

IMPLEMENTASI SISTEM TEMU KEMBALI INFORMASI PADA DOKUMEN MUTU

Abdul Rahman Kadafi

Sistem Informasi, STMIK Nusa Mandiri Jakarta
e-mail: abdurrahman.kadafi@gmail.com

ABSTRACT

In the quality management system, documentation of quality documents becomes one of the main points in the quality management requirements. Documents on the quality control system management can be stored in the form of digital documents that can be accessed by interested parties. Implementation of information retrieval in search of quality documents on the quality management system is expected to assist the user in finding relevant information in the shortest time for large amounts of data. The purpose of this paper is to compile a collection of documents as part of the information retrieval system testing tool for Indonesian text documents in the quality document and build search engines to help conduct document retrieval using the Information Retrieval Method. Implement the information retrieval system by designing a web-based search engine, which is expected to help search the documents in a collection of quality documents. Search can be labeled by providing keywords with the number of keys one or more or can also combine more than one keyword. The use of search engines can be utilized by management, quality teams, and employees to search for documents in accordance with the process to be implemented.

Keywords: *information retrieval, quality document, search documents.*

ABSTRAK

Dalam sistem manajemen mutu pendokumentasian dokumen mutu menjadi salah satu hal utama dalam persyaratan manajemen mutu. Dokumen pada manajemen sistem kendali mutu dapat disimpan dalam bentuk dokumen digital yang dapat diakses oleh pihak yang berkepentingan. Implementasi temu kembali informasi dalam pencarian dokumen mutu pada sistem manajemen mutu diharapkan dapat membantu pengguna dalam menemukan informasi yang relevan dengan kebutuhan dalam waktu singkat untuk jumlah data yang besar. Tujuan dari tulisan ini adalah untuk menyusun koleksi dokumen sebagai bagian perangkat pengujian sistem temu kembali informasi untuk dokumen teks berbahasa Indonesia dalam dokumen mutu dan membangun mesin pencari untuk membantu melakukan pencarian dokumen menggunakan metode Sistem Temu Kembali Informasi. Implementasi sistem temu kembali informasi dengan merancang mesin pencari berbasis web, yang diharapkan dapat membantu melakukan pencarian dokumen dalam koleksi dokumen mutu. Pencarian dapat dilakukan dengan memberikan kata kunci dengan jumlah kunci satu atau lebih atau dapat juga menggabungkan lebih dari satu kata kunci. Penggunaan mesin pencari dapat dimanfaatkan oleh manajemen, tim mutu, dan pegawai untuk mencari dokumen yang sesuai dengan proses yang akan dilaksanakan.

Kata Kunci: *Temu kembali informasi, dokumen mutu, pencarian dokumen.*

I. PENDAHULUAN

Penyimpanan dokumen secara digital berkembang dengan pesat seiring meningkatnya penggunaan komputer sebagai media yang membantu dalam mengoperasikan dan menyimpan data. Dalam berbagai aspek manusia hidup berhubungan dengan data, data dibuat dan dikirimkan di seluruh dunia dalam hitungan detik. Pertumbuhan data pada sistem jaringan bertambah secara besar-besaran. Kebutuhan akan pengelolaan data dan hal yang terkait dengannya semakin meningkat. Salah satu bagian yang penting dalam pengelolaan data adalah proses pencarian informasi yang diinginkan oleh pengguna atau biasa disebut dengan temu kembali informasi [1]. Banyaknya dokumen yang tersimpan dalam media penyimpanan, menimbulkan kesulitan dalam pencarian dokumen yang dikehendaki [2].

Produksi dan volume data yang besar dapat memunculkan masalah untuk mengakses informasi yang diinginkan secara akurat dan cepat. Meskipun sebagian besar dokumen digital tersimpan dalam bentuk

teks dan berbagai algoritma yang efisien untuk pencarian teks telah dikembangkan, masih terjadi kesulitan menemukan dokumen yang berhubungan dengan kata kunci tertentu dengan hasil yang tepat dan akurat. Pencarian terhadap seluruh isi dokumen yang tersimpan bukan solusi yang tepat mengingat pertumbuhan ukuran data yang tersimpan pada umumnya meningkat secara signifikan. Temu kembali informasi diharapkan dapat membantu pengguna dalam menemukan informasi yang relevan dengan kebutuhan pengguna dalam waktu singkat. Banyak teknik-teknik temu kembali informasi yang tergantung pada bahasa yang digunakan dalam dokumen. Untuk mengembangkan teknik-teknik temu kembali informasi bagi dokumen mutu, dibutuhkan perangkat pengujian untuk dokumen mutu. Salah satunya adalah suatu koleksi dokumen mutu dalam implementasi sistem kendali mutu sebagai pendekatan seragam dalam evaluasi sistem temu kembali informasi.

