

PROSIDING SEMMAU 2017

SEMINAR NASIONAL & KONFERENSI SISTEM INFORMASI, INFORMATIKA & KOMUNIKASI

TEMA :

E - BUSINESS SEBAGAI DAYA DUKUNG INDUSTRI KREATIF

Kupang, 25 November 2017

BUKU 3
ISBN: 978-602-73628-0-2



STIKOM UYELINDO KUPANG

PROSIDING SEMMAU 2017

Penulis,
Pemakalah SEMMAU 2017

Penerbit,
STIKOM UYELINDO KUPANG

KOMITE

Penulis:

Pemakalah Seminar Nasional & Konferensi Sistem Informasi, Informatika & Komunikasi (SEMMAU 2017)

ISBN : 978-602-73628-0-4

Komite Program:

Prof. Daniel Herman Fredy Manongga, M.Sc., Ph.D. (UKSW)
Prof. Dr. Ir. Eko Sediyono, M. Kom (UKSW)
Prof. Mustafid (UNDIP)
Prof. Dr.Ir. Kuswara Setiawan, M.T. (UPH)
Dr. Ir. Rila Mandala, M.Eng. (ITB)
Dr. Achmad Nizar, S. Kom., M.Kom. (UI)
Ir. Dana Indra Sensuse, M.Lis., Ph.D. (UI)

Penyunting:

Max ABR. Soleman Lenggu. S. Kom., M.T.
Skolastika Siba Igon, S. Kom., M.T
Reza Hardi Nugroho
Henry Max Matchless Ratmo
Fransiskus Xaverius Pey Tae
Eko D. Rihibiha
Yohana Stefania Pipa Wea
Imanuel M. Laka
Ritwan Banu
Maissy P. Babar
Frialdhy S. Ketty
Maestryn A. Taeko
Muhammad Bdariyadi
Barnabas Sarbunan

Desain Sampul:

Max Lenggu

Redaksi :

Dapur Semmau
Lembaga Penelitian, Publikasi dan Pengembangan pada Masyarakat
Jl. Perintis Kemerdekaan 1, Kayu Putih, Kupang, NTT, Indonesia.
Telp.(0380)8554501, Fax (0380) 8554501
Email : semmau@uyelindo.ac.id
http://www.semmau.uyelindo.ac.id.

Penerbit :

Sekolah Tinggi Manajemen Informatika & Komputer (STIKOM) Uyelindo Kupang.
Jl. Perintis Kemerdekaan 1, Kayu Putih, Kupang, NTT, Indonesia.
Telp.(0380)8554501, Fax (0380) 8554501
Email : stikom@uyelindo.ac.id
http://www.uyelindo.ac.id.

Cetakan ketiga November 2017
Hak Cipta di Lindungi Undang-undang
Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara apapun tanpa ijin tertulis dari penerbit.

KATA PENGANTAR

Segala puji dan syukur selayaknya tercurah kehadirat Allah Yang Maha Kuasa yang tanpa henti mengucurkan rahmat dan karuniaNya, baik kurunia sehat, rejeki, kecerdasan, kemauan, dan bahkan juga karunia dalam bentuk kesadaran dan kemampuan bersyukur kepadaNya, dan dengan ijinnya Prosiding Seminar Nasional danKonferensi Sistem Informasi, Teknik Informatika, dan Komunikasi (SEMMAU) tahun 2017 dengan Tema "*E-BUSSINES* SEBAGAI DAYA DUKUNG INDUSTRI KREATIF" dapat kami terbitkan.

Buku Prosiding ini berisi sekumpulan *Paper* dari hasil penelitian ilmiah yang telah diseleksi, untuk dipresentasikan dalam kegiatan Seminar Nasional dan Konferensi Sistem Informasi, Teknik Informatika, dan Komunikasi (SEMMAU) tahun 2017 dan bertempat di *Ballroom* Swiss Belinn Kristal Hotel Kupang Nusa Tenggara Timur pada tanggal 25 November 2017, kegiatan ini diikuti oleh peserta pemakalah yang berasal dari berbagai perguruan tinggi yang tersebar di kawasan Nusa Tenggara Timur (NTT), maupun di luar NTT, yang terdiri dari 26 makalah dari para peserta pemakalah.

Seminar Nasional yang bertemakan "*E-BUSSINES* SEBAGAI DAYA DUKUNG INDUSTRI KREATIF" ini menghadirkan pembicara utama berkelas nasional yakni Prof. Dr. Ir. Eko Sediyono, M. Kom.

