p-ISSN: 1410-9840 & e-ISSN: 2580-8850 http://journals.usm.ac.id/index.php/jprt/index

SISTEM TEMU KEMBALI INFORMASI UNTUK PENCARIAN JUDUL TUGAS AKHIR BERBASIS KATA KUNCI

B. Very Christioko¹, April Firman Daru²

¹, dan ² Program Studi Teknik Informatika Universitas Semarang ¹ very@usm.ac.id, ² firman@usm.ac.id

Abstrak

Kesamaan atau kemiripan judul pada saat mahasiswa mengambil mata kuliah Tugas Akhir (TA) dapat mungkin terjadi. Hal ini merupakan masalah yang serius dalam penelitian bagi mahasiswa saat mengerjakan TA atau skripsi.Saat ini di Fakultas TIK, data judul TA yang sudah selesai masih hanya disimpan ke dalam sebuah basis data pada saat pengajuan dan belum digunakan sebagai sumber rujukan untuk mengecek kesamaan atau kemiripan. Untuk itu dibutuhkan sebuah alat untuk menemukan informasi judul-judul yang sudah ada pada database menggunakan mesin pencari. Dikarenakan mesin pencari merupakan bentuk praktek dari penerapan teknik temu kembali informasi pada koleksi text dengan skala besar maka perlu adanya sebuah mesin pencari judul Tugas Akhir pada Fakultas TIK yang menerapkan teknik temu kembali informasi dengan model *Boolean retrieval*.

Kata Kunci: Boolean retrieval, Mesin Pencari, Teknik temu kembali.

Abstract

The similarity or resemblance of the title at the time students take the Final Task subject may be possible. This is a serious problem in research for students while working on Final Task or minithesis. Currently in the Faculty of Information Technology and Communication, the completed Final Task title data is still only stored into a database at the time of submission and has not been used as a referral source to check for similarities or similarities. For that we need a tool to find information titles that already exist in the database using search engines. Because the search engine is a form of practice of applying information retrieval techniques on large-scale text collections it is necessary to have a search engine title Final Task on the Faculty of Information Technology and Communication that apply information retrieval techniques with Boolean retrieval model.

Keywords: Boolean retrieval, Search engine, Technique of Retrieval

I. Latar Belakang

Kualitas tertinggi dari sebuah karya tulis diukur dari orisinalitas atau keaslian. Keaslian yang diamaksudkan disini lebih pada kejujuran dalam mengemukakan tulisannya. Tidak semua tulisan dalam karya ilmiah adalah berasal dari diri sendiri melainkan milik orang lain yang berupa kata, kalimat, paragraph, ide, gagasan atau pendapat, maka secara etika ilmiah kita wajib mencantumkan sumber dengan jujur dan objektif dari mana mendapatkan kutipan tersebut, bila ini tidak dilakukan maka akan memunculkan masalah plagiasi (Mulyana, 2010).

Plagiasi merupakan masalah serius dalam penelitian termasuk bagi mahasiswa saat mengerjakan mata kuliah Tugas Akhir (TA) atau skripsi.TA merupakan mata kuliah yang harus diambil mahasiswa sebelum lulus kuliah. Untuk dapat memulai mengerjakan TA mahasiswa harus mengajukan judul yang tidak memiliki kesamaan atau kemiripan dengan judul TA yang sudah ada untuk menghindari Plagiasi. Permasalahan yang terjadi adalah belum adanya sistem untuk melakukan deteksi kesamaan atau kemiripan judul TA. Saat ini pengelolaan data judul Tugas Akhir masih hanya disimpan ke dalam sebuah

basis data pada saat pengajuan dan belum digunakan sebagai sumber rujukan untuk mengecek kesamaan atau kemiripan.

Perkembangan internet saat ini menjadikan kebutuhan untuk menemukan informasi secara cepat dan tepat semakin tinggi.Kebutuhan ini mendorong orang-orang di perusahaan maupun universitas menggunakan sebuah mesin pencari untuk membantu menemukan informasi baik berupa text, gambar maupun video.Mesin pencari merupakan bentuk praktek dari penerapan teknik temu kembali informasi pada koleksi text dengan skala besar. Komponen dalam mesin pencari mendukung 2 (dua) fungsi utama yaitu indexing process dan query process, indexing process membangun struktur yang memungkinkan pencarian, sedang query process menggunakan struktur-struktur tersebut dan permintaan pengguna untuk menghasilkan daftar peringkatdokumen atau teks (W. Bruce Crof, 2015). Boolean retrieval merupakan salah satu model temu kembali informasi yang digunakan pada mesin pencari. Model ini juga biasa dikenal dengan istilah exact-match retrieval dikarenakan dokumen atau teks akan ditemukan jika benar-benar sesuai dengan spesifikasi pencarian (W. Bruce Crof, 2015). Model Boolean retrieval ini sesuai digunakan untuk melakukan pencarian judul-judul penelitian maupun karya ilmiah seperti yang sudah dilakukan oleh Christioko (2012).

Oleh karena itu mesin pencari untuk menemukan judul TA menjadi sebuah kebutuhan dengan tujuan menghindari kesamaan atau kemiripan.