Pencarian yang digunakan pada dokumen mutu biasanya bersifat *word matching* berdasarkan kata kunci pada judul atau penyusun. Hal ini menyebabkan munculnya dokumen yang hanya relevan berdasarkan judulnya atau penyusunnya saja, tetapi tidak berdasarkan isi dari dokumen tersebut. Sehingga banyak dokumen relevan yang tidak terpanggil. Penerapan sistem temu kembali informasi terhadap koleksi dokumen, diharapkan mampu menjawab kebutuhan tersebut [3]. Pada dasarnya jajaran koleksi atau susunan koleksi di rak adalah untuk mempermudah menemukan kembali setiap koleksi yang ada [4].

Sistem temu kembali informasi atau *Information Retrieval Sistem* (IRS) bertujuan untuk memberitahukan keberadaan (atau ketidakberadaan) dan keterangan dokumen-dokumen yang berhubungan dengan permintaan dari pengguna, bukan memberitahukan mengenai masalah yang ditanyakan. Alasan utamanya adalah karena IRS menangani teks bahasa alami yang tidak selalu terstruktur dengan baik dan bersifat ambigu [5]. Beberapa metode sistem temu-kembali informasi antara lain metode: *Boolean*, Probabilistik, Ruang Vektor, Fuzzy, P-Norm, Jaringan Inferensi [6]. Hasil evaluasi kecepatan proses diperoleh bahwa peringkasan teks otomatis mempengaruhi tingkat kecepatan pengukuran kemiripan dokumen, untuk dokumen dengan jumlah 100 kata ke atas. Sedangkan untuk dokumen dibawah 100 kata pengaruhnya sangat kecil [7]. Perbedaan antara temu kembali informasi dan pencarian data adalah, sistem temu kembali informasi terutama berhubungan dengan pencarian informasi yang isinya tidak memiliki struktur, yang kemudian ditemukan hasil suatu informasi yang dibutuhkan atau relevan dengan kata kunci [8].

Volume informasi teks yang berkembang eksponensial menyebabkan kesulitan dalam proses temu kembali informasi utamanya pada model perolehan informasi linear berbasis *word matching* yang umumnya menjadi tidak efektif. Faktor sinonim dari kata menjadi penyebab munculnya dokumen tidak relevan dalam perolehan, sebaliknya faktor *polisemy* menyebabkan banyak dokumen yang relevan tidak terpanggil. Penerapan *clustering* dokumen dipercaya dapat meningkatkan kinerja berdasar satu *hypothesis* bahwa dokumen yang relevan terhadap suatu *query* cenderung berada dalam kluster yang sama [9]. Tulisan ini bertujuan untuk merancang implementasi sistem temu kembali informasi yang digunakan dalam koleksi dokumen mutu.

II. METODE PENELITIAN

Dalam implementasi sistem temu kembali informasi untuk koleksi dokumen mutu berbahasa Indonesia dengan menggunakan dua bagian utama yaitu *Indexing subsistem*, dan *Searching subsistem* (*matching sistem*), menggunakan beberapa tahapan sebagai berikut:

1. Kata buangan

Pada tahap kata buangan, dilakukan penentuan daftar kata-kata buangan yang tidak digunakan dalam penyaringan indeks. Pendefinisian kata buangan memilah kata sambung atau kata penghubung yang bukan sebagai kata kunci pencarian. Pembangunan *Upload Engine* digunakan untuk parsing dokumen melalui proses *upload*.

2. Koleksi Dokumen

Koleksi dokumen yang diperlukan dalam implementasi sistem temu kembali adalah dokumen yang memiliki tema yang sama. Pada tahap koleksi dokumen, dokumen yang dikumpulkan sebanyak 20 dokumen untuk jenis dokumen panduan, prosedur dan instruksi kerja dengan format .doc (Microsoft Word). File yang dikumpulkan akan diindeks berdasarkan judul dan ruang lingkup dokumen dengan cara meng-upload dokumen ke sistem *database* yang sudah disiapkan.

3. Parsing Dokumen

Proses parsing untuk dokumen dilakukan pada Judul dan ruang lingkup dokumen dengan cara membuang kata buangan yang tidak perlu dijadikan sebagai indeks. Setiap kata yang berada pada judul dan ruang lingkup dokumen dapat disimpan di dalam database menjadi indeks.

4. Data / Indeks

Daftar indeks dari setiap dokumen yang terdapat pada koleksi dokumen mutu dihasilkan dari tahap parsing dokumen. Dalam daftar indeks sudah tidak mengandung kata-kata buangan. Daftar indeks akan digunakan sebagai kata kunci pencarian informasi dengan menggunakan kata yang sesuai dengan yang diinputkan pada mesin pencari dokumen.