Ucapan terima kasih kami sampaikan kepada Reviewer Paper dan pihak-pihak yang telah membantu penyelenggaraan Seminar Nasional dan Konferensi Sistem Informasi, Teknik Informatika, dan Komunikasi (SEMMAU) tahun 2017 ini. Semoga prosiding ini dapat bermanfaat dan dapat digunakan dengan sebaik-baiknya.

Akhir kata, jika ada yang kurang berkenan selama penyelenggaraan kegiatan seminar maupun dalam penerbitan buku prosiding ini mohon dimaafkan. Semoga apa yang telah kita lakukan ini bermanfaat bagi kemajuan bangsa dan negara dimasa depan. Amin.

Kupang, November 2017

Panitia,

Sumarlin

DAFTAR ISI

	Halaman
OPTIMALISASI PEMAHAMAN MATERI TEKNOLOGI INFORMASI DAN KOMUNIKASI DENGAN KONTEN BERBASIS MULTIMEDIA (STUDI KASUS PADA SMP SATU ATAP MORO – ENDE). *Agustinus Lambertus Suban.*	**394 - 399**
APLIKASI PENILAIAN DOSEN DAN STAF PADA STIKOM ARTHA BUANA KUPANG BERBASIS ANDROID. *Ahmad Haidaroh, Fajar Riski Maulidan.*	**400 - 408**
PETA RUTE ANGKUTAN UMUM KOTA KUPANG BERBASIS *MOBILE GEOGRAPHIC INFORMATION SYSTEM* *Benyamin Jago Belalawe, Benediktus Yoseph Bhae, Petrus Katemba.*	**409 - 413**
ANALISIS POPULARITAS WEBSITE PEMERINTAH KOTA DI SUMATERA SELATAN. *Dien Novita, Lisa Amelia Fransen*	**414 -419**
E-COMMERCE SEBAGAI UPAYA PENGEMBANGAN USAHA KECIL DAN MENENGAH DALAM PEMASARAN KERIPIK HASIL INOVASI DI KOTA BOYOLALI. *Donna Setiawati.*	**420 - 427**
SISTEM INFORMASI PUBLIK RUMAH TANGGA MISKIN PADA KABUPATEN SUMBA TIMUR. *Edwin Ariesto Umbu Malahina, Emanuel Safirman Bata*	**428 - 435**
CARA PEMBERANTASAN HAMA TANAMAN KAKAO DENGAN METODE *MULTIMEDIA DEVELOP LIFE CYCLE* GUNA MENINGKATKAN PRODUKSI TANAMAN KAKAO BERBASIS *ANDROID.* *Febriyanti Alwisye Wara, Imelda Dua Reja.*	**436 - 439**
PERANCANGAN SISTEM INFORMASI PELINTAS BATAS WILAYAH NEGARA RI DAN RDTL. *Fransiskus M.H. Tjiptabudi, Skolastika S. Igon.*	**440 - 446**
SISTEM PENDUKUNG KEPUTUSAN UNTUK PENENTUAN MINAT DENGAN METODE *ANALYTICAL HIERARCHY PROCESS.* *Gregorius William Meno, Dony M. Sihotang, Tiwuk Widiastuti.*	**447 - 452**
APLIKASI ANDROID SEBAGAI MEDIA PEMBELAJARAN KARDIOVASKULER PADA PEREDARAN DARAH MANUSIA. *Imelda Dua Reja, Febriyanti Alwisye Wara, Bastian Texaniwin Nakoda.*	**453 - 458**
PEMANFAATAN APLIKASI PEMBELAJARAN BUDAYA TIMOR TENGAH SELATAN SEBAGAI UPAYA MENINGKATKAN RASA CINTA TANAH AIR. *Mardhalia Saitakela, Skolastika S. Igon.*	**459 – 462**

PENGARUH PENERAPAN SISTEM INFORMASI AKUNTANSI TERHADAP KINERJA INDIVIDU PEGAWAI DI KOPERASI PEGAWAI NEGERI SIPIL BHAKTI HUSADA. **463 – 467**
Maria Florentina Rumba, Yosafat Pati Koten

SISTEM TEMU BALIK INFORMASI DOKUMEN TEKS MENGGUNAKAN VECTOR SPACE MODEL. **468 – 473**
Mariam Fatima Somu, Paulina Aliandu, Paskalis Andi Nani.

PENENTUAN LARVA TIRAM MUTIARA TERBAIK SEBAGAI PENGHASIL MUTIARA UNGGUL DENGAN LOGIKA *FUZZY.* **474 – 481**
Marselina Dorce Tlaan, Marinus I.J. Lamabelawa.

PEMETAAN TEMPAT WISATA KABUPATEN ROTE NDAO BERBASIS WEB GIS. **482 – 486**
Menhya Snae, Max ABR S Lenggu, Benediktus Y. Bhae.

