

Penerapan Model *Information Retrieval* Untuk Pencarian Konten Pada Perpustakaan Digital

Eka Fitriani¹, Richardus Eko Indrajit², Riska Aryanti³

¹ STMIK Nusa Mandiri
Jakarta, Indonesia
Fitranieka817@gmail.com

² ABFI Institute Perbanas
Jakarta, Indonesia
indrajit@post.harvard.edu

³ STMIK Nusa Mandiri
Jakarta, Indonesia
Riskaary22@gmail.com

Abstract - Efforts to improve the quality of human resources is to develop the interest of users to read. Libraries have a great responsibility for improving the user in reading habits therefore the digital library is very positive berpengaruh to users in reading. Information and communication technology (ICT) has become an integral part of global life. Therefore the need for a library with a new look and facilitate users to obtain information from libraries. Digital library is the source of user information for library data search. The development of technology makes it easier for users to get the information they need. Digital libraries have a lot of content that causes search difficulties based on the interests of users. By using the *N-Gram Method* and *Spelling Correction index* applied in the search program will allow users to find the information needed, especially interest-based search. This search process utilizes text information stored in digital library documents. In text documents, information is obtained directly from the content of the document in search of the digital library.

Kata Kunci: *Perpustakaan Digital, information retrieval, N-Gram, Spelling Correction.*

I. PENDAHULUAN

Perkembangan dalam dunia perpustakaan dari segi penyimpanan data dan dokumen yang masih menggunakan perpustakaan tradisional yang hanya terdiri dari kumpulan koleksi-koleksi buku tanpa adanya katalog dari perpustakaan tersebut. Dengan berjalannya perkembangan teknologi informasi yang begitu pesat dalam halnya dunia perpustakaan saat ini muncul yang disebut perpustakaan digital.

Perpustakaan mempunyai tanggungjawab yang besar untuk peningkatan pengguna dalam kebiasaan membaca maka dari itu perpustakaan digital ini sangatlah berpengaruh positif terhadap pengguna dalam gemar membaca.

Perpustakaan digital adalah perpustakaan yang mempunyai koleksi buku-buku yang sebagian besar tampilannya dalam format digital dan bisa di akses dengan komputer.

Perpustakaan digital ini sangat bermanfaat untuk para pengguna dalam pemberian informasi dari sumber-sumber yang berasal dari koleksi perpustakaan digital tersebut. Perpustakaan digital ini secara ekonomis lebih menguntungkan di bandingkan perpustakaan tradisional.

Oleh karena itu berkaitan dengan pencarian konten pada perpustakaan digital ini , dengan menggunakan metoda *N-Gram* adalah solusi yang tepat karena pengecekan mulai dari konten itu tersebut. Temu balik informasi akan memberikan informasi-informasi yang relevan untuk memenuhi kebutuhan pengguna dalam pencarian konten. Pengguna dapat mengakses dan mendownload konten yang di perlukannya.

A. *Information Retrieval*

Temu balik informasi merupakan sebuah metode untuk mengambil data terstruktur yang tersimpan dalam koleksi dokumen, kemudian menyediakan informasi

yang diperlukan. Tujuan dari Information Retrieval yaitu untuk mengambil informasi serta menampilkan dokumen yang relevan dengan user dalam jumlah yang sangat besar (query) (Witanti et al., 2016).

Temu balik informasi memiliki keunggulan untuk menemukan informasi secara optimal sehingga secara otomatis kebutuhan user dapat terpenuhi. Pada prinsipnya temu *information retrieval* dapat memformulasikan sekumpulan dokumen dari sebuah pernyataan (request atau query) (Yang & Ada, 2016).

Temu balik informasi adalah sebagai media layanan untuk memperoleh informasi atau sumber yang dibutuhkan oleh pengguna. Temu balik informasi berfungsi sebagai perantara kebutuhan informasi pengguna dengan sumber yang tersedia.

B. Perpustakaan Digital

Perpustakaan digital adalah selingkup materi multimedia dalam bentuk digital yang sangat tertata rapi, ditujukan untuk keuntungan populasi pengguna yang terstruktur, untuk memfasilitasi akses pada isi (content) dan dilengkapi dengan fasilitas untuk melakukan navigasi ke jaringan global (Collier 1997). (Winarko, 2009).

