

STUDY LITERATUR: QUERY EXPANSION PADA TEMU KEMBALI INFORMASI (INFORMATION RETRIEVAL)

Rifa'atus Shalihah¹, Sitti Najmia Rifai², Arga Wiradarma³, A. Ainun Nurfajrin. S⁴

^{1,3}Program Studi Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Halu Oleo
Kota Kendari, Indonesia

²Program Studi Komunikasi dan Penyiaran Islam, Fakultas Ushuluddin, Adab dan Dakwah, Institut Agama Islam Negeri Kendari
Kota Kendari, Indonesia

⁴Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Halu Oleo
Kota Kendari, Indonesia

rifaatus.shalihah@aho.ac.id¹, sittinajmia17@iainkendari.ac.id², arga.wiradarma@aho.ac.id³,
ainunnurfajrin@aho.ac.id⁴

Received : Month, Year	Accepted : Month, Year	Published : Month, Year
------------------------	------------------------	-------------------------

Abstract

Information Retrieval plays a crucial role in assisting users in locating relevant documents within increasingly large data collections. One of the main challenges in the retrieval process is the mismatch between the terms used in a user's query and the vocabulary contained in documents. Query expansion is widely employed to enrich the initial query by adding related terms, thereby broadening the representation of the user's information needs. This article presents a literature review of several commonly used query expansion methods, including thesaurus-based approaches, WordNet, association rules, and relevance feedback. This review provides a comprehensive overview of the various query expansion techniques that have been developed within the field of Information Retrieval and serves as a reference for future research aimed at advancing more effective retrieval methods.

Keywords: *Information retrieval, Query Expansion, Thesaurus, Wordnet, Association Rule, Relevance Feedback*

Abstrak

Temu balik informasi (Information Retrieval) berperan penting dalam membantu pengguna menemukan dokumen yang relevan dari koleksi data yang semakin besar. Salah satu permasalahan utama dalam proses pencarian adalah perbedaan istilah antara query yang diberikan pengguna dan kosakata yang terdapat dalam dokumen. Query expansion menjadi pendekatan yang banyak digunakan untuk memperkaya query awal dengan istilah tambahan agar representasi kebutuhan informasi pengguna menjadi lebih luas. Artikel ini menyajikan tinjauan literatur mengenai beberapa metode query expansion yang umum digunakan, termasuk pendekatan berbasis thesaurus, WordNet, association rule, dan relevance feedback. Kajian ini memberikan gambaran menyeluruh tentang variasi teknik query expansion yang telah dikembangkan dalam ranah Information Retrieval dan dapat menjadi referensi bagi penelitian selanjutnya terkait pengembangan metode pencarian yang lebih efektif.

1. PENDAHULUAN

Mesin pencari (*Search Engine*) merupakan salah satu sistem pencarian informasi yang paling banyak digunakan dan memungkinkan pengguna untuk menemukan informasi yang relevan dengan *query*. Penggunaan sistem temu balik informasi (*Information Retrieval*) menjadi alternatif dan banyak digunakan dalam pembuatan mesin pencari. Hal ini dikarenakan sistem temu balik informasi mampu memproses Bahasa alami (*Natural Language*) serta merangking daftar dokumen sesuai dengan tingkat relevansi antara *query* dan dokumen yang ditampilkan.

Munculnya internet memicu kelahiran berbagai mesin pencari, bersamaan dengan perkembangan awal *World Wide Web* pada tahun 1990. Perkembangan ini kemudian disusul oleh hadirnya *Web Crawler* pada tahun 1994 yang mulai memanfaatkan pola klik pengguna dan log kueri untuk menyaring hasil pencarian. Setelah itu, berbagai mesin pencari seperti Archie, AltaVista, dan Google bermunculan, membawa inovasi seperti PageRank yang menggunakan hyperlink sebagai penanda relevansi. Saat ini, *Information Retrieval* (IR) terus berkembang dengan dukungan kecerdasan buatan, pemahaman semantik, dan algoritme yang semakin maju untuk menghasilkan pencarian informasi yang lebih cepat dan akurat. Abad ke-20 hingga awal abad ke-21 menjadi periode penting yang mengubah cara manusia mengakses dan memperoleh informasi (Sanderson & Croft, 2016).