DETEKSI CALON KREDITUR MOTOR DENGAN NAÏVE BAYES CLASSIFIER (STUDI KASUS: PT. FIF CABANG KUPANG). **487 – 494**
Miransyah Koroh, Marlinda Vasty Overbeek.

PENERAPAN GOOGLE MAPS DALAM MENENTUKAN LOKASI-LOKASI WIFI CORNER DI KOTA KUPANG BERBASIS ANDROID. **495 – 498**
Muhammad Harits Ardiyanto, Yohanes Suban Belutowe.

SISTEM TEMU BALIK DOKUMEN TEKS MENGGUNAKAN METODE *BOOLEAN* DAN TERM WEIGHT TF.IDF. **499 – 504**
Ni Putu Anggi Yuliani, Paulina Aliandu, Paskalis Andi Nani.

PENERAPAN METODE *SMART (SIMPLE MULTI ATTRIBUTE RATING TECHNIQUE)* DAN ALGORITMA *K-NN (K-NEAREST NEIGHBOR)* DALAM PENENTUAN STATUS KESEHATAN BAYI BARU LAHIR DI RUMAH SAKIT BHAYANGKARA KUPANG **505 – 511**
Omar Bilham Tamonob, Kornelis Letelay, Sebastianus Mola.

SISTEM PENDUKUNG KEPUTUSAN PENENTUAN CALON PENERIMA BEASISWA PADA PROGRAM STUDI TEKNIK INFORMATIKA UNIVERSITAS NUSA NIPA MENGGUNAKAN METODE AHP. **512 – 515**
Petrus Wolo, Stefania Memen Tupen, Yosep P. Minggo.

FUTURE COMPUTING PROFESSION. **516 – 522**
Raul Bernardino, Hasibun Asikin

INFORMASI LOKASI DAERAH PENGHASIL KOMODITI UNGGULAN DI KABUPATEN KUPANG BERBASIS ANDROID. **523 – 527**
Remerta Noni Naatonis, Skolastika S. Igon.

PENERAPAN CUSTOMER RELATIONSHIP MANAGEMENT PADA APLIKASI PEMASARAN KAIN TENUN IKAT KHAS FLORES BERBASIS ANDROID. **528 – 534**
Sumarlin, Dewi Anggraini.

DESAIN MODEL INFORMASI DANA DESA BERBASIS WEB (STUDI KASUS DESA TANINI KECAMATAN TAKARI KABUPATEN KUPANG). **535 – 543**
Yohanes Payong, Antonio Suares, Venansius A. K. Ga'a

ANALISA CITRA UNTUK KLASIFIKASI KAYU BANGUNAN. **544 – 547**
Yohanes Suban Belutowe

IMPLEMENTASI METODE *SIMPLE ADDITIVE WEIGHTING* UNTUK SELEKSI PENERIMAAN BEASISWA BIDIK MISI DI POLITEKNIK NEGERI SAMARINDA. **548 – 554**
Yusni Nyura, Damar Nurcahyono.

SISTEM TEMU BALIK DOKUMEN TEKS MENGGUNAKAN METODE *BOOLEAN* DAN TERM WIGHT TF.IDF

Ni Putu Anggi Yuliani[1], Paulina Aliandu[2], Paskalis Andi Nani[3]

[123] Jurusan Teknik Informatika, Unika Widya Mandira Kupang
[1] putuanggi28@gmail.com, [2] paulinaaliandu@gmail.com, [3] paskalisnani@gmail.com