Pemanfaatan teknologi ini khususnya pada perpustakaan digital diakui mampu meningkatkan efisiensi dalam pengelolaan pada perpustakaan dan selanjutnya dapat memberikan kemudahan bagi pustakawan dalam perpustakaan tersebut. Seperti yang sudah diketahui bahwa kini sudah memasuki "era quantum" era di mana perubahan pengetahuan, cara hidup, kebudayaan, dan peradaban manusia memiliki kecepatan dan akselerasi yang tinggi (Medan, 2006).

Perkembangan perpustakaan digital bagi pengelola perpustakaan dapat membantu pekerjaan di perpustakaan melalui fungsi otomatisasi perpustakaan, sehingga proses pengelolaan perpustakaan lebih efektif dan efisien. Fungsi otomatisasi perpustakaan menitikberatkan pada bagaimana mengontrol sistem administrasi layanan secara otomatis atau terkomputerisasi. Sedangkan bagi pengguna dapat membantu mencari sumber informasi yang diinginkan dengan menggunakan catalog on-line yang dapat diakses melalui intranet maupun internet, sehingga pencarian informasi dapat dilakukan kapan dan dimana pun ia berada (Dana et al., 2008).

Perpustakaan merupakan pusat penyedia informasi yang multifungsi, sehingga dikenal sebagai unit pelayanan informasi yang sangat dibutuhkan oleh

masyarakat secara luas karena perpustakaan memang terikat dalam tatanan masyarakat yang luas. Pesatnya teknologi maka dibuatkan perpustakaan online untuk lebih mempermudah atau mempercepat pengguna dalam mencari buku atau mendapatkan informasi.

C. Metode N-Gram

N-Gram adalah metode populer yang dapat diimplementasikan dalam pengembangan kata ataupun sebuah karakter. Metode ini digunakan untuk mengambil penggalan karakter huruf sejumlah n dalam sebuah kata secara berkesinambungan (Andri, Ariana, & Andriani, 2014).

Untuk melakukan pengecekan Ejaan yang sesuai dengan EYD dalam sebuah dokumen meliputi berbagai langkah. Proses input dokumen yang akan diperiksa oleh metode N-gram yaitu dengan cara memarsing berbagai kalimat yang ada dalam dokumen. (Andri et al., 2014).

D. Information Retrieval menggunakan Metode N-Gram

Segala jenis informasi dan sumber-sumber informasi yang dimiliki oleh perpustakaan secara terus-menerus akan bertambah sejalan dengan berkembangnya informasi. Pada perpustakaan digital ini maka diperlukan sistem temu balik informasi.

Information retrieval yang akan mempermudah dalam pencarian data dibutuhkan suatu algoritma yang dapat dengan mudah membantu dalam proses pencariannya diterapkan algoritma *information retrieval* dengan metoda N-Gram dan Spelling Correction untuk pengkoreksian dari konten yang telah di sediakan dalam database.

II. METODOLOGI PENELITIAN

Dalam penelitian ini metode penelitian yaitu dengan mengumpulkan data-data penulis dibutuhkan untuk menyelesaikan penulisan ini, penulis menggunakan metode studi pustaka dengan mencari dari berbagai referensi, mengumpulkan data-data, mempelajari dan menganalisa data referensi tersebut secara langsung dari buku yang berhubungan dengan penelitian yang akan di bahas pada penulisan ini.

2.1. Populasi Data

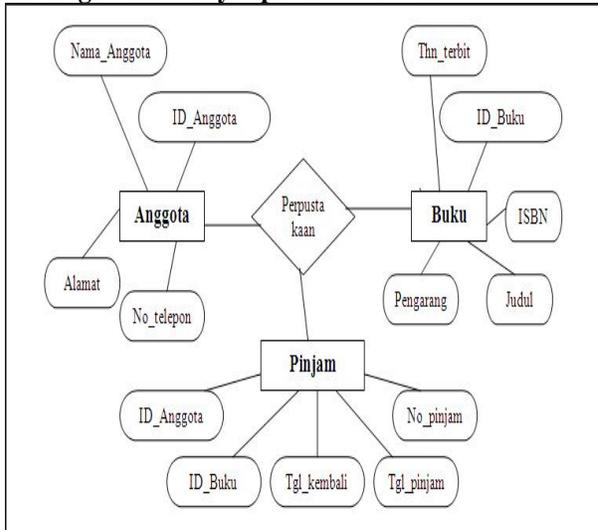
Populasi data atau data yang di kumpulkan pada penelitian ini adalah berhubungan dengan data-data perpustakaan, seperti dari data buku, berdasarkan dari

populasi atau sampel data buku perpustakaan itu tersebut.