II. KAJIAN PUSTAKA

2.1. Plagiasi Tugas Akhir

Mulyana (2010) menyatakan bahwa plagiasi dalam dunia akademik (khususnya dosen dan mahasiswa) sangat mungkin terjadi apabila etika ilmiah dalam mencantumkan sumber dengan jujur dan objektif dimana kutipan-kutipan berupa kata, kalimat, paragraph, ide, gagasan, atau pendapat yang merupakan milik orang lain tidak dilakukan. Maka civitas akademika diharapkan memegang teguh etika akademik yang berlaku. Berikut merupakan hasil penelitian yang dilakukan Mulyana tentang jenis plagiarism Tugas Akhir Skripsi (TAS) Mahasiswa.

Tabel 1. Jenis Plagiarisme TAS Mahasiswa

Kasus	Jenis Plagiarisme	Tindakan Pencegahan
Plagiarisme Skripsi	Duplikasi judul Duplikasi substansi Duplikasi teori Duplikasi data Duplikasi referensi	Mengembalikan usulan TAS dan atau merevisi Mengharuskan pengutipan primer Membatalkan karya tulis

Plagiarisme dalam bentuk duplikasi judul menyangkut kesamaan atau kemiripan judul dari karya tulis mahasiswa lain. Artinya, beberapa proposal dan atau skripsi dapat mempunyai kesamaan atau kemiripan dalam hal judul, topik, data, dan kajian teori.

2.2. Mesin Pencari

Sebuah mesin pencari adalah penerapan praktis dari teknik pengambilan informasi untuk koleksi teks dengan skala besar.Istilah "Mesin Pencari" awalnya digunakan untuk mengacu ke perangkat keras khusus untuk pencarian teks.Mulai pertengahan tahun 1980-an, secara bertahap mulai merujuk ke "information retrieval system" sebagai nama dari perangkat lunak yang melakukan proses membandingkan dokumen dan menghasilkan daftar ranking dokumen (Crof, 2015).

2.3. Arsitektur Mesin Pencari

Sebuah arsitektur dari mesin pencari didesain untuk memastikan bahwa sistem akan memuaskan kebutuhan atau tujuan (Crof, 2015). Dua tujuan utama dari mesin pencari adalah:

- 1. Effectiveness (quality),kemampuan uuntuk mengambil kumpulan dokumen yang paling relevan dalamm sebuah permintaan.
- 2. *Efficiency (speed)*, kemampuan untuk memproses permintaan dari user secepat mungkin.

Komponen mesin pencari mendukung dua fungsi utama, yang biasa disebut dengan *indexing process* dan *query process.Indexing process* membangun struktur yang memungkinkan untuk pencarian, dan *query process* menggunakan struktur tersebut dan permintaan pengguna untuk memproduksi daftar ranking dokumen. Gambar berikut menunjukkan bagan dari *indexing process*.

Komponen utama dari queryprocess userinteraction, ranking, dan evaluation. Komponen user interaction menyediakan antarmuka antara pengguna yang melakukan pencarian dan mesin pencari.Tugas utama komponen ini adalah menerima query pengguna dan mengubahnya ke bentuk index terms. Tugas lain adalah mengambil daftar dokumen terangking dari mesin pencari dan mengatur hasilnya untuk dilihat oleh pengguna. Komponen ranking merupakan inti dari mesin pencari. Komponen ini membawa query dari pengguna yang sudah diterjemahkan dan menciptakan daftar ranking dari dokumen menggunakan basis skor pada model temu kembali. Sedangkan komponen evaluation bertugas untuk mengukur dan memonitor effectiveness dan efficiency.

2.4. Temu Kembali Informasi (Information Retrieval)

Menurut Jansen, B. J. and Rieh, S. (2010) Sebuah objek adalah sebuah entitas yang diwakili oleh informasi dalam basis data . User queries dicocokkan database informasi. Hasil information dengan retrieval(IR) yang dikembalikan mungkin sesuai dengan query, sehingga iasanya hasilnya diperingkat. Peringkat ini berasal dari hasil perbedaan kunci pencarian informasi pencarian dibandingkan dengan mencari database. Sedangkan pengertian IR menurut Manning et al, (2008) adalah "Information retrieval (IR) is finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers)". Dapat disimpulkan bahwa IR adalah proses menemukan materi yang biasanya dokumen dari data yang tidak terstruktur yang memenuhi kebutuhan informasi dari sebuah kumpulan koleksi yang tersimpan pada komputer. Berikut skema konsep dasar dari sebuah IR

Sistem Temu-Balik Informasi (Information Retrieval) digunakan untuk menemukan kembali informasi-informasi yang relevan terhadap kebutuhan pengguna dari suatu kumpulan informasi secara otomatis. Salah satu aplikasi umum dari sistem temu kembali informasi adalah search-engine atau mesin pencarian yang terdapat pada jaringan internet. Pengguna dapat mencari halaman-halaman Web yang dibutuhkannya melalui mesin tersebut.

Ukuran efektifitas pencarian ditentukan oleh precision dan recall. *Precision* adalah rasio jumlah dokumen relevan yang ditemukan dengan total jumlah dokumen yang ditemukan oleh *search-engine*.