Abstrak

Kepolisian Sektor (Polsek) Satarmese merupakan markas kepolisian yang bertugas untuk melindungi dan mengayomi masyarakat di Kecamatan Satarmese, Kabupaten Manggarai, Nusa Tenggara Timur. Setiap bulannya ada 3 hingga 5 kasus pelanggaran hukum yang harus ditangani Polsek Satarmese. Setiap kasus memiliki dokumen yang harus diolah dan disimpan oleh Polsek Satarmese. Dokumen yang tersimpan sejak tahun 1997 hingga tahun 2016 berjumlah lebih dari 500 dokumen. Dokumen-dokumen tersebut hanya disimpan menggunakan arsip yang belum terkomputerisasi sehingga menyulitkan petugas kepolisian di Polsek Satarmese jika ingin mencari dokumen yang diperlukan untuk penyelidikan atau keperluan lainnya. Untuk mempercepat proses pencarian maka diperlukan sebuah sistem yang dapat mengidentifikasi dan memanggil suatu dokumen dalam suatu penyimpanan seperti sistem temu balik (*information retrieval*). Ada beberapa metode dalam sistem temu balik informasi, salah satunya adalah Metode Boolean. Kelebihan metode Boolean dari metode lainnya adalah kecepatannya dalam proses pencocokan. Metode Boolean merupakan suatu cara mengekspresikan keinginan pemakai ke sebuah *query* dengan memakai operator-operator Boolean yaitu: "*and*","*or*", dan "*not*". Pada Metode Boolean, *query* diproses sesuai dengan operator yang digunakan dan dokumen diperingkat berdasarkan bobot dari dokumen menggunakan metode pembobotan tf-idf. Masing-masing term (kata) diberi bobot baik terhadap dokumen tempat *term* tersebut berada maupun terhadap dokumen-dokumen lain yang ada. Sistem temu balik dokumen pelanggaran hukum ini berbasis web sehingga dapat diakses dimana saja dan dapat membantu petugas kepolisian mencari informasi dari dokumen yang disimpan secara cepat, serta dokumen dari hasil pencarian dapat diurutkan berdasarkan bobot informasinya. Performansi menunjukan *exact match* dalam temu balik dimana *precission* dan *recall* bernilai 1.

***Kata kunci**: sistem temu kembali, boolean, tf-idf, hukum, polsek*

1. PENDAHULUAN

Sebagai makhluk sosial manusia selalu hidup berdampingan dan berinteraksi satu sama lain. Untuk menjaga kedamaian dalam proses interaksi tersebut maka diperlukan aturan untuk mengatur individu atau kelompok yang disebut hukum. Namun seiring berjalannya waktu tentu timbullah permasalahan antar berbagai pihak dan munculah pelanggaran atas beberapa aturan-aturan yang telah ditentukan, disinilah aparat penegak hukum seperti polisi diperlukan. Pihak kepolisian kemudian akan menyelesaikan permasalahan tersebut dan juga mengurus dokumen terkait pelanggaran hukum yang telah dilakukan.

Kepolisian Sektor (Polsek) Satarmese merupakan markas kepolisian yang bertugas untuk melindungi dan mengayomi masyarakat. Polsek Satarmese adalah bagian dari Kepolisian Resort (Polres) Manggarai yang memiliki kekuasaan wilayah hukum di Kecamatan Satarmese berpusat di Desa Iteng. Wilayah hukum yang harus dilindungi oleh Polsek Satarmese terdiri dari 23 desa yaitu Desa Lolang, Jaong, Gololambo, Ngkaer, Pongkor, Papang, Ulu Belang, Umung, Ponggeok, Wewo, Lungar, Mocok, Wae Ajang, Koak, Tado, Golo Muntas, Langgo, Satar Loung, Legu, Iteng, Paka, Gara dan Tal. Begitu banyak desa yang harus diayomi Polsek Satarmese berbanding lurus dengan masalah pelanggaran hukum yang harus diurus. Setiap waktu dokumen pelanggaran hukum terus meningkat. Dalam sebulan jumlah tindak pidana yang mencapai 3 sampai 5 kasus.

Dokumen-dokumen pelanggaran hukum dari setiap periode harus disimpan untuk dijadikan arsip. Adapun sistem yang berjalan dalam pendataan dan penyimpanan koleksi dokumen pelanggaran hukum tersebut, hanya menggunakan arsip yang belum terkomputerisasi dan disimpan pada lemari/rak buku dan tidak disusun berdasarkan tahun ataupun kriteria pelanggaran, sehingga apabila petugas yang ingin mencari dokumen tersebut untuk keperluan seperti penyelidikan ulang, pembuatan laporan tahunan dan

pengecekan pengaduan dari pihak tertentu maka harus membaca dan mencari satu persatu di rak penyimpanan. Arsip yang tersimpan sejak tahun 1997 hingga tahun 2016 berjumlah lebih dari 500 dokumen. Hal ini menyebabkan petugas mengalami kesulitan dalam mencari secara manual atau memeriksa satu persatu dokumen pelanggaran hukum yang diperlukan. Selain itu, pencarian seperti ini tentu akan sangat membosankan dan menyita banyak waktu mengingat banyaknya arsip yang tersedia.

Permasalahan temu balik informasi dari kumpulan dokumen melahirkan suatu ilmu yang diberi nama temu kembali informasi (*information retrieval*). Temu kembali informasi berkaitan dengan representasi, penyimpanan, struktur dan akses dari dokumen-dokumen yang bertujuan untuk memudahkan pencarian sebuah informasi. Representasi dari dokumen itu nantinya harus mudah diakses oleh pengguna untuk mendapatkan informasi. Pengguna *information retrieval system* sangat bervariasi sesuai dengan kebutuhan yang berbeda-beda.