2.2. Format dan Karakteristik Data

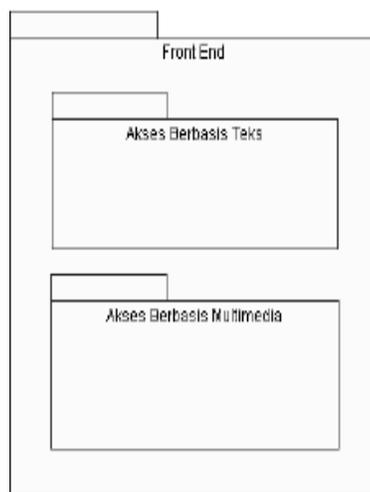
Format data yang digunakan pada pencarian ini adalah non terstruktur, dalam perpustakaan bisa berdasarkan dari nama penulis, judul buku, tahun terbit. Jadi dalam karakteristiknya bisa menggunakan data-data perpustakaan.

2.3 Organisasi Penyimpanan Informasi



Gambar 1. Rancangan ERD

Pencarian Dokumen Teks dan Multimedia

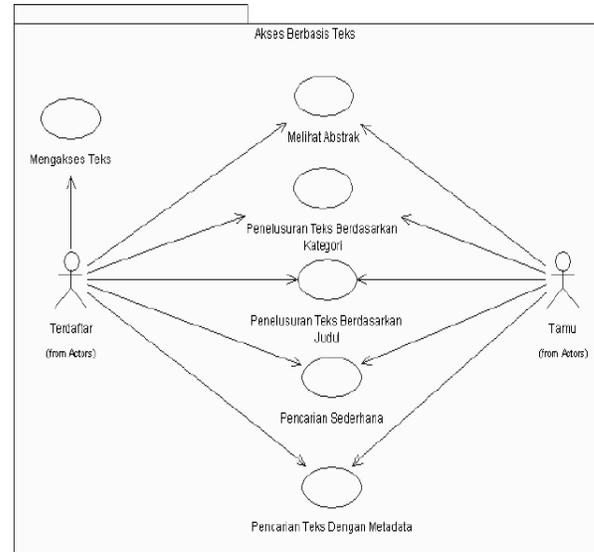


Gambar 2. Front End

Gambar 2 menunjukkan gambaran front end dari sistem perpustakaan digital untuk proses pengaksesan dan pencarian untuk dokumen teks dan multimedia.

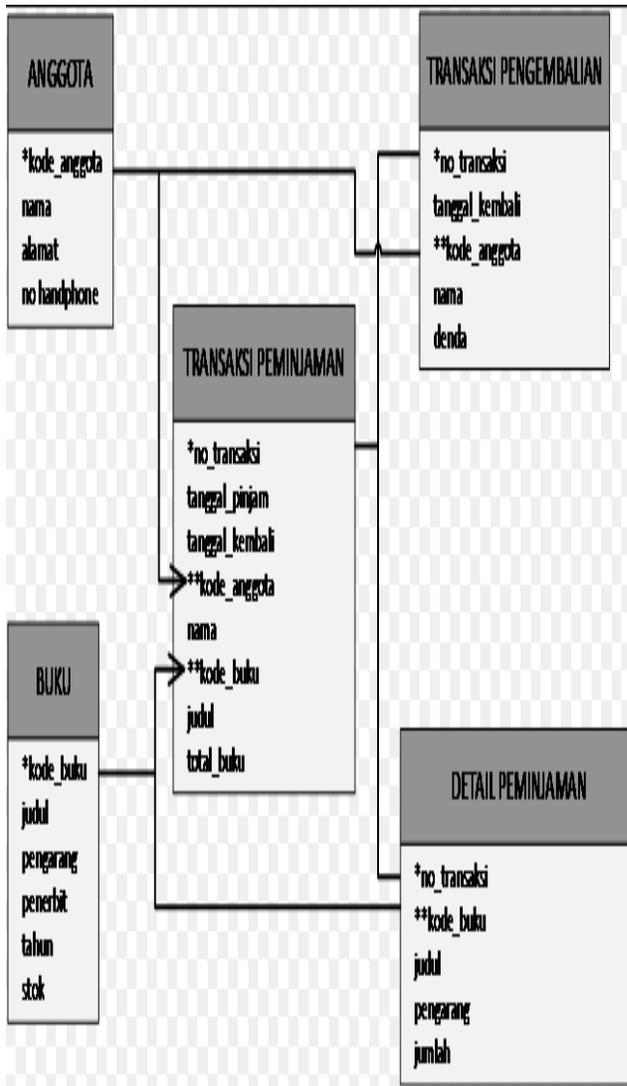
Sistem ini dikembangkan dengan database SQL Server[2].