5. Matriks Dokumen Indeks

Matriks dokumen indeks merupakan hasil dari indeks dan dokumen dengan cara mengukur frekuensi kemunculan kata dalam indeks pada setiap dokumen.

6. Query Dokumen

Pengujian terhadap koleksi indeks yang dirancang dalam proses parsing dokumen dilakukan dengan menggunakan kata kunci pencarian. Untuk melakukan query dibangun sebuah mesin pencari yang dapat digunakan untuk memasukkan kata kunci pencarian tunggal dan ganda.

III. HASIL DAN PEMBAHASAN

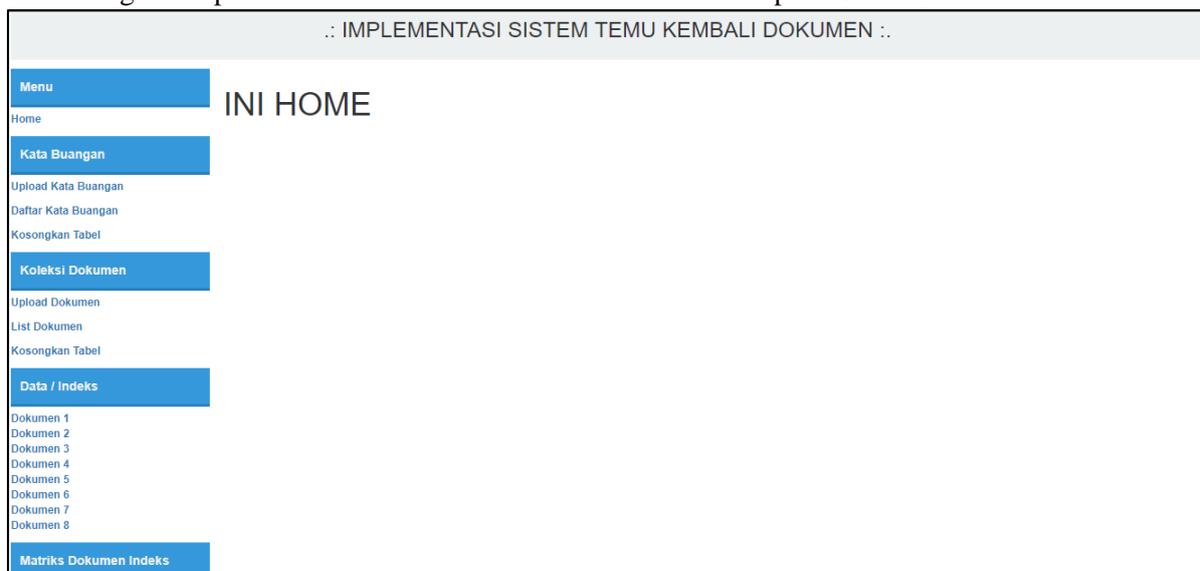
1. Kata buangan / stop list

Kata-kata buangan yang sudah dipilah diupload dan disimpan ke dalam tabel ‘Common’ pada sistem database yang sudah dibuat, menggunakan modul “upload kata buangan”. Tabel ‘common’ mempunyai spesifikasi sebagai berikut:

Field	Type
word	varchar(255)
no	int(1)