Beberapa penelitian yang berkaitan dengan sistem temu balik informasi pernah dilakukan [1] yang menghasilkan sebuah sistem yang dapat mencari informasi dari isi *file* dokumen yang disimpan di dalam sistem. Penelitian ini dikembangkan dengan metode *waterfall*. Metode pembobotan yang digunakan dalam penelitian ini adalah tf-idf. Penelitian [2] dengan judul "Perancangan *Information Retrieval System* Dengan Metode *Extended Boolean* Dan *Savoy*" yang menghasilkan sebuah sistem yang dapat melakukan pencarian dokumen yang relavan dengan kata kunci yang dimasukkan oleh pengguna dan juga dokumen yang ditampilkan telah terurut berdasarkan perhitungan bobot. Sistem ini dapat mencari *file* dengan ekstensi .txt dan .doc. Selain itu penelitian [3] menghasikan sistem temu balik informasi pencarian skripsi menggunakan pembobotan tf.idf dan *vector space model*. Sedangkan penelitian [4] menggunakan pembobotan tf.idf namun untuk *similarity* menggunakan metode *jaccard similarity.* Sistem ini diimplementasikan dalam pencocokan dokumen dengan tafsir Al-Quran yang berada dalam *database* dan mencari padanan yang sesuai dengan dokumen yang diinputkan tersebut.

Penelitian yang dilakukan dengan judul "Sistem Temu Kembali Informasi Dokumen Pelanggaran Hukum Dengan Metode *Boolean*". Penelitian ini mengacu pada penelitian [3] dengan judul "*Information Retrieval System* Dokumen Teks Abstrak Skripsi Dengan Metode Pembobotan Tf-Idf Dan Pengukuran *Similarity* Menggunakan *Vector Space Model*" namun dengan kasus dan metode yang berbeda. Lake menggunakan metode *Vector Space Model* sedangkan metode yang digunakan pada penelitian ini adalah Metode *Boolean*. Studi kasus yang digunakan [3] adalah dokumen teks abstrak di Perpustakaan Program Studi Teknik Informatika Universitas Widya Mandira Kupang; sedangkan penelitian ini mengambil studi kasus dokumen pelanggaran hukum di Polsek Satarmese. Namun kedua penelitian sama-sama menggunakan pembobotan tf-idf. Penelitian lain yang menjadi acuan penelitian ini adalah penelitian dari [2] dengan judul "Perancangan *Information Retrieval System* Dengan Metode *Extended Boolean* Dan *Savoy*". Penelitian [2] menggunakan Metode *Extended Boolean* dan *Savoy* yang berbeda dengan Metode *Boolean*.

Metode-metode temu kembali informasi sudah dikenal sejak lama, salah satunya adalah Metode *Boolean*. Metode *Boolean* merupakan suatu cara mengekspresikan keinginan pemakai ke sebuah *query* dengan memakai operator-operator Boolean yaitu: "*and*","*or*", dan "*not*". Pada Metode *Boolean*, *query* diproses sesuai dengan operator yang digunakan dan dokumen diperingkat berdasarkan bobot dari dokumen. Penggunaan Metode Boolean dalam studi kasus dokumen hukum dianggap sesuai. Salah satu algoritma pembobotannya adalah algoritma tf-idf (*term frequency- inverse document frequency*) yang dipengaruhi oleh frekuensi kemunculan kata pada sebuah dokumen dan frekuensi dari dokumen yang memiliki kata tersebut.