Pengguna yang dapat mengakses *front end* adalah pengguna terdaftar dan pengguna tamu. Pengguna terdaftar adalah pengguna yang harus terautentikasi terlebih dahulu. Pengguna tamu adalah pengguna yang tidak perlu diautentikasi atau dengan kata lain tidak melakukan login ke sistem.

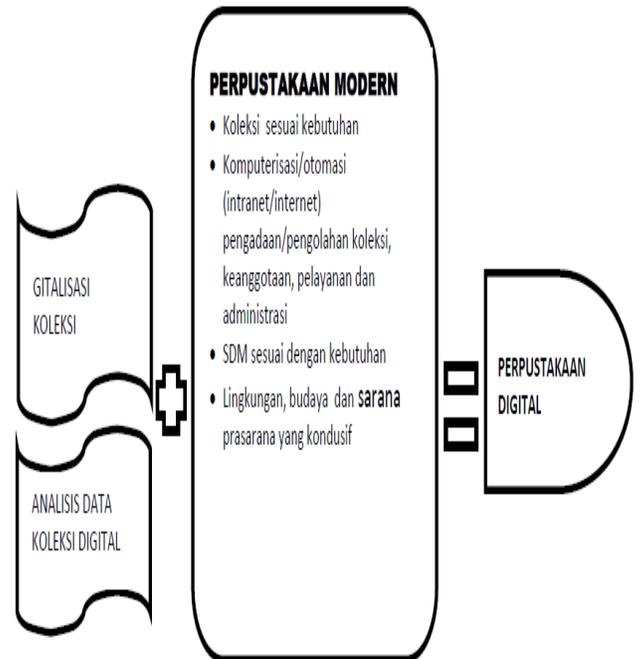


Gambar 3. Use Case Diagram Akses Berbasis Teks

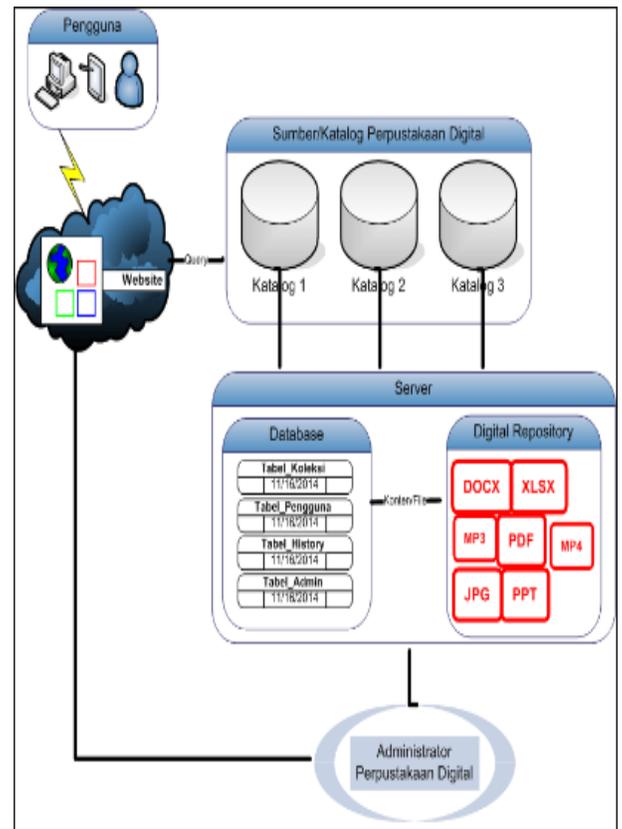
Pada Gambar 3.dapat dilihat kalau pengguna terdaftar dan pengguna tamu dapat melihat abstrak dokumen teks, melakukan penelusuran teks berdasarkan kategori, melakukan penelusuran teks berdasarkan judul, melakukan pencarian sederhana dan pencarian teks dengan metadata. Hasil dari proses pencarian ini adalah daftar dokumen teks yang dicari. Untuk melihat isi dokumen secara keseluruhan, hanya dapat dilakukan oleh pengguna terdaftar melalui use case mengakses teks.