Precision mengindikasikan kualitas himpunan jawaban, tetapi tidak memandang total jumlah dokumen yang relevan dalam kumpulan dokumen.

$$Precision = \frac{|\{relevant\ documents\} \cap \{documents\ retrieved\}|}{|\{documents\ retrieved\}|}$$

Recall adalah rasio jumlah dokumen relevan yang ditemukan kembali dengan total jumlah dokumen dalam kumpulan dokumen yang dianggap relevan.

$$\text{Recall} = \frac{|\{\text{relevant documents}\} \cap \{\text{documents retrieved}\}|}{|\{\text{relevant documents}\}|}$$

Dalam Information Retrieval, mendapatkan dokumen yang relevan tidaklah cukup. Tujuan yang harus dipenuhi adalah bagaimana mendapatkan dokumen relevan dan tidak mendapatkan dokumen yang tidak relevan. Tujuan lainnya adalah bagaimana menyusun dokumen yang telah didapatkan tersebut ditampilkan terurut dari dokumen yang memiliki tingkat relevansi lebih tingi ke tingkat relevansi rendah. Penyusunan dokumen terurut tersebut disebut sebagai perangkingan dokumen. Model Ruang Vektor dan Model Probabilistik adalah 2 model pendekatan untuk melakukan hal tersebut.

Model ruang vektor dan model probabilistik adalah model yang menggunakan pembobotan kata dan perangkingan dokumen. Hasil retrieval yang didapat dari model-model ini adalah dokumen terangking yang dianggap paling relevan terhadap query.Terdapat beberapa cara atau metode dalam melakukan pembobotan kata pada metode TF-IDF, yaitu melalui skema pembobotan query dan dokumen. Dalam model ruang vektor, dokumen dan query direpresentasikan sebagai vektor dalam dalam ruang vektor yang disusun dalam indeks term, kemudian dimodelkan dengan persamaan geometri. Sedangkan model probabilistik membuat asumsi-asumsi distribusi term dalam dokumen relevan dan tidak relevan dalam order estimasi kemungkinan relevansi suatu dokumen terhadap suatu query.

2.5. Model Boolean Retrieval (exact-matchretrieval)

Model Boolean retrieval adalah model temu kembali informasi yang mengekspresikan query menggunakan bentuk ekspresi Boolean seperti and, or, dan not (Manning et al, 2008). Dalam model ini diperlukan sebuah proses pengindekan sehingga akan membentuk sebuah matrik *term-document incidence*. Pengindekan adalah operasi offline untuk mengumpulkan data yang berisi kata yang terdapat pada teks, sehingga pada saat pencarian kita hanya perlu mengakses index. Contoh berikut adalah matrik dari kumpulan dokumen berupa karya dari Shakespeare. Elemen Matrix (t, d) bernilai 1 jika peran pada kolom d mengandung kata pada baris t, bernilai 0 jika sebaliknya.

	Antony and Cleopatra	Julius Caesar	The Tempest	Hamlet	Othello	Macbeth
Antony	1	1	0	0	0	1
Brutus	1	1	0	1	0	0
Caesar	1	1	0	1	1	1
Calpurnia	0	1	0	0	0	0
Cleopatra	1	0	0	0	0	0
mercy	1	0	1	1	1	1
worser	1	0	1	1	1	0

Gambar1. Contoh matrik term-document incidence

Selanjutnya model ini memproses query dari pengguna yang menggunakan ekspresi Boolean yaitu AND, OR dan NOT. Berikut contoh pemrosesan query oleh model Boolean Retrieval.

Query: "Brutus AND Caesar AND NOT Calpunia"

Tahap 1: Index Awal

	Antony and Cleopatra	Julius Caesar	The Tempest	Hamlet	Othello	Macbeth
Antony	1	1	0	0	0	1
Brutus	1	1	0	1	0	0
Caesar	1	1	0	1	1	1
Calpurnia	0	1	0	0	0	0
Cleopatra	1	0	0	0	0	0
mercy	1	0	1	1	1	1
worser	1	0	1	1	1	0

Tahap 2: Operasi NOT dijalankan

	Antony and Cleopatra	Julius Caesar	The Tempest	Hamlet	Othello	Macbeth
Antony	1	1	0	0	0	1
Brutus	1	1	0	1	0	0
Caesar	1	1	0	1	1	1
¬Calpurnia	1	0	1	1	1	1
Cleopatra	1	0	0	0	0	0
mercy	1	0	1	1	1	1
worser	1	0	1	1	1	0

Tahap 3: Perhitungan operasi AND

	Antony and Cleopatra	Julius Caesar	The Tempest	Hamlet	Othello	Macbeth
Antony	1	1	0	0	0	1
Brutus	1	1	0	1	0	0
Caesar	1	1	0	1	1	1
¬Calpurnia	1	0	1	1	1	1
Cleopatra	1	0	0	0	0	0
mercy	1	0	1	1	1	1
worser	1	0	1	1	1	0
AND	1	0	0	1	0	0

Hasil dari perhitungan operasi AND diatas mengembalikan dokumen "Antony and Cleopatra" dan "Hamlet".

III. METODE PENELITIAN

3.1. Metode Pengumpulan Data

Teknik yang digunakan dalam pengumpulan data untuk keperluan penelitian adalah sebagai berikut:

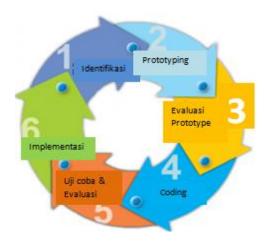
- a. Wawancara dan observasi, teknik ini digunakan untuk mendapatkan data primer yaitu melalui wawancara dengan petugas perpustakaandi fakultas teknologi informasi dan komunikasi untuk mendapatkan data judul-judul tugas akhir yang sudah tercatat di perpustakaan.
- b. Studi pustaka, teknik ini digunakan untuk mendapatkan data sekunder tentang obyek penelitian dan referensi tentang teori pendukung dan metode melalui buku-buku dan majalah ilmiah.