Sistem Temu Kembali Informasi merupakan sistem yang dirancang untuk menyimpan, mencari, serta mengelola berbagai bentuk informasi. Jenis informasi yang ditangani sangat beragam, mulai dari teks (termasuk data numerik), hingga audio, video, dan bentuk multimedia lainnya. Proses pemodelan informasi menjadi aspek yang sangat penting karena membantu peneliti dalam merancang dan mengembangkan sistem informasi yang benar-benar efisien ketika digunakan (Mabrouk et al., 2017).

Dalam proses temu kembali informasi, sistem perlu mengekstraksi dokumen-dokumen yang relevan dari kumpulan data yang besar dan kompleks. Seiring bertambahnya volume data serta meningkatnya kebutuhan pengguna akan hasil pencarian yang lebih akurat, metode pencarian tradisional menjadi kurang mampu menyediakan hasil yang memenuhi standar kualitas yang diharapkan (Jun-Feng & Wei-Ming, 2019). Berbagai penelitian yang berbeda tentang Sistem Temu Balik Informasi telah dilakukan dengan tujuan yang sama yaitu meningkatkan relevansi temu kembali informasi.

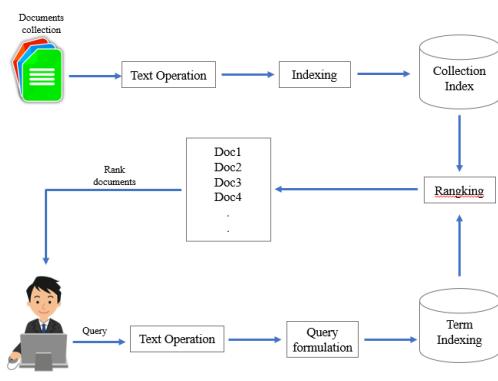
Dalam Sistem Temu Kembali Informasi, pencarian dokumen umumnya dilakukan menggunakan kata kunci atau *query*. Salah satu permasalahan yang sering muncul ialah ketidakmampuan pengguna dalam merumuskan kebutuhan informasinya secara tepat melalui *query*, sehingga hasil pencarian tidak sesuai dengan harapan. Jumlah dokumen relevan yang ditemukan sangat dipengaruhi oleh kualitas *query* yang diberikan. Ketidaksesuaian kosakata antara kebutuhan pengguna dan dokumen yang tersimpan juga dapat mengakibatkan tidak ditemukannya dokumen apa pun. Oleh karena itu, sistem temu kembali informasi yang efektif harus mampu mengatasi perbedaan istilah serta kesenjangan bahasa antara *query* dan dokumen. Salah satu pendekatan yang dapat digunakan untuk mengatasi masalah tersebut adalah *query expansion*, yaitu teknik yang memperluas atau menambah kata dalam *query* untuk meningkatkan kecocokan antara *query* dan dokumen (Pal et al., 2015).

Ekspansi *query* (*Query expansion*) adalah teknik yang digunakan untuk meningkatkan presisi pencarian. Ekspansi *Query* diyakini dapat mengatasi masalah yang terkait dengan representasi *query* pengguna dengan memodifikasi *query* untuk meningkatkan kualitas *query* pengguna (Ooi et al., 2015). Ide dasar penerapan ekspansi *query* yaitu dengan menggunakan hasil dari *query* awal untuk memformulasi ulang *query* dan melakukan penelusuran *second-pass* untuk mendapatkan hasil presisi yang lebih tinggi (Piramuthu et al., 2013).

Terdapat berbagai macam metode yang sudah dilakukan untuk memperluas kueri pada temu balik informasi. Penelitian ini difokuskan untuk mengkaji secara komprehensif literatur terkait metode-metode *query expansion* yang telah dikembangkan. Kajian ini diharapkan dapat memberikan gambaran yang jelas mengenai perkembangan terkini dalam bidang temu balik informasi sekaligus menjadi dasar bagi penelitian lebih lanjut dalam merancang sistem pencarian yang lebih adaptif, akurat, dan mampu memenuhi kebutuhan pengguna secara optimal.