Gambar 1. Tabel common

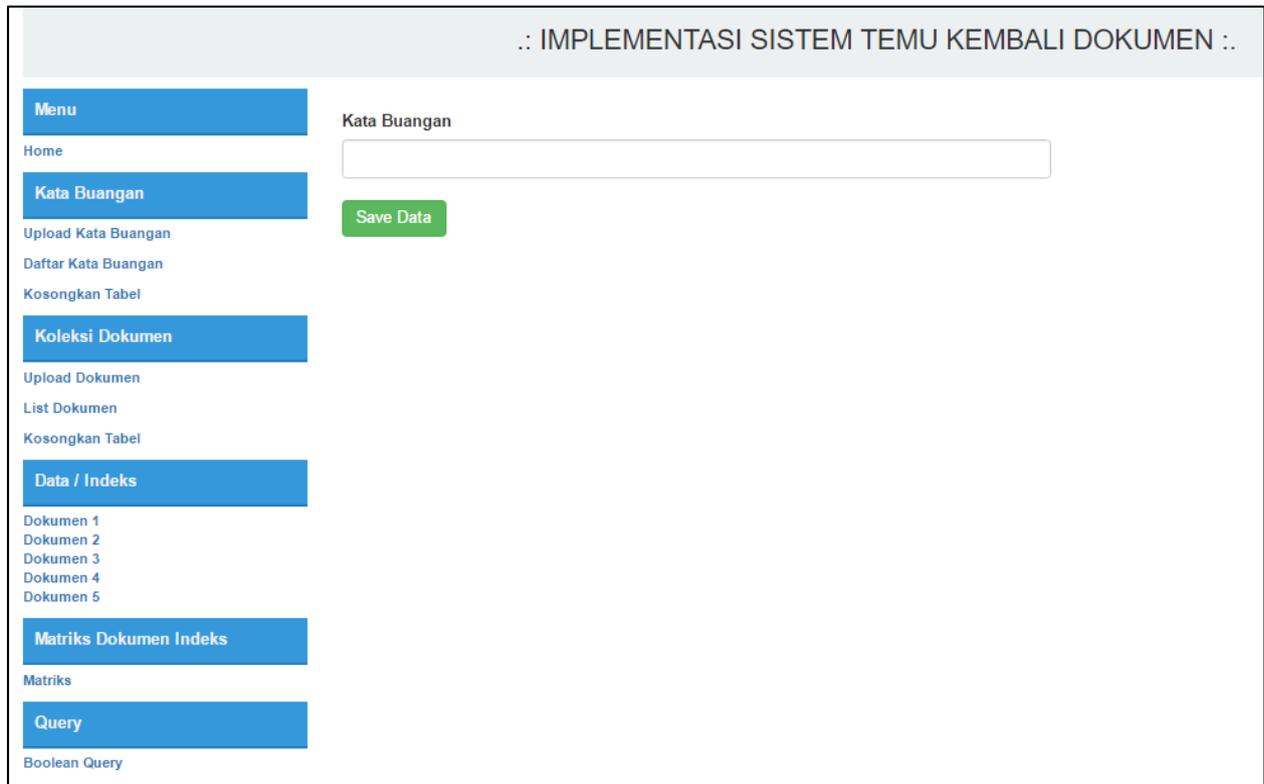
Rancangan tampilan beranda sistem temu kembali dokumen seperti Gambar 2 di bawah ini:



Gambar 2. Tampilan Beranda

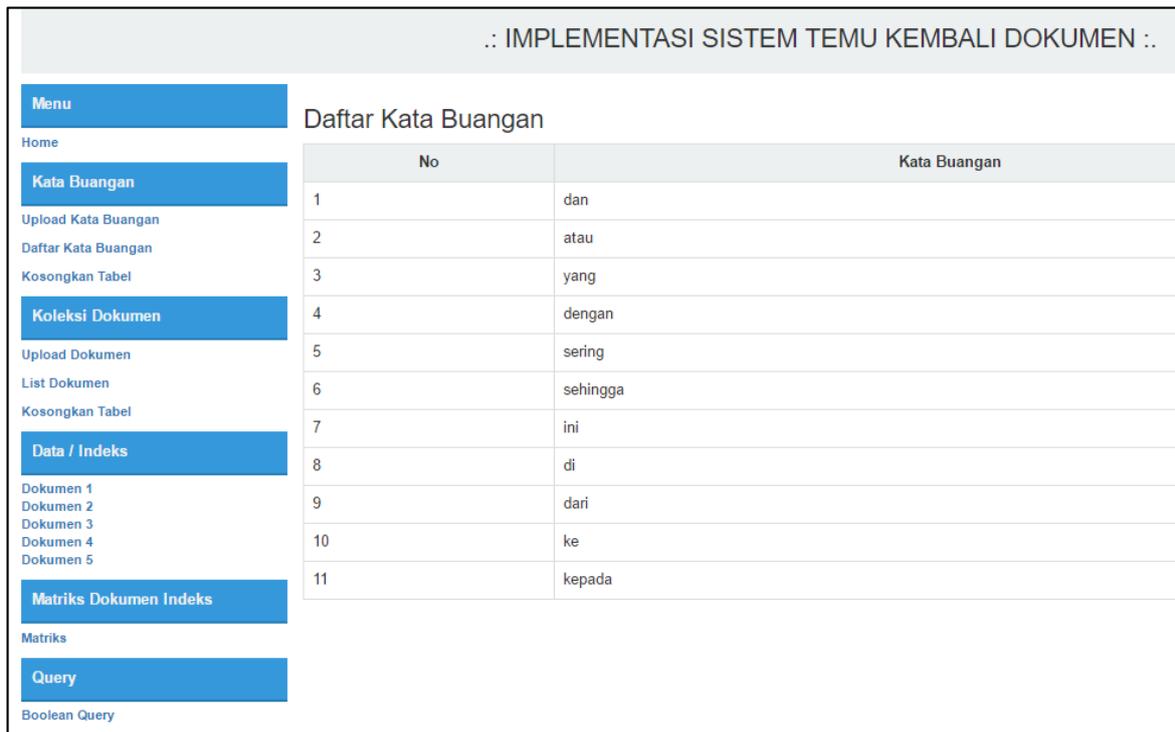
Pada tampilan beranda terdapat beberapa menu yang dapat dipilih yaitu kata buangan, koleksi dokumen, data/indeks. Matriks dokumen, dan query.