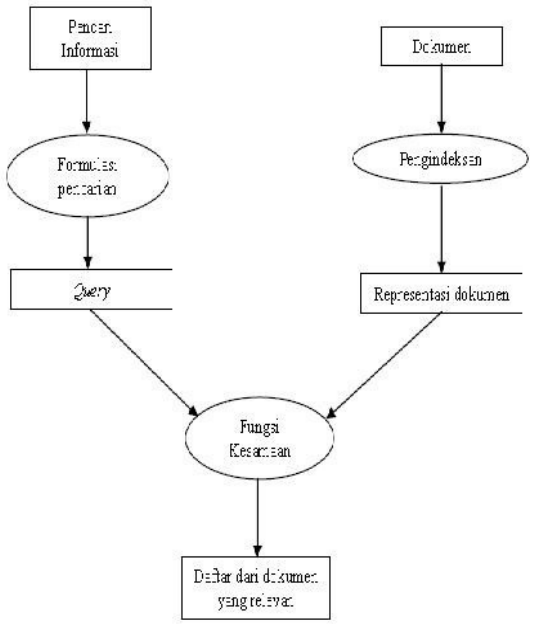
2. KAJIAN LITERATUR

2.1 Sistem Temu Kembali (*Information Retrieval Systems*)

Menurut Lancaster [5]: "sebuah sistem temu kembali informasi tidak memberitahu (yakni tidak mengubah pengetahuan) pengguna mengenai masalah yang ditanyakannya, sistem tersebut hanya memberitahukan keberadaan (atau ketidakberadaan) dan keterangan dokumen-dokumen yang berhubungan dengan permintaannya". Sedangkan Greengrass [6], menyatakan bahwa *Information Retrieval* adalah sebuah disiplin ilmu yang berhubungan dengan pencarian data tidak terstruktur, khususnya dokumen-dokumen tekstual, dalam respon terhadap sebuah *query* atau pernyataan topik.

Baeza-Yates dan Riberto Neto [7] menyatakan bahwa *Information Retrieval* (IR) merupakan sistem untuk merepresentasikan, menyimpan, mengorganisasikan, dan memperoleh informasi. *Information retrieval* adalah bagaimana menemukan suatu dokumen dari dokumen-dokumen tidak terstruktur yang memberikan informasi yang dibutuhkan dari koleksi dokumen yang sangat besar yang tersimpan dalam komputer [8]. Pada dasarnya sistem temu kembali informasi merupakan sistem yang berfungsi untuk menemukan informasi yang relevan dengan kebutuhan pemakai.Untuk lebih

jelasnya mengenai kerangka sistem temu kembali informasi dapat dilihat pada Gambar 1.



Gambar 1 Kerangka Sistem Temu Kembali Informasi Sederhana [9]

2.2 Bagian-Bagian Sistem Temu Kembali

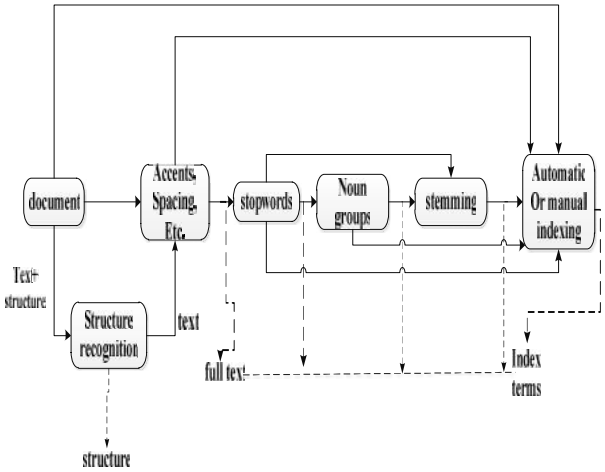
Baeza-Yates [7] membagi IR ke dalam lima bagian, yaitu *text operations, indexing, query operations, searching* dan *ranking*.

2.2.1 Text Operations/ Preprocessing

Text operations (pengoperasian teks) adalah proses transformasi dokumen dan *query* menjadi kata-kata indeks. Dalam suatu dokumen terdapat beberapa kata yang memiliki makna lebih penting dibandingkan kata-kata lainnya. Karena itulah *preprocessing* terhadap teks dalam suatu koleksi dokumen dianggap perlu dalam menentukan kata yang akan digunakan sebagai *index terms.* Tahap *preprocessing* mencakup pengoperasian teks seperti penghapusan *markup*, penghapusan *stopwords*, dan *stemming* (pembentukan kata dasar).

Penghapusan *markup* adalah sebuah proses mengubah sekelompok karakter (teks dari dokumen) menjadi sekelompok kata yang memiliki kemungkinan menjadi kata-kata indeks [7]. Kata-kata yang sering muncul dalam dokumen pada suatu koleksi *(corpus)* yang tidak mewakili indeks untuk melakukan pencarian disebut dengan *stopwords*. Kata-kata yang termasuk dalam kelompok *stopwords* ini biasanya akan disaring dari kelompok kata yang akan dijadikan sebagai indeks. Dengan membuang stopwords, ukuran struktur *indexing* dapat dikurangi dan dapat mengurangi *recall* dokumen yang tidak sesuai.

Stem adalah bagian kata yang tinggal setelah menghilangkan imbuhannya (awalan dan akhiran) [7]. Menurut Manning [8], tujuan dari proses *stemming* adalah untuk mengurangi bentuk inflektif dan bentuk-bentuk turunan yang berhubungan dari sebuah kata terhadap bentuk dasar yang umum. Tahapan *preprocessing* dapat dilihat pada gambar 2.