2. INFORMATION RETRIEVAL SYSTEM

Information Retrieval (IR) merupakan suatu cara yang digunakan untuk menemukan kembali (*retrieve*) informasi-informasi yang relevan terhadap kebutuhan pengguna dari suatu kumpulan informasi secara otomatis. Salah satu aplikasi umum dari IR system adalah *search engine* atau mesin pencarian yang terdapat pada jaringan internet. Pengguna dapat mencari halaman-halaman *web* yang dibutuhkannya melalui *search engine*. Contoh lain dari IR system adalah sistem informasi perpustakaan. IR system terutama berhubungan dengan pencarian informasi yang isinya tidak memiliki struktur. Ekspresi kebutuhan pengguna yang disebut *query*, juga tidak memiliki struktur. Hal ini yang membedakan IR system dengan sistem basis data. Dokumen adalah contoh informasi yang tidak terstruktur. Isi dari suatu dokumen sangat tergantung pada pembuat dokumen tersebut. Sebagai suatu sistem, IR system memiliki beberapa bagian yang membangun sistem secara keseluruhan. Adapun representasi yang menjelaskan tentang IR system pada Gambar 1.



Gambar 1. Proses Sistem Temu Kembali Informasi

Berikut penjelasan yang merujuk pada Gambar 1 tentang bagian dari IR system (Bunyamin, 2005):

1. *Text Operations* (operasi terhadap teks) yang meliputi pemilihan kata-kata dalam *query* maupun dokumen (*term selection*) dalam pentransformasian dokumen atau *query* menjadi *term index* (indeks dari kata-kata).
2. *Query formulation* (formulasi terhadap *query*) yaitu memberi bobot pada indeks kata-kata *query*.
3. *Rangking* (perangkingan), mencari dokumen-dokumen yang relevan terhadap *query* dan mengurutkan dokumen tersebut berdasarkan kesesuaianya dengan *query*.
4. *Indexing* (pengindeksan), membangun basis data indeks dari koleksi dokumen. Dilakukan terlebih dahulu sebelum pencarian dokumen dilakukan.

3. QUERY EXPANSION

Query merupakan formulasi kebutuhan informasi pengguna yang merujuk pada permintaan atau pertanyaan yang diajukan oleh pengguna untuk mencari informasi tertentu dari suatu koleksi dokumen atau basis data. *Query* berbasis kata kunci sangat populer digunakan (Baeza-Yates & Ribeiro-Neto, 1999). *Query Expansion* adalah proses merumuskan kembali *query* untuk meningkatkan kinerja pencarian dalam operasi pencarian informasi. Ada cukup banyak peneliti yang sangat percaya bahwa ketidakakuratan *query* yang dibentuk oleh beberapa model kata kunci adalah alasan utama ketidakefektifan sistem temu kembali informasi (Carpineto & Romano, 2012).

Menambahkan istilah tambahan ke dalam *query* dapat dilakukan secara otomatis, manual, atau dengan bantuan pengguna. Perluasan kueri manual bergantung pada masukan pengguna untuk memutuskan istilah mana yang akan ditambahkan ke *query* asli, sedangkan perluasan kueri otomatis adalah teknik yang mengandalkan penimbangan istilah. Istilah dengan bobot tertinggi akan ditambahkan ke *query* asli (Ooi et al., 2015).

Beberapa metode yang digunakan untuk memperluas kueri (*query expansion*) antara lain:

1. *Thesaurus*

Tesaurus (*Thesaurus*) adalah sebuah daftar kata atau istilah yang terorganisir dan

disusun secara hierarkis serta menunjukkan hubungan semantik antar kata (asosiatif, atau kesetaraan) (Kombolo et al., 2022). *Thesaurus* merupakan sebuah rujukan yang berisi kata-kata yang memiliki pertalian makna berupa kesamaan, perlawanan, hubungan superordinat dan subordinat, atau hubungan antarbagian. Kesamaan makna (sinonim) dapat digunakan untuk memperkaya *query* pada proses Temu Kembali Informasi. Diharapkan dengan pemanfaatan sinonim untuk memperkaya *query* dapat menemukan dokumen yang lebih banyak dibanding sebelum menggunakan pengayaan *query* (Putra et al., 2018).

