan dan diletakkan pada posisi teratas dalam daftar hasil temu kembali tidak berbeda secara signifikan. Tanpa ekspansi kueri, sistem temu kembali mempunyai kinerja dengan parameter precision bernilai 0,503, recall 0,597 sehingga parameter F-measure bernilai 0,543. Ekspansi kueri dengan sinonim, hipernim dan holonim memberikan parameter precision sebesar 0,474, recall 0,566 sehingga F-measure bernilai 0,513. Penelitian ini menunjukkan bahwa relevansi hasil temu kembali sedikit menurun jika sistem temu kembali dilengkapi ekspansi kueri dengan sinonim, hipernim ataupun keduanya.

## 6. REFERENSI

- [1] Carpineto, C., dan Romano, G., 2012. "A survey of automatic query expansion in

- information retrieval.” *ACM Computing Surveys (CSUR)*, 44(1), 1.
- [2] Xu, J., dan Croft, W. B., 1996. “Query expansion using local and global document analysis,” *Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval*, Agustus 1996, ACM, hal. 4-11.
- [3] Pal, D., Mitra, M., dan Datta, K., 2013. “Query Expansion Using Term Distribution and Term Association,” arXiv preprint arXiv:1303.0667.
- [4] Lesk, M. E., dan Salton, G., 1968. “Relevance assessments and retrieval system evaluation,” *Information storage and retrieval*, 4(4), hal. 343-359.
- [5] Salton, G. dan Buckley, C., 1988. “Improving Retrieval Performance by Relevance Feedback,” *Readings in Information Retrieval*, 24(5).
- [6] Kim, J. Y. dan Croft, W. B., 2012. “A field relevance model for structured document retrieval,” *Advances in Information Retrieval*, Springer Berlin Heidelberg, hal. 97-108.
- [7] Zhang, S., Guan, D., dan Yang, H., 2013. “Query change as relevance feedback in session search,” *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, Juli 2013, ACM, hal. 821-824.
- [8] Mitra, M., Singhal, A., dan Buckley, C., 1998. “Improving automatic query expansion,” *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, ACM, Agustus 1998, hal. 206-214.
- [9] Pusat Bahasa, 2009. *Tesaurus Alfabetis Bahasa Indonesia*, Bandung: Mizan.
- [10] Danesi, M., 2009. *Dictionary of Media and Communications*, London, ME Sharpe.
- [11] Wei, C. P., Hu, P. J. H., Tai, C. H., Huang, C. N., & Yang, C. S., 2007. “Managing word mismatch problems in information retrieval: A topic-based query expansion approach,” *Journal of Management Information Systems*, 24(3), hal. 269-295.
- [12] Saneifar, H., Bonniol, S., Poncelet, P., & Roche, M., 2014. “Enhancing passage retrieval in log files by query expansion based on explicit and pseudo relevance feedback,” *Computers in Industry*, 65(6), hal. 937-951.
- [13] Zhou, D., Lawless, S., & Wade, V., 2012. “Improving search via personalized query expansion using social media,” *Information retrieval*, 15(3-4), hal. 218-242.
- [14] Maxwell, T., & Schafer, B., 2010. “Natural language processing and query expansion in legal information retrieval: Challenges and a response,” *International Review of Law, Computers & Technology*, 24(1), hal. 63-72.