Gambar 4. LRS Perpustakaan



Gambar 5. Perancangan Perpustakaan digital



Gambar 6. Skema Pencarian Pada Perpustakaan

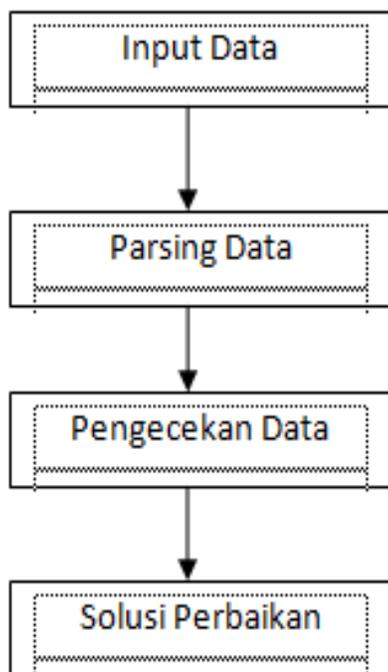
2.4 Sintaks Pencarian Informasi

Sintaks pada perpustakaan ini menggunakan kata kunci dari konten perpustakaan, dapat memasukkan nama penulis dan judul buku. Pencarian akan dilakukan berdasarkan pencarian yang sering banyak dicari itu akan muncul otomatis.

2.5 Algoritma Pencarian

Algoritma dalam pencarian ini :

Langkah pertama adalah dengan melakukan proses input data yang akan diperiksa, Langkah kedua adalah data yang telah diinputkan akan dilakukan proses parsing yaitu proses memecah kalimat-kalimat yang ada dalam dokumen sehingga membentuk kumpulan-kumpulan kata, Langkah ketiga adalah akan dilakukan proses pengecekan kata pada data tersebut. Setelah proses perbandingan kata dalam data dengan kata yang ada akan diberikan solusi perbaikan. Langkah keempat adalah proses solusi dari hasil pengecekan data yang telah di inputkan. Secara garis besar alur dari proses koreksi dapat dilihat pada Gambar 3.



Gambar 7. Alur Proses Koreksi

2.6. Kehandalan Algoritma

Dalam pencarian konten menggunakan metode N-Gram adanya pengkoreksian per ejaan kata, sehingga kata yang di cari harus berdasarkan perbandingan dengan konten yang sesuai di database. Dengan user memasukan kata kunci untuk pencarian , maka sistem akan segera merespon pencarian tersebut berdasarkan key yang di masukan dengan lebih cepat dan akurat.

III. HASIL DAN PEMBAHASAN

3.1. Implementasi

Implementasi dari sistem perpustakaan digital ini, disini adanya data buku, data anggota, pencarian buku.

kdbarang	nmbarang	pengarang	penerbit	tterbit
BK001	Visual Basic	Ulus Rusmak	Elex Media	2010
BK002	Pegnantar B	Dr.Francis T	Raja Grafinc	2009
BK003	Bisnis Mode	Sentot Imam	Graha Ilmu	2010
BK004	EDP Audit	Marmah Har	Erlangga	2010
BK005	Sistem Infor	Deni Dama	Rosdaa	2013
BK006	C++ Builder	Imam Heryal	Informatika	2015

Gambar 8. Penginputan Data Buku

	kdanggota	Nama	Alamat	Tlp
	AG001	Audi Hendriar	Perum Karaba	08569236709
	AG002	Anita Wulandi	Perum Peruri	08125678234
	AG003	Emi Rustiani	Perum Karaba	08774567345
	AG004	M Ihsan A	Grand Taruma	08576789456
	AG007	Rachmawati	Wadas	09875678456
	AG006	Risky Indiani	Wadass	09875678456

Gambar 9. Penginputan data anggota

kdbarang	nmbarang	pengarang	penerbit	lterbit
BK001	Visual Basic	Uus Rusmak	Elex Media	2010