2. WordNet

WordNet adalah sebuah kamus yang merupakan pengembangan dari kamus kata *Thesaurus*, dikembangkan oleh Universitas Princeton dan berisi pemodelan leksikal Bahasa Inggris. Didalam *WordNet* terdapat *synset* (*synonym set*), didalamnya terdapat kata-kata yang memiliki keterkaitan dan relasi *semantic* dalam berbagai kategori kata seperti kata benda, kata kerja, kata sifat, dan kata keterangan (Zhang et al., 2009)

3. Association Rule

Association Rule merupakan salah satu teknik yang ada dalam *data mining* yang bertujuan untuk mendapatkan aturan asosiasi atau hubungan antar sekumpulan *item*. Aturan asosiasi dapat diperoleh dari berbagai sumber data, termasuk yang berasal dari database transaksional, gudang data, serta dari area penyimpanan informasi lainnya. Secara umum, data yang diolah adalah homogen (Gunawan & Mustofa, 2016). Langkah pertama pencarian aturan asosiasi diperoleh dari *item set* yang sering muncul bersamaan. Pentingnya suatu aturan asosiasi dapat diketahui dengan dua parameter yaitu *Support* dan *Confidence*. *Support* adalah persentase terjadinya kombinasi *item* atau *support count* dari jumlah *item* yang muncul dalam satu set transaksi, dan kepercayaan adalah hubungan yang kuat antara *item* dalam aturan asosiasi. Analisis asosiasi didefinisikan sebagai proses untuk menemukan semua aturan asosiatif yang memenuhi persyaratan dukungan minimum, dan persyaratan kepercayaan minimum (Afuan et al., 2019).

Support adalah ukuran yang menunjukkan tingkat dominasi dari probabilitas *item* atau *set item* dalam *database* transaksional. Persamaan (1) merupakan formula dalam menentukan nilai dari *support*.

$$Support = \frac{(X \cup Y)count}{n} \quad (1)$$

Dimana $(X \cup Y)count$ adalah jumlah transaksi yang mengandung x dan y , dan n adalah total transaksi. Parameter lainnya adalah *confidence* yaitu nilai ukuran seberapa besar valid tidaknya suatu *Association Rules*. Sebuah *Association Rules* dengan *confidence* sama atau lebih besar dari minimum *confidence* dapat dikatakan sebagai valid *association rule*. Persamaan (2) untuk menentukan nilai *confidence*, dimana $X count$ adalah jumlah kemunculan *item X* pada seluruh transaksi. Persamaan (2) merupakan formula dalam menentukan *confidence*.

$$Confident = \frac{(X \cup Y)count}{X.count} \quad (2)$$

4. Relevance Feedback

Relevance feedback, teknik ini memodifikasi suatu *query* yang sering diimplementasikan pada *information retrieval*. Cara kerjanya ialah dengan memilih *term* penting dalam dokumen yang dianggap dokumen relevan oleh pengguna, serta menambahkan *term* penting ke dalam proses modifikasi *query* (Putri et al., 2019).

Berikut ini merupakan metode *relevance feedback* (Pamungkas et al., 2015), diantaranya:

a. User Judgement (*Explicit Relevance Feedback*)

Pada metode ini membutuhkan penilaian relevansi dokumen pada *query* tertentu. Penilaian yang diberikan, yaitu dengan penilaian terhadap suatu dokumen dan menentukan mana dokumen relevan maupun tidak.

b. User Behavior (*Implicit Relevance Feedback*)

Metode ini berhubungan dengan perilaku pengguna, yaitu seperti mencatat dokumen terpilih maupun tidak dan durasi waktu untuk melihat dokumen maupun proses selama

- melakukan pencarian hingga *scrolling* halaman.
- c. *Top K Relevance (Blind/Pseudo Relevance Feedback)* Metode ini hanya melibatkan pengguna saat proses keputusan dan dokumen relevan yang ditampilkan secara otomatis, karena tanpa memerlukan *feedback* dari pengguna.

4. EVALUASI INFORMATION RETRIEVAL

Dalam bidang *Information Retrieval* (IR), evaluasi dilakukan untuk menilai efektivitas suatu sistem dalam menemukan serta menampilkan informasi yang sesuai dengan kebutuhan pengguna. Proses evaluasi ini merupakan aspek mendasar yang digunakan untuk menilai kinerja sekaligus membandingkan berbagai model IR yang dikembangkan. Evaluasi dilakukan untuk mengukur dan menganalisis kinerja dari model yang telah dikembangkan.

Evaluasi peforma efektivitas pada sistem klasifikasi teks dengan menggunakan standar *confussion matrix* yang berisi informasi klasifikasi yang sebenarnya dan merupakan prediksi klasifikasi oleh sistem (Pamungkas et al., 2015)

Tabel 1. *Confusion Matrix*

	<i>Actual Positive</i>	<i>Actual Negative</i>
<i>Predicted Positive</i>	<i>True Positive</i>	<i>False Positive</i>
<i>Predicted Negative</i>	<i>False Negative</i>	<i>True Negative</i>

Keterangan:

True Positive (TP): menunjukkan perangkingan sistem merupakan dokumen yang sesuai dengan *query*.