Gambar 2 Tahap *Preprocessing* Pada IR [7]

2.2.2 Indexing

Indeks adalah bahasa yang digunakan di dalam sebuah buku konvensional untuk mencari informasi berdasarkan kata atau istilah yang mengacu kedalam suatu halaman. Pada sistem temu kembali informasi, indeks ini nantinya yang digunakan untuk merepresentasikan informasi di dalam sebuah dokumen [1].

2.3 Metode Boolean

Model *Boolean* pada sistem temu kembali ini berdasarkan teori himpunan dan aljabar *Boolean*. Model *Boolean* menyediakan sebuah *framework* yang mudah dimengerti oleh *user* umum dari sistem IR. Selain itu, *query* dimasukkan sebagai ekspresi *Boolean* memiliki semantik yang tepat. Dokumen adalah himpunan dari istilah (*term*). Istilah (*term*) dalam sebuah *query* dihubungkan dengan menggunakan operator AND, OR atau NOT. Metode ini merupakan metode yang paling sering digunakan pada mesin penelusur (*search engine*) karena kecepatannya.

2.4 Pembobotan Tf.Idf

Metode *term frequency – inverse document frequency atau yang lebih dikenal dengan tf-idf* merupakan metode untuk menghitung bobot setiap kata yang paling umum digunakan pada *information retrieval*. Metode ini juga terkenal efisien, mudah dan memiliki hasil yang akurat [10]. Pada model pengindeksan yang berdasarkan pada frekuensi istilah dapat diperkirakan bahwa istilah-istilah indeks terbaik adalah istilah-istilah yang sering muncul dalam dokumen individual tetapi jarang muncul dalam keseluruhan koleksi. Sebuah penanda

kepentingan jenis ini yang umum adalah perkalian (*tf·idf*) dengan bobot *wij* sebuah istilah *Ti* dalam dokumen *Di*, didefinisikan sebagai frekuensi istilah dikalikan dengan fungsi *inverse document frequency*.

Term frequency (tf) merupakan frekuensi kemunculan *term* (t) pada dokumen (d). *Document frequency* suatu *term* adalah banyaknya dokumen dalam koleksi, yang mengandung kemunculan *term* tersebut. *Inverse document frequency* (idf) suatu term adalah invers dari df yaitu 1/df. Nilai idf suatu *term* yang tinggi menunjukkan bahawa term tersebut penting karena tidak banyak dipakai di dokumen-dokumen dalam koleksi. Sebaliknya nilai idf suatu *term* yang rendah menunjukkan bahwa *term* tersebut tidak penting karena banyak dipakai di dokumen – dokumen dalam koleksi.

Metode tf-idf akan menghitung nilai *term frequency* (tf) dan *inverse document frequency* (idf) dengan formula yang biasa digunakan [11] adalah:

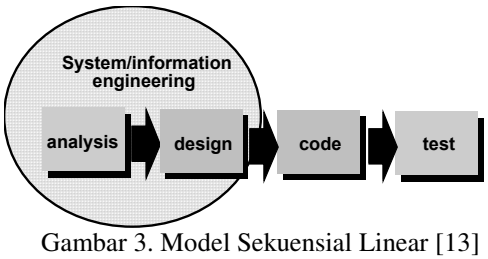
$$W_{ij} = tf_{ij} * \log(N/df_j) \quad (1)$$

Dimana:

N = total jumlah dokumen dalam koleksi
tf_{ij} = frekuensi *term* j dalam dokumen i
W_{ij} = bobot *term* j dalam dokumen i
df_j = jumlah dokumen yang berisi *term* j

3. METODE PENELITIAN

Pada penelitian ini, digunakan teknik pengembangan sistem model sekuensial linear yang mengusulkan sebuah pendekatan pengembangan perangkat lunak yang sistematik dan sekuensial. Metode *System Development Life Cycle* (SDLC) atau daur hidup pengembangan sistem merupakan metodologi klasik yang digunakan untuk mengembangkan, memelihara, dan menggunakan sistem [12].



Gambar 3. Model Sekuensial Linear [13]

Metodologi ini mencakup sejumlah fase atau tahapan yang akan dilakukan dalam penelitian ini yakni:

a. Perencanaan

Tahapan ini merupakan kegiatan awal untuk membuat sistem yang akan dibangun dan proses pengumpulan data. Kegiatan tersebut meliputi: (i) Observasi yaitu dengan melakukan *survey* langsung terhadap permasalahan yang diambil di Kepolisian Sektor (Polsek) Satarmese; (ii)Wawancara yaitu teknik pengumpulan data dengan mengadakan tanya jawab/wawancara langsung dengan petugas di Polsek Satarmese; (iii) Studi kepustakaan, yaitu dengan menambah literatur-literatur dari buku, jurnal dan media internet yang berkaitan dengan *information retrieval, preprocessing* dokumen teks, Metode *Boolean,* perancangan *database*, serta panduan cara membuat aplikasi.