3.2. MetodePengembangan Sistem

Metode pengembangan sistem yang digunakan dalam penelitian ini adalah metode *prototype*. Metode *prototype* adalah proses dimana permintaan pengguna diubah ke dalam sistem yang bekerja, secara terus menerus diperbaiki melalui kerja sama antara pengguna dan pembuat sistem.

Metode ini sering digunakan dalam pengembangan sistem, karena terdapat interaksi yang baik antara pengguna dan pembuat sistem.

Sering terjadi *gap* antara pengguna dan pembuat sistem. Dimana pengguna hanya mendefinisikan secara umum apa saja yang dikehendakinya tanpa menyebutkan secara detail, output apa yang dibutuhkan, pemrosesan dan data apa saja yang dibutuhkan, kendala lapangan, dan lain sebagianya. Dan sebaliknya, dari sisi pembuat sistem, kurang memperhatikan kondisi kemampuan pengguna, kondisi lapangan, dan kebutuhan pengguna secara detail.



Gambar 2. Model Prototype

Dengan menggunakan metode *prototype*, pengguna dilibatkan secara aktif dalam proses pengembangan sistem.

Adapun tahap – tahap dalam metode prototype ini adalah sebagai berikut :

1. Identifikasi kebutuhan

Pembuat sistem, bersama sama dengan pengguna mendifinisikan format seluruh perangkat lunak, mengidentifikasikan seluruh kebutuhan, dan garis besar sistem yang akan dibuat.

2. Membangun prototyping

Adalah proses pembuatan rancangan sementara yang berfokus pada penyajian kepada pelanggan.

3. Evaluasi prototyping

Langkah ini melibatkan pengguna. Dimana pengguna akan memberikan penilaian, apakah prototype nya sudah sesuai dengan kebutuhan dan keinginan pelanggan. Jika sudah , maka pembuat sistem akan melanjutkan ke langkah berikutnya. Jika belum, maka akan diulang dari langkah 1,2 dan 3.

4. Coding

Pada tahap ini, setelah semua desain sesuai dengan kebutuhan pengguna, maka akan diterjemahkan ke dalam bahasa pemrograman yang akan dipakai dalam sistem.

5. Menguji Sistem dan Evaluasi

Setelah program bantu jadi, maka diujicobakan pada sistem. Dan diamati apakah sudah sesuai dan sudah dapat mengatasi kebutuhan serta kendala lapangan. Uij coba ini menjadi evaluasi bersama antara pengguna dan pembuat sistem. Jika dirasa masih ada yan kurang sesuai, maka akan diulang dari langkah ke 4. Jika sudah cocok, dilanjutkan dengan tahap ke 6.

6. Implementasi Sistem

Program bantu yang sudah jadi dan diterima oleh pengguna, siap diimplementasikan.

IV. HASIL & PEMBAHASAN

4.1. Data Penelitian

Dalam penelitian ini data yang peneliti kumpulkan adalah judul-judul Tugas Akhir (TA) dengan stuktur sebagai berikut:

Tabel 2. Struktur judul TA

Field	Keterangan					
Kategori	Kategori judul (Tugas					
	Akhir/TA,Kerja Praktek/KP)					
Progdi	Program Studi (Teknik					
	Informatika/TI, Sistem					
	Informasi/SI)					
Tahun	Tahun TA disahkan					
Nama	Nama mahasiswa					
Nim	Nim mahasiswa					
Judul	Judul TA					
Pembimbing 1	Pembimbing 1 TA					
Pembimbing 2	Pembimbing 2 TA					

Data judul yang peneliti peroleh dari petugas perpustakaan adalah data dari tahun 2007 – 2017 sebanyak 1695 record dengan judul TA dari progdi Teknik Informatika (TI)sebanyak 1272 record dan progdi Sistem Informasi (SI) sebanyak 423 record. Berikut sebaran data judul yang peneliti peroleh.

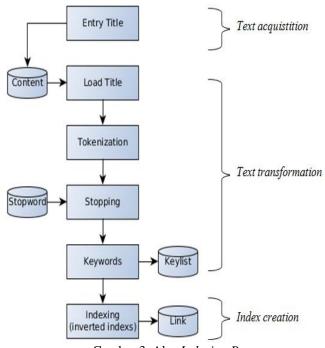
Tabel 3. Sebaran	data	sesuai	tahun	dan	program	studi
------------------	------	--------	-------	-----	---------	-------

Tahun Judul	Program Studi			
TA	TI	SI		
2007	10	3		
2008	12	7		
2009	31	8		
2010	39	9		
2011	28	7		
2012	75	23		
2013	173	38		
2014	256	89		
2015	285	87		
2016	271	102		
2017	92	50		
Jumlah	1272	423		

4.2. Komponen Mesin Pencari

4.2.1. *Indexing Process*

Indexing process merupakan fungsi utama dalam komponen mesin pencari yang mempunyai fungsi untuk membangun struktur yang memungkinkan untuk pencarian. *Indexing process* terdiri dari 3 komponen utama yaitu, *text acquistition*, *text transformation*, *dan index creation*.