False Prositive (FP): menunjukkan dokumen dalam hasil perangkingan sistem tidak sesuai dengan *query*.

False Negative (FN): menunjukkan dokumen tidak termasuk dalam perangkingan sistem dan harusnya sesuai dengan *query*.

True Negative (TN): menunjukkan dokumen tidak termasuk perangkingan sistem dan memang seharusnya tidak sesuai *query*.

Hasil dari temu kembali berupa dokumen akan dilakukan pengukuran dengan menggunakan *precision* dan *recall*. *Recall* adalah proporsi

jumlah dokumen yang dapat ditemukan kembali oleh sebuah proses pencarian di sistem IR. *Precision* adalah proporsi jumlah dokumen yang ditemukan dan dianggap relevan untuk kebutuhan si pencari informasi. *Recall* (perolehan) berhubungan dengan kemampuan sistem untuk memanggil dokumen yang relevan. *Precision* (ketepatan) berkaitan dengan kemampuan sistem untuk tidak memanggil dokumen yang tidak relevan (Lestari, 2016).

$$Precision = \left[\frac{TP}{TP+FP} \right] \quad (3)$$

$$Recall = \left[\frac{TP}{TP+FN} \right] \quad (4)$$

KESIMPULAN

Information retrieval (IR) merupakan bidang yang terus berkembang dan sangat penting dalam era digital saat ini. Sistem IR memungkinkan pencarian dan pengambilan informasi yang relevan dari kumpulan data besar, baik dalam bentuk teks maupun multimedia, dengan tujuan utama memenuhi kebutuhan informasi pengguna secara efisien dan akurat.

Berdasarkan studi literatur yang telah dilakukan, Salah satu teknik lanjutan pada sistem temu balik informasi adalah perluasan kueri atau *query expansion*.

Query expansion dapat dilakukan dengan menambahkan istilah tambahan ke dalam *query* dapat dilakukan secara otomatis, manual, atau dengan bantuan pengguna. Perluasan kueri manual bergantung pada masukan pengguna untuk memutuskan istilah mana yang akan ditambahkan ke *query* asli, sedangkan perluasan kueri otomatis adalah teknik yang mengandalkan penimbangan istilah. Istilah dengan bobot tertinggi akan ditambahkan ke *query* asli.

Beberapa metode yang dapat digunakan untuk *expansion query* secara otomatis yaitu *thesaurus* dengan memanfaatkan daftar kata yang terorganisir sinonim bisa memperkaya *query*. *WordNet* yang merupakan perkembangan dari *thesaurus* yang memiliki sinonim set. *Association rule* metode yang melihat keterkaitan antar kata yang menjadikan *query* menjadi beragam, dan *relevance feedback* yang menggunakan top K dokumen relevan untuk memperkaya *query*.

PERNYATAAN PENGHARGAAN

Segala puji dan rasa syukur Penulis haturkan kepada Tuhan Yang Maha Esa atas limpahan rahmat, kekuatan, dan kesempatan sehingga penelitian ini dapat diselesaikan dengan baik. Penulis menyadari bahwa terselesaikannya penelitian ini tidak terlepas dari dukungan, bantuan, serta kontribusi berbagai pihak. Oleh sebab itu, dengan penuh ketulusan Penulis menyampaikan apresiasi dan ucapan terima kasih yang mendalam kepada semua pihak yang telah memberikan bantuan, baik dalam bentuk bimbingan, saran, maupun informasi yang sangat berharga selama proses penyusunan penelitian ini. Penulis juga memahami bahwa tidak semua pihak dapat disebutkan satu per satu, namun setiap bentuk dukungan yang diberikan memiliki arti penting dalam kelancaran penelitian ini.

Pada akhirnya, Penulis mengucapkan terima kasih kepada seluruh pihak yang telah berperan dalam proses penyusunan penelitian ini. Besar harapan Penulis agar hasil penelitian ini dapat memberikan manfaat, memperluas wawasan, serta menjadi kontribusi positif bagi perkembangan dunia pendidikan dan penelitian di masa yang akan datang.