Tampilan dari modul ‘Upload Kata Buangan’ seperti gambar 3 berikut:



Gambar 3. Tampilan Upload Kata Buangan

Pada menu “upload kata buangan”, dapat digunakan untuk memasukkan kata-kata buangan yang akan dihapus dalam proses pembuatan indeks dokumen. Kata buangan yang dihapus meliputi kata penghubung, kata keterangan, kata penunjuk, dan beberapa kata lainnya yang bukan merupakan kata kunci. Setelah kata-kata buangan terupload, dapat dilihat melalui menu ‘Daftar Kata Buangan’, seperti Gambar 4 berikut:



Gambar 4. Tampilan Menu Daftar Kata Buangan

2. Koleksi Dokumen

Sepuluh dokumen mutu yang telah disiapkan sebagai sampel diupload ke dalam *database* menggunakan menu “Upload Dokumen”, berupa *form* isian judul dan ruang lingkup dari setiap dokumen yang ada. Dokumen secara otomatis akan tersaring dari kata-kata buangan yang sudah ditentukan sebelumnya dan disimpan ke dalam tabel ‘Content’, ‘key table’, dan ‘link’. Struktur dari ketiga tabel sebagai berikut:

Field	Type
<u>contid</u>	mediumint(9)
title	text
abstract	longtext

Gambar 5. Tabel Konten

Pada tabel konten terdapat tiga *field* id sebagai indeks pada tabel, title yang berisi judul dokumen dan abstrak yang akan diisikan deskripsi singkat dari dokumen yang akan disimpan.

Field	Type
<u>keyid</u>	mediumint(9)
keyword	varchar(100)