Gambar 3. Alur *Indexing Process*

4.2.1.1. Text Acquisition

Tugas utama dari text acquistition adalah untuk mengidentifikasi dan membuat dokumen untuk siap dicari. Komponen ini akan lebih sering membangun koleksi berupa document data store pada komputer yang mengandung teks dan metadata untuk semua dokumen. Untuk persiapkan Document data store judul-judul Tugas Akhir dibuat sebuah tabel pada server MySQL dengan struktur sebagai berikut:

Nama Tabel : Content

Tujuan : Untuk menyimpan data-data judul

Tabel 4. Struktur Tabel Content

Nama Field	Tipe	Keterangan
contid	mediumint(9)	Id content
prodi	enum('TI', 'SI')	Nama program studi
kategori	enum('TA', 'KP')	Kategori judul
tahun	int(4)	Tahun judul
nama	varchar(100)	Nama mahasiswa
nim	char(13)	Nim mahasiswa
judul	varchar(255)	Nama judul
pembimbing1	varchar(100)	Pembimbing ke-1
pembimbing2	varchar(100)	Pembimbing ke-2
parsing	enum('Y', 'T')	Status parsing

Terdapat 2 (dua) metode untuk membuat document data sore judul-judul pada tabel "Content" yaitu:

Import Judul TA ke tabel 'Content' di MySQL Server.

Terdapat 2 (dua) metode untuk membuat document data sore judul-judul pada tabel "Content" yaitu:

Import Judul TA ke tabel 'Content' di MySQL Server.

C D kategortahun	E nama	F	
	nama		
			judul
	JANU ARIYANTO WIBOWO		Pembelajaran Multimedia pengenalan bahan - bahan kimia dalam pro
	RYFKI MAULANA		Pemanfaatan Sensor cahaya untuk menghitung jumlah Orang Dalam
	YOSEP AGUNG PRAMUNDAYA		sistem informasi Administrasi pembayaran sekolah pada SMP MARSU[
	WAHYU IKHA SULISTYANTO	G.211.02.0022	TUTORIAL PEMBUATAN ANIMASI IKLAN BANNER DENGAN MACROMEDI
TA 2007	Wida Adhi Hapsara	G.211.02.0001	Pemanfaatan Macromedia Flash sebagai Media Informasi Dalam Peran
	BUDI SETIAWAN	G.241.05.0016	sistem informasi CRM (Customer Relationship Management) Berbasis
TA 2007	ARI ANDHINI	G.211.02.0011	Macromedia Flash MX Profesional 2004 Sebagai Sarana Pembelajaran
	USWATUN HASANAH	G.211.02.0024	Penggunaan Macromedia flash Mx sebagai sarana pembelajaran Anat
TA 2007	EPI NOPIANTI	G.211.02.0023	pembelajaran pijat refleksi dengan menggunakan BORLAND - DELPHI
TA 2007	Galih Septyawan	G.211.02.0007	Multimedia sebagai sarana pembelajaran Bahasa Inggris Mengenai Ar
TA 2008	CATUR ADI PURNOMO		UJIAN MASUK MAHASISWA BERBASIS KOMPUTER MENGGUNAKAN MAP
TA 2008	MUHAMAD AMRULLAH	G.211.02.0016	Sistem pembelajaran berbasis WEB pada Fkultas Hukum UNIVERSITA
TA 2008	UDI HERVIANTO	G.211.03.0023	Papan Informasi Kehadiran pada Ruang Dekan Fakultas TIK. Berbasis N
TA 2008	HARISANTO ARIF SUBENO	G.241.04.0009	Sistem Informasi Manajemen Sumber Daya Manusia pada PT. MITRA P
TA 2008	Rochmad Adiono Utomo	G.241.05.0017	penyandian file gambar BMP dengan metode Enkripsi & DEKRIPSI
TA 2008	VIENDA WINTARI		sistem informasi penggajian karyawan pada SMPN 1 MRANGGEN DEM
TA 2008	WISNU ADHI MAHENDRA		PROGRAM BANTU PELAJARAN PENGENALAN TO BE dalam Bahasa Ingg
TA 2008	YUDI PRAMONO	G.211.02.0005	Game Sibonie berbasis Macromedia Flash untuk Sekolah Dasar
TA 2008	BIMO HARYO SETYOKO	G.241.05.0008	pengendalian penenangan rumah melalui port paralel menggunakan
TA 2008	AHMAD QOHAR WAHYUDDIN	G.211.02.0004	PENGENALAN KOSAKATA BAHASA INDONESIA, BAHASA INGGRIS, DAN
TA 2008	ROCHMAD WIBOWO	G.241.05.0005	Sistem Informasi Keuangan Sekolah di SMA MUHAMMADIYAH SALATIG
TA 2008	R. Gatot NURHADI Priyono		Pembelajaran interaktif menarik komputer menggunakan SWISHMAX
TA 2009	HAFIS MADYASTA		Sistem Informasi Geografis Letak SPBU Yang Sudah Terakreditasi ' Pasl
TA 2009	YAZID RAHMANULLAH		pembelajaran jaringan komputer menggunakan Swistimax 2004
	TTA 2007 TTA 2007 TTA 2007 TTA 2007 TTA 2007 TTA 2007 TTA 2008	TA 2007 RYFKI MAULANA TA 2007 YOSEP AGUNG PRAMUNDAYA TA 2007 WAHYU IKHA SULISTYANTO TA 2007 Wida Adhi Hapsara TA 2007 BUDI SETIAWAN TA 2007 ARI ANDHINI TA 2007 EPI NOPIANTI TA 2007 Galih Septyawan TA 2008 CATUR ADI PURNOMO TA 2008 MUHAMAD AMRULLAH TA 2008 HARISANTO ARIF SUBENO TA 2008 ROCHMAD WINTARI TA 2008 WISNU ADHI MAHENDRA TA 2008 YUDI PRAMONO TA 2008 BIMO HARIYO SETYOKO TA 2008 AHMAD QOHAR WAHYUDDIN TA 2008 ROCHMAD QOHAR WAHYUDDIN TA 2008 R. Gatot NURHADI Priyono TA 2008 R. Gatot NURHADI Priyono TA 2008 R. Gatot NURHADI Priyono	TA 2007 RYFKI MAULANA G.211.03.0018 TA 2007 YOSEP AGUNG PRAMUNDAYA G.241.05.0003 TA 2007 WAHYU IKHA SULISTYANTO G.211.02.0022 TA 2007 Mda Adhi Hapsara G.211.02.0001 TA 2007 BUDI SETIAWAN G.241.05.0016 TA 2007 ARI ANDHINI G.211.02.0011 TA 2007 LUSWATUN HASANAH G.211.02.0024 TA 2007 EPI NOPIANTI G.211.02.0023 TA 2007 Galih Septyawan G.211.02.0007 TA 2008 CATUR ADI PURNOMO G.241.05.0009 TA 2008 MUHAMAD AMRULLAH G.211.02.0016 TA 2008 WILLAM G.211.02.0016 TA 2008 ROCHMAD AMRULLAH G.211.02.0016 TA 2008 ROCHMAD AGUNG G.241.05.0010 TA 2008 ROCHMAD AGUNG G.241.05.0017 TA 2008 ROCHMAD AGUNG G.241.05.0017 TA 2008 WISNU ADHI MAHENDRA G.211.02.0017 TA 2008 WISNU ADHI MAHENDRA G.211.02.0017 TA 2008 BIMO HARYO SETYOKO G.241.05.0008 TA 2008 BIMO HARYO SETYOKO G.241.05.0008 TA 2008 ROCHMAD AGUNG G.241.05.0008 TA 2008 ROCHMAD AGUNG G.241.05.0008 TA 2008 BIMO HARYO SETYOKO G.241.05.0008 TA 2008 ROCHMAD WIBOWO G.241.05.0005 TA 2008 ROCHMAD WIBOWO G.241.05.0005 TA 2008 R GAGON WIBOWO G.241.05.0005