DAFTAR PUSTAKA

- Afuan, L., Ashari, A., & Suyanto, Y. (2019). Query expansion in information retrieval using frequent pattern (FP) growth algorithm for frequent itemset search and association rules mining. *International Journal of Advanced Computer Science and Applications*, 10(2), 263–267. <https://doi.org/10.14569/ijacsa.2019.0100235>
- Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern information retrieval*.
- Bunyamin, H. (2005). Algoritma Umum Pencarian Informasi Dalam Sistem Temu Kembali Informasi Berbasis Metode Vektorisasi Kata dan Dokumen. *Jurnal Informatika*, 1(2), 85–92.
- Carpinetto, C., & Romano, G. (2012). A survey of automatic query expansion in information retrieval. *ACM Computing Surveys*, 44(1). <https://doi.org/10.1145/2071389.2071390>
- Gunawan, R., & Mustofa, K. (2016). Pencarian Aturan Asosiasi Semantic Web Untuk Obat Tradisional Indonesia. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi (JNTETI)*, 5(3), 192–200. <https://doi.org/10.22146/jnteti.v5i3.256>
- Jun-Feng, S., & Wei-Ming, Z. (2019). Research on Ontology Based Information Retrieval Model. *Journal of Nanjing* http://en.cnki.com.cn/Article_en/CJFDTO-TAL-NJDZ200502011.htm
- Kombolo, M., Yon, J., Landrieu, F., Richon, B., Aubin, S., & Hocquette, J.-F. (2022). A new semantic resource responding to the principles of open science: The meat thesaurus as an IT tool for dialogue between sector actors. *Meat Science*, 192(February), 108849. <https://doi.org/10.1016/j.meatsci.2022.108849>
- Lestari, N. P. (2016). Uji Recall and Precision Sistem Temu Kembali. *Universitas Airlangga*.
- Mabrouk, D., Rady, S., BadrNagwa, & Khalifa, M. (2017). *A survey on Information Retrieval System's Modeling Using Term Dependencies and Term Weighting*. *Icicsi*, 321–328.
- Ooi, J., Ma, X., Qin, H., & Liew, S. C. (2015a). A survey of query expansion, query suggestion and query refinement techniques. *2015 4th International Conference on Software Engineering and Computer Systems, ICSECS 2015: Virtuous Software Solutions for Big Data*, August, 112–117. <https://doi.org/10.1109/ICSECS.2015.733094>
- Ooi, J., Ma, X., Qin, H., & Liew, S. C. (2015b). A survey of query expansion, query suggestion and query refinement techniques. *2015 4th International Conference on Software Engineering and Computer Systems, ICSECS 2015: Virtuous Software Solutions for Big Data*, October, 112–117. <https://doi.org/10.1109/ICSECS.2015.733094>
- Pal, D., Mitra, M., & Bhattacharya, S. (2015). *Exploring Query Categorisation for Query Expansion: A Study*. September. <http://arxiv.org/abs/1509.05567>
- Pamungkas, Z. Y., Indriati, & Ridok, A. (2015). Query Expansion pada Sistem Temu Kembali Informasi Dokumen Berbahasa Indonesia Menggunakan Pseudo Relevance Feedback. *Jurnal Pengembangan Teknologi Informasi Dan*

- Ilmu Komputer (J-PTIIK) Universitas Brawijaya.*
- Piramuthu, R., Bhardwaj, A., Di, W., & Sundaresan, N. (2013). Visual search: A large-scale perspective. In *Handbook of Statistics* (Vol. 31). Elsevier B.V. <https://doi.org/10.1016/B978-0-444-53859-8.00011-4>
- Putra, F. N., Effendi, A., & Arifin, A. Z. (2018). Pembobotan Kata pada Query Expansion dengan Tesaurus dalam Pencarian Dokumen Bahasa Indonesia. *Linguistik Komputasional*, 1.
- Putri, C. A. A., Indriati, & Supianto, A. A. (2019). Query Expansion Pada LINE TODAY Dengan Algoritme Extended Rocchio Relevance Feedback. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 3, No 1, 177–185.
- Sanderson, M., & Croft, W. B. (2016). The History of Information Retrieval. *Journal of the ACM*.
- Zhang, J., Deng, B., & Li, X. (2009). Concept Based Query Expansion Using WordNet. *Beijing: Dept. Electronic Engineering, Tsinghua Univ, Beijing, 100084, China*.