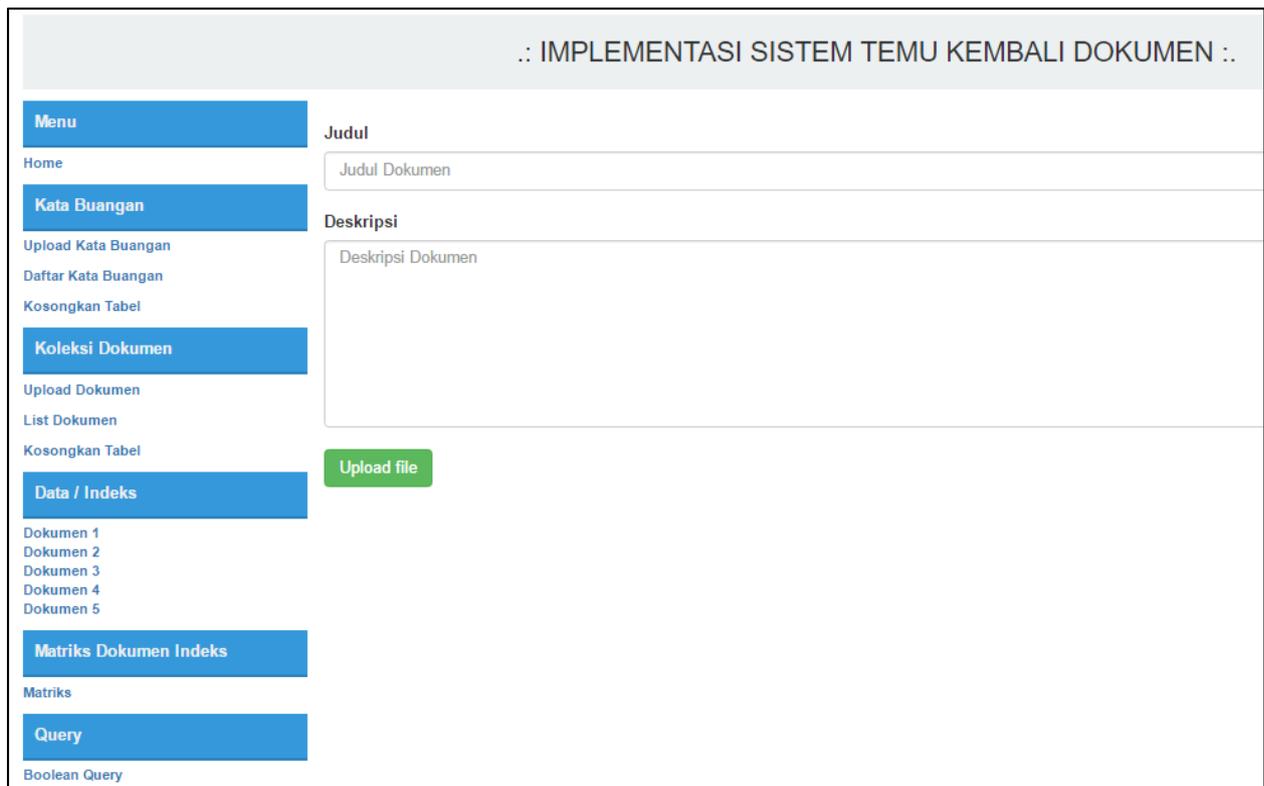
Gambar 6. Tabel *key table*

Pada Gambar 6 dapat dijelaskan bahwa *key table* sebagai tabel yang difungsikan untuk menyimpan kata kunci dalam pencarian dokumen. Pada *key table* terdiri dari 2 *field* yaitu *keyid* yang digunakan sebagai id tabel dan *keyword* yang digunakan untuk menyimpan kata yang digunakan sebagai indeks dalam pencarian dokumen. Untuk menghubungkan antara tabel deskripsi dan *key table*, digunakan tabel link, sebagaimana tampak pada Gambar 7 berikut:

Field	Type
keyid	mediumint(9)
contid	mediumint(9)

Gambar 7. Tabel link

Tampilan dari menu ‘Upload Dokumen’ seperti Gambar 8, digunakan sebagai laman untuk memasukkan judul dan deskripsi singkat (ruang lingkup) dokumen.



Gambar 8. Tampilan Menu Upload Dokumen

Setelah dokumen diupload, dapat dilihat melalui menu “List dokumen”(Gambar 9) yang berisi tentang daftar dokumen yang sudah diupload ke dalam database. Tampilan pada list dokumen adalah judul dokumen yang sudah diinputkan.

: IMPLEMENTASI SISTEM TEMU KEMBALI DOKUMEN :			
List Dokumen			
No	Judul	Deskripsi	
1	Panduan Surat Menyurat	Surat ialah alat komunikasi tertulis untuk menyampaikan informasi kepada dan dari pihak lain, baik perorangan, swasta maupun dinas yang dibuat oleh atau ditujukan kepada Yayasan atau Unit di bawahnya	
2	Pedoman pelaksanaan rapat	Rapat yang efektif adalah alat ampuh dalam suatu manajemen yang baik, terutama sebagai salah satu mekanisme kontrol dan koordinasi. Ada dua fakta yang melandasi nilai khas rapat.Pertama, rapat memungkinkan interkomunikasi di antara semua pesertanya. Tidak hanya satu arah, melainkan dua arah.Kedua, dengan adanya komunikasi antara peserta yang jumlahnya mencukupi di saat bersamaan, maka masalah keterbatasan waktu dan perolehan hasil akan dapat diatasi secara efisien	
3	Prosedur penerimaan tamu	Prosedur Penerimaan Tamu Kunjungan Belajar ini meliputi instansi/lembaga pendidikan/sekolah eksternal di luar SITNurul Fikri yang hendak berkunjung untuk keperluan silaturahmi dan kunjungan belajar.	
4	Prosedur pengendalian dokumen	Prosedur ini berlaku di lingkungan Sekolah Islam Terpadu Nurul Fikri yang meliputi kegiatan penyusunan, pengidentifikasian, penerbitan, pemeliharaan, penarikan, perubahan atau revisi dan pemusnahan dokumen.	
5	IK pembuatan draf dokumen	Instruksi Kerja ini berlaku di lingkungan Sekolah Islam Terpadu Nurul Fikri yang meliputi kegiatan membuat draf prosedur.	