Gambar 4. Data judul pada spreadsheet

Data judul-judul diatas kemudian diimport ke dalam tabel untuk membentuk koleksi dokumen judul-judul TA. Gambar di bawah ini merupakan tampilan hasil import ke dalam tabel "Content".

contid	prodi	kategori	tahun	nama	nim	judul
1	TI	TA	2007	JANU ARIYANTO WIBOWO	G.211.02.0006	Pembelajaran Multimedia pengenalan bahan - bahan k
2	TI	TA	2007	RYFKI MAULANA	G.211.03.0018	Pemanfaatan Sensor cahaya untuk menghitung jumlah
3	TI	TA	2007	YOSEP AGUNG PRAMUNDAYA	G.241.05.0003	sistem informasi Administrasi pembayaran sekolah p
4	TI	TA	2007	WAHYU IKHA SULISTYANTO	G.211.02.0022	TUTORIAL PEMBUATAN ANIMASI IKLAN BANNER DENGAN MAC
5	TI	TA	2007	Wida Adhi Hapsara	G.211.02.0001	Pemanfaatan Macromedia Flash sebagai Media Informa
6	TI	TA	2007	BUDI SETIAWAN	G.241.05.0016	sistem informasi CRM (Customer Relationship Manage
7	TI	TA	2007	ARI ANDHINI	G.211.02.0011	Macromedia Flash MX Profesional 2004 Sebagai Saran
8	TI	TA	2007	USWATUN HASANAH	G.211.02.0024	Penggunaan Macromedia flash Mx sebagai sarana pemb
9	TI	TA	2007	EPI NOPIANTI	G.211.02.0023	pembelajaran pijat refleksi dengan menggunakan BOR
10	TI	TA	2007	Galih Septyawan	G.211.02.0007	Multimedia sebagai sarana pembelajaran Bahasa Ingg
11	TI	TA	2008	CATUR ADI PURNOMO	G.241.05.0009	UJIAN MASUK MAHASISWA BERBASIS KOMPUTER MENGGUNAKA
12	TI	TA	2008	MUHAMAD AMRULLAH	G.211.02.0016	Sistem pembelajaran berbasis WEB pada Fkultas Huku

Gambar 5. Data judul pada document data store

Input ke tabel melalui halaman input data judul TA pada sistem

Selain membentuk koleksi dokumen judul dengan melakukan import, user juga disediakan sebuah form untuk entri judul seperti gambar di bawah ini.