Gambar 10. Pencarian pada perpustakaan

No	Titik	Format	ISBN/ISSN	Detail
1	Akuntansi syariah di Indonesia (Risalah - Ruyyat, Si)		978770113206	Detail
2	Microsoft excel 2010 dan Myeji untuk membuat aplikasi akuntansi		978770120526	Detail
3	Akuntansi untuk perusahaan jasa (Maha, Evi)		979107892	Detail
4	Metode memahami jimat penyelesaian (Prabowo, Theano Ai - Purba, Despoti Rosdani)		978770212268	Detail
5	MYOB V.18 penyelesaian kasus pendekatan bukti transaksi (Tondro, Anwar - Saputra, Ryan Bastian)		978770293124	Detail
6	Akuntansi sektor publik (organisasi non laba) (Renyonojo, Murnido)		9786027523367	Detail
7	Akuntansi kreatif, utang dan modal		9786028446031	Detail

Gambar 11. Katalog Perpustakaan

Proses pencarian berbasis konten untuk dokumen berbasis teks

Penelitian ini memanfaatkan indeks dalam melakukan pencarian dokumen. Pencarian tersebut dilakukan berdasarkan kata kunci yang dimasukkan oleh pengguna. Mesin pencari melakukan proses pencarian kata dalam *index* sehingga dapat ditemukan pada dokumen mana kata tersebut muncul. Selanjutnya, dokumen hasil pencarian yang telah diurutkan berdasarkan skor ditampilkan pada layar. Semakin besar skor suatu dokumen, semakin besar pula relevansi dokumen tersebut dengan kata kunci yang dimasukkan.

Uji coba pencarian dokumen berbasis konten pada sistem ini dilakukan dengan memasukkan berbagai kata kunci sebagai kriteria pencarian. Dari proses tersebut didapatkan kondisi apakah pencarian dapat menghasilkan dokumen yang sesuai dengan kata kunci atau tidak. Dokumen hasil pencarian merupakan dokumen yang telah tersimpan dalam *server* sistem.

IV. KESIMPULAN

Dari hasil pengujian pada pencarian konten pada perpustakaan digital ini, didapatkan kesimpulan : Dapat mempercepat dalam proses pencarian karena adanya pengoreksian pada setiap konten yang dicari, Pemanfaatan mesin pencari dengan pengindekan dapat mempercepat proses pencarian karena waktu yang dibutuhkan mesin pencari hanya ½ dari waktu yang dibutuhkan pencarian normal, Pencarian dokumen teks berbasis konten dengan memanfaatkan mesin pencari dapat memudahkan untuk mengetahui relevansi antara dokumen hasil pencarian dan kata kunci.

REFERENSI

Andri, Ariana, S., & Andriani, M. (2014). Aplikasi Koreksi Kesalahan Berbasis Pada Tulisan Berbahasa Indonesia Untuk Meningkatkan Kualitas Penulisan Karya Ilmiah.

Seminar Nasional Aplikasi Dan Teknologi,
(November), 169–172.

Dana, T., Teknik, J., Fakultas, I., Industri, T., Atmajaya, U.,
& Information, P. (2008). PENGEMBANGAN
DIGITAL LIBRARY PERPUSTAKAAN
UNIVERSITAS ATMAJAYA YOGYAKARTA,
2008(semnasIF), 260–263.

Medan, U. N. (2006). Perpustakaan Digital sebagai Wujud
Penerapan Teknologi Informasi di Perguruan Tinggi
Supriyanto Jurusan Sejarah, 2(1), 32–36.

Winarko, B. (2009). PERPUSTAKAAN DIGITAL DI
INDONESIA DAN FITUR-FITUR YANG TERSEDIA
Bambang Winarko 1) 2), 18(20), 45–49.

Witanti, W., Rahmanto, H., Renaldi, F., Informatika, P. S.,
Jenderal, U., Yani, A., ... Negation, C. (2016).
PEMBANGUNAN SISTEM TEMU BALIK
INFORMASI (INFORMATION RETRIEVAL)
DALAM PEMILIHAN PEMAIN SEPAK BOLA
BERKUALITAS, 2016(Sentika), 18–19.

Yang, P., & Ada, T. (2016). INFORMATION RETRIEVAL
DOKUMEN TESIS UNTUK, 12(2), 105–115.