Gambar 9. List Dokumen

3. Hasil Indexing

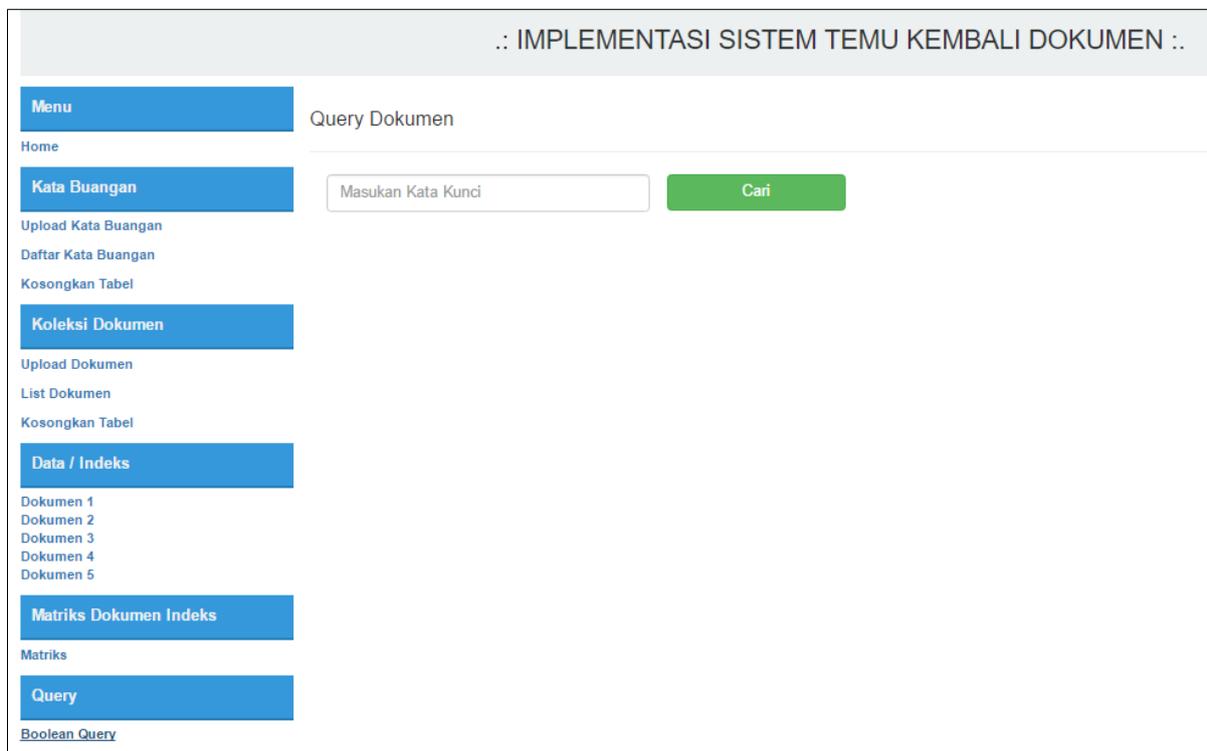
Pada dokumen yang sudah diupload akan dilakukan indeks dokumen. Indeks dokumen dengan cara memisahkan kata-kata kunci yang terdapat pada deskripsi singkat yang berisi tentang ruang lingkup dokumen. Contoh tampilan indeks dokumen sebagaimana ditampilkan pada gambar 10.