Gambar 6. Form input judul TA

4.2.1.2 Text Transformation

Komponen Text Transformation bertugas untuk merubah kumpulan dokumen ke dalam bentuk index terms berupa kata-kata. Metode yang digunakan adalah Tokenization Stoppingyaitu memotong judul menjadi beberapa bagian yang disebut tokendengan menghilangkan kata-kata umum (stop words). Daftar stop words ini disimpan ke dalam tabel "Stopword" yang kemudian digunakan untuk menghilangkan kata-kata umum pada saat Tokenization. Hasil dari Tokenization berupa index terms disimpan ke dalam tabel "Keylist". Berikut struktur tabel untuk tahap Text Tranformation.

a. Stopword

Struktur tabel untuk menyimpan kata-kata umum (kata buangan) yang digunakan sebagai stopword dalam membentuk index terms adalah sebagai berikut:

Nama Tabel : Stopword

Tujuan : Untuk menyimpan kata-kata

umum

Tabel 5. Struktur Tabel Stopword

Nama Field	Tipe	Keterangan
id	int(2)	Id stopword
word	Varchar(255)	Kata umum
no	Int(1)	Jumlah

Berikut beberapa kata-kata umum (stopword) yang akan dihilangkan dari judul-judul TA pada proses *Tokenization*sesuai dengan Tala (2003):

Keywords (inde terms)

ada	diantara	Melalui	pun	para
adanya	apa	Terlalu	mau	pasti
adalah	apaan	Lama	maupun	pastilah
adapun	mengapa	Lamanya	semaunya	per
agak	apabila	selama	memang	pernah
agaknya	apakah	selama-	mereka	pula
agar	apalagi	lamanya selamanya	merekalah	

Struktur tabel untuk menyimpan *keyword* hasil *tokenization* adalah sebagai berikut:

Nama Tabel : Keylist

Tujuan : Untuk menyimpan keyword

Tabel 6. Struktur Tabel Content

Nama Field	Tipe	Keterangan
Keyed	mediumint(9)	Id key
keyword	varchar(100)	keyword

Contoh proses text transformation

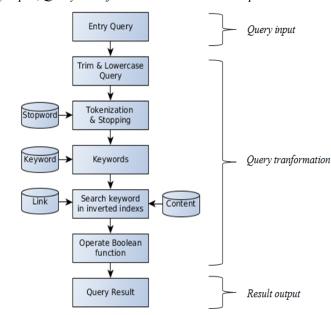
Judul TA:sistem informasi penggajian karyawan pada SMPN 1 MRANGGEN DEMAK dengan menggunakan Visual Basic 6.0

Tahap 1:	Tahap 2:	Tahap 3:
<u>Tokenization</u>	Stopping	<u>keywords</u>
sistem	sistem	sistem
informasi	informasi	informasi
penggajian	penggajian	penggajian
karyawan	karyawan	karyawan
pada	pada	smpn
smpn	smpn	mranggen
1	1	demak
mranggen	mranggen	visual

demak	demak	basic
dengan	dengan	
menggunakan	menggunakan	
visual	visual	
basic	basic	
6.0	6.0	

4.2.2. Query Process

Query process tersusun dari komponen *user interface* yang menyediakan antarmuka antara sesorang yang melakukan pencarian dan mesin pencari. Salah satu tugas utamanya adalah menerima query pengguna dan menerjemahkannya ke dalam bentuk index terms (*keyword*). Komponen ini terdiri dari *Query input*, *Query transformation* dan *Results Output*.



Gambar 7. Alur Querying Process

b. Query input

Komponen *query input* menyediakan antarmuka dan *parser* untuk *query*. Untuk melakukan pencarian user mengakses halaman pencarian judul kemudian memasukkan kata kunci dan memilih operator *Boolean* (AND atau OR).



Gambar 8. Halaman pencarian

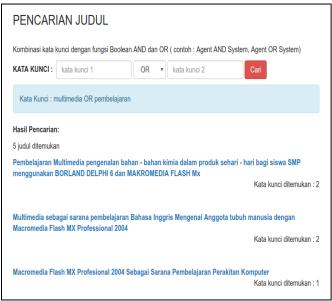
c. Query Transformation

Komponen query transformation menggunakan proses sederhana yang melibatkan metode pemrosesan teks untuk mentransformasi *query* menjadi keywordkeyword yang kemudian dibandingkan dengan indeks (inverted indexs) kemudian menerapkan operator Boolean untuk memperoleh hasil pencarian. Metode yang digunakan adalah tokenization dan stopping pada query sedangkanoperator Boolean yang digunaan adalah AND dan OR.

d. Result Output

Komponen results output bertanggungjawab untuk membentuk tampilan hasil pencarian. Setelah pencarian menemukan judul atau dokumen yang sesuai dengan query, halaman hasil ditampilkan dengan informasi sebagai berikut:

- 1. Jumlah judul yang ditemukan,
- 2. Judul-judul yang ditemukan
- 3. Jumlah referensi kata kunci yang ditemukan pada setiap judul
- 4. Informasi detil setiap judul



Gambar 9. Halaman hasil pencarian

Untuk mengetahui informasi detil tiap judul yang ditemukan, pengguna dapat melakukan klik pada judul yang kemudian akan muncul halaman berikut ini.

Program Studi	TI
Kategori	TA
Tahun	2007
Nama	JANU ARIYANTO WIBOWO
Nim	G.211.02.0006
Judul	Pembelajaran Multimedia pengenalan bahan - bahan kimia dalam produk sehari - hari bagi siswa SMP menggunakan BORLAND DELPHI 6 dan MAKROMEDIA FLASH Mx
Pembimbing 1	
Pembimbing 2	

Gambar 10. Halaman detil hasil pencarian

4.2.3. Pengujian dan Pengukuran

Pengujian dilakukan dengan mencari judul dengan kata kunci "sistem AND informasi" pada koleksi dokumen dengan jumlah 32 dokumen yang terdiri dari 13 dokumen judul tahun 2007 dan 19 dokumen judul tahun 2008. Hasil dari pencarian ditemukan 11 judul sebagai berikut.