Gambar 10. Tampilan Indeks Dokumen

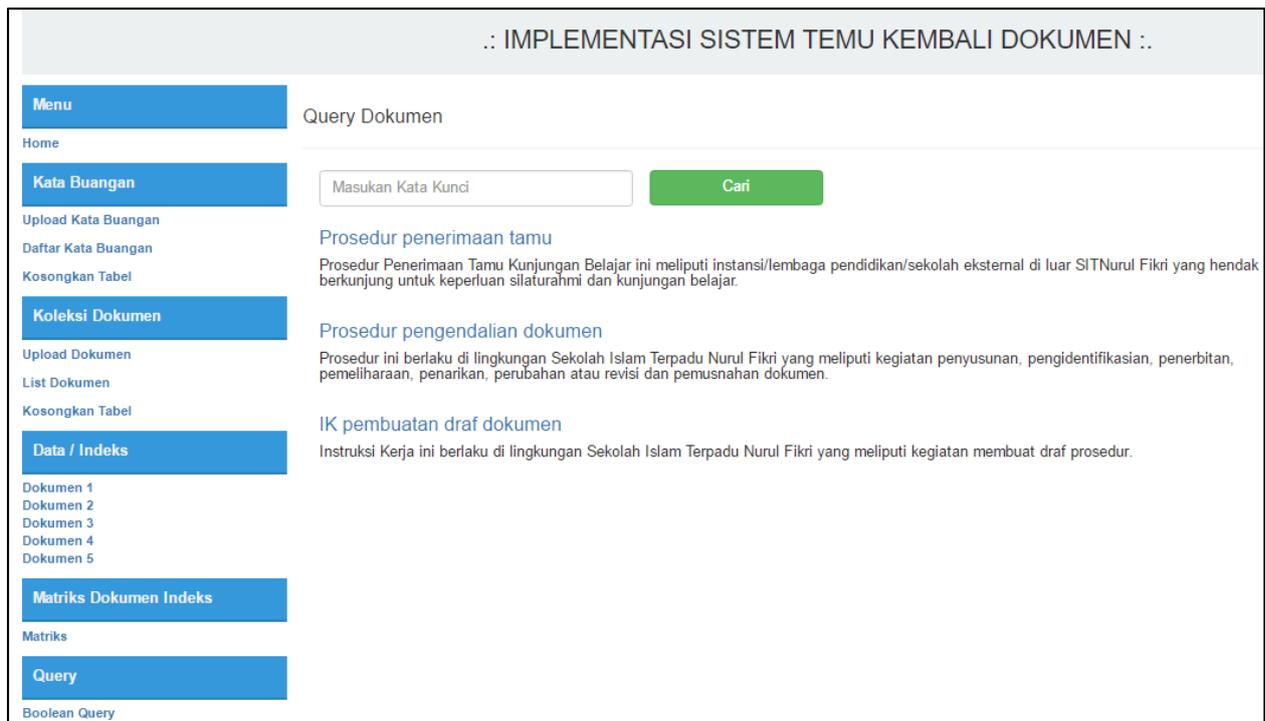
4. Hasil Query

Untuk melakukan pencarian terhadap dokumen yang sudah diinputkan ke dalam database, menggunakan menu Query, yang akan menampilkan sebagaimana Gambar 11 berikut.



Gambar 11. Input Kata Kunci

Dalam uji coba implementasi sistem temu kembali, pada tampilan menu query dapat diinputkan kata tertentu sebagai pencarian dokumen yang akan dicari, dan menghasilkan tampilan sebagaimana pada Gambar 12 berikut.



Gambar 12. Tampilan hasil pencarian dokumen

Pada Gambar 12 merupakan contoh hasil pencarian menggunakan kata “prosedur”. Setelah dienter atau klik tombol cari, maka akan tampil judul dan deskripsi singkat dari dokumen yang mengandung kata prosedur, baik pada nama dokumen maupun pada deskripsi dokumen.

IV. KESIMPULAN

Dari pengujian yang dilakukan, dapat disimpulkan bahwa menerapkan konsep-konsep dalam sistem temu kembali informasi dan membangun suatu mesin pencari dapat membantu dalam melakukan pencarian dokumen dalam suatu koleksi dokumen. Pencarian dapat dilakukan dengan memberikan kata kunci dengan jumlah kunci satu atau lebih atau dapat juga menggabungkan lebih dari satu kata kunci.

DAFTAR PUSTAKA

- [1] Marzuki, Imam, “Temu Kembali Informasi Big Data menggunakan K-Menas dan Clusterring”, Smatika Jurnal, Vol 5. No. 2. ISSN. 2087-0256. STMIK Indonesia Malang, 2015.
- [2] Mumahhad, dkk, “Sistem Temu Kembali Informasi dalam Dokumen Menggunakan Metode Latent Semantic Indexing”, Jurnal Masyarakat Informatika, Volume 3, Nomor 5, ISSN 2083 – 4930, 2012
- [3] Rachman, Hastarita Fika, “Sistem Temu Kembali Informasi Menggunakan Model Ruang Vektor dan Inverted Index”, Madura, Jurnal SimanteC, Vol 2, No 3 Juni 2012, ISSN 2088-2130, 2012.
- [4] Magdalena, 2015, “Pemanfaatan Sarana Temu Kembali Informasi Khususnya Katalog UPT Perpustakaan UNIMA Oleh Mahasiswa”, Journal “Acta Diurna” Volume IV. No.1. Tahun 2015.
- [5] Very Christioko, Bernadus, “Implementasi Sistem Temu Kembali Informasi Studi Kasus: Dokumen Teks Berbahasa Indonesia (Implementation of Information Retrieval System Case Study: Text Document in Indonesian Language)”, Jurnal Transformatika, Vol. 10, No. 1, 2012.
- [6] Yahma Wisyani, “Query Ganda Pada Sistem Temu Kembali Informasi Berbasis Jaringan Inferensi”, MAKARA, SAINS, VOL. 8, NO. 2, AGUSTUS 2004: 76-84, 2004.
- [7] Putung, Carter dkk, “Penerapan Sistem Temu Kembali Informasi Pada Dokumen Skripsi”, E-journal Teknik Informatika, Volume 8, No 1 (2016), ISSN : 2301-8364, 2016.
- [8] Iyan Mulyana, Andi Chairunnas, Aries Maesya, “Identifikasi Plagiasi Karya Ilmiah berbasis Temu Kembali Informasi Menggunakan Algoritam Edit Distance Melalui Peringkasan Teks Otomatis”, Seminar Nasional Teknologi, Bogor, 2015.
- [9] Hamzah, Amir, “Temu kembali Informasi Berbasis Kluster Untuk Sistem Temu Kembali Informasi Teks Bahasa Indonesia”, Jurnal Teknologi, Vol 2. No. 1, Akprind, Yogyakarta, 2009.