Kata Kunci: sistem AND informasi

Hasil Pencarian: 11 judul ditemukan

- sistem informasi Administrasi pembayaran sekolah pada SMP MARSUDIRINI MARIA GORETTI SEMARANG
 - Kata kunci ditemukan: 2
- 2 sistem informasi CRM (Customer Relationship Management) Berbasis web pada PT. AIMS PERDANA
 - Kata kunci ditemukan: 2
- 3 sistem Informasi Penjualan baju dan accessoies secara online pada DISTRO TREESECOND SEMARANG
 - Kata kunci ditemukan: 2
- 4 Sistem Informasi Simpan Pinjam pada Koperasi JITU ARTHA PRATAMA dengan Menggunakan Pendekatan Relational Database dan Bahasa Pemrograman Visual Borland Delphi 6.0 Kata kunci ditemukan: 2
- 5 Sistem Informasi Manajemen Sumber Daya Manusia pada PT. MITRA PENGUSAHA INDONESIA Kata kunci ditemukan : 2
- 6 sistem informasi penggajian karyawan pada SMPN 1 MRANGGEN DEMAK dengan menggunakan Visual Basic 6.0
 - Kata kunci ditemukan: 2
- 7 Sistem Informasi Keuangan Sekolah di SMA MUHAMMADIYAH SALATIGA dengan VISUAL BASIC VERSI 6.
 - Kata kunci ditemukan: 2
- 8 Sistem Informasi Eksekutif Penjualan Menggunakan Visual Foxpro 9.0 (Studi Kasus PT. Elnusa Petrofin Semarang)
 - Kata kunci ditemukan: 2
- 9 Sistem Informasi rawat jalan pasien pada poliklinik umum NUGRAHA SEMARANG dengan menggunkan VISUAL FOXPRO 9.0 Kata kunci ditemukan : 2
- 10 Sistem Informasi perpustakaan pada SMA N 14 Semarang dengan menggunakan Visual Fox Pro 9.0 Kata kunci ditemukan : 2
- Sistem Informasi penjualan benang sitetik berbasis WEB pada PT. POLYSINDO EKA PERKASA KALIWUNGU

Kata kunci ditemukan: 2

Judul yang relevan sesuai keinginan pengguna adalah nomor 2, 3, 4, 5, 8, 9, 11

Pengukuran hasil pencarian

	Relevant	Not	Total
		Relevant	
Retrieved	8 (A)	3 (B)	11
Not Retrieved	0 (C)	21 (D)	21
Total	8	24	32

Keterangan:

Koleksi: A + B + C + D = 32Jml Relevan: A + C = 8Jml Tidak Relevan: B + D = 21Jml Terpilih: A + B = 11Jml tidak terpilih: C + D = 21

Maka:

Precision = A/(A+B) = 8/(8+3) = 72,73 (100%) Recall = A/(A+C) = 8/(8+0) = 100 (100%) F1 = 2rp/(r+p) = (2x72,73x100)/(72,73+100) = 14,546/172,73 = 84,21 %

V. SIMPULAN DAN SARAN

a. Kesimpulan

Berdasarkan hasil dari tahapan-tahapan dalam penelitian ini dapat diambil kesimpulan sebagai berikut:

- 1. Implementasi temu kembali informasi pada mesin pencari judul Tugas Akhir dapat memberikan hasil pencarian yang sesuai dengan kata kunci dari pengguna dengan nilai *Precision* sebesar 72,73 %, *Recall* sebesar 100% dan nilai gabungan *F1* (precision dan recall) sebesar 84,21 %.
- Model penemuan kembali informasi menggunakan Boolean retrieval (AND dan OR) dapat memberikan hasil pencarian yang lebih khusus sesuai dengan operator Boolean yang dipilih.

b. Saran

Hasil penelitian ini masih memiliki beberapa kekurangan, untuk itu peneliti memberikan saran untuk penelitian selanjutnya sebagai berikut:

- 1. Penerapan operator NOT pada model penemuan kembali informasi menggunakan Boolean retrieval.
- 2. Penggunaan model Boolean retrieval dapat dikembangkan ke penggunaan operator Boolean pada query yang bertingkat.

DAFTAR PUSTAKA

Christioko, B.V. . 2012. Implementasi Sistem Temu Kembali Informasi Studi Kasus: Dokumen Teks Berbahasa Indonesia. Transformatika,10(1), 1 – 10.

Croft, W.B., Metzler, D., & Strohman, T.. 2015. Search Engines Information Retrieval in Practice. USA: Pearson Education, Inc.

Jansen, B. J. and Rieh, S. 2010. The Seventeen Theoretical Constructs of Information Searching and Information Retrieval. Journal of the American Society for Information Sciences and Technology, 61(8), 1517-1534.

Manning, C. D., Raghavan P., & Schütze H. 2009. An Introduction to Information Retrieval. England: Cambridge University Press.

Mulyana. 2010. Pencegahan Tindak Plagiarisme Dalam Penulisan Skripsi Upaya Memperkuat Pembentukan Karakter di Dunia Akademik. Cakrawala Pendidikan, XXIX, Edisi Khusus Dies Natalis UNY.

Sommerville, Ian. 2011. Software Engineering 9ed. USA: Pearson Education, Inc.

Tala, F.Z. 2003.A Study of Stemming Effects on InformationRetrieval in Bahasa Indonesia.Tesis Master of Logic Project.Universiteit van AmsterdamThe Netherlands