b. Analisis

Pada tahapan ini dilakukan analisa terhadap hal-hal atau kebutuhan yang diperlukan untuk pembuatan sistem temu kembali dokumen hukum dengan Metode *Boolean* yang meliputi:

b.1 Analisis Kebutuhan Sistem

Analisis kebutuhan sistem dilakukan untuk mengetahui fasilitas yang harus disediakan atau dimiliki oleh sistem agar dapat melayani kebutuhan pengguna sistem. Fungsi utama dari sistem ini adalah untuk membantu pengguna dalam mencari dokumen dari koleksi dokumen pelanggaran hukum yang sesuai dengan *query* yang diinginkan.

b.2 Analisis Peran Sistem

Sistem yang dibangun mempunyai peranan sebagai berikut:

1. Sistem yang dibuat dapat menginput, menyimpan, melihat, mengubah dan menghapus teks yang dalam hal ini adalah dokumen-dokumen pelanggaran hukum di Polsek Satarmese.
2. Sistem ini juga melakukan pencocokan antara himpunan dokumen yang tersimpan dengan *query* yang dimasukan pengguna serta operator *Boolean* yang dipilih ("AND" dan "OR").
3. Sistem dapat memberikan *output* berupa hasil pencarian dokumen pelanggaran hukum yang telah teurut sesuai tingkat kemiripan dengan *query*

b.3 Analisis Peran Pengguna

Sistem ini memiliki dua pengguna yaitu administrator dan *user* petugas Polsek Satarmese. Administrator menginput, melihat, menambah, menghapus dokumen pelanggaran hukum, *stopwords*, simbol dan data *user*. *User* yang menggunakan sistem ini untuk melakukan pencarian (*search*) dokumen pelanggaran hukum berdasarkan *query* yang dimasukkan.

c. Perancangan Sistem

Terdapat 3 (tiga) komponen perangkat yang dibutuhkan untuk membantu kinerja sistem agar tujuan dari sistem tersebut dapat tercapai. Perangkat tersebut adalah perangkat keras (*hardware*) berupa komputer, perangkat lunak (*software*) berupa

program dan perangkat manusia (*brainware*). Hasil dari perancangan sistem adalah sebuah sistem yang didalamnya terdapat informasi yang berguna.

Sistem akan dibangun pada sistem operasi *Windows* dengan bahasa pemrograman PHP (*Hypertext Preprocessor*). Model aliran data akan digambarkan dengan *Data Flow Diagram* (DFD). *Flowchart* akan digunakan untuk memperlihatkan urutan dan hubungan antar proses beserta instansinya. Perancangan basis data akan menggunakan *Entity Relationship Diagram* (ERD).

d. Pengkodean

Setelah membuat perancangan sistem temu kembali dokumen hukum dengan metode *Boolean*, maka tahap selanjutnya yakni mengimplementasikan hasil dari perancangan tersebut ke dalam PHP sebagai bahasa pemrogramannya serta pemilihan *platform* sistem operasi yang digunakan yakni sistem operasi *Windows*.

e. Pengujian dan analisis

Dalam penelitian ini proses uji coba dilakukan dengan metode pengujian *black box.* Pengujian *black box* hanya mengamati hasil eksekusi dan memeriksa fungsional dari perangkat lunak. Tujuan metode pengujian ini adalah mencari kesalahan pada fungsi yang salah atau hilang sehingga menemukan cacat yang mungkin terjadi pada saat pengkodean.

4. HASIL DAN PEMBAHASAN

Implementasi dan pengujian pada sistem temu balik yang dibahas meliputi uji performansi (*Precision* dan *Recall*), perbandingan waktu pengindeksan, dan waktu pencarian.

4.1 Uji Performansi

Pengujian performansi akan dilakukan dengan metode *recall* dan *precision*. Pengujian dilakukan dengan memasukan kata kunci sesuai dengan deskripsi. Sampel pengujian diambil sebanyak 4 buah menggunakan logika "*and"* dan "*or"*. sesuai tertera pada Tabel 1 dan jumlah dokumen tersimpan adalah 10.

Tabel 1. Tabel Deskripsi Dokumen

Query	Dok. Relevan	Hasil Pencarian		
		Relevan	Tak Relevan	Tot.
Tersangka AND Fulbertus (KK01)	1	1	0	1
Pencurian OR Penganiayaan (KK02)	6	6	0	6
Tersangka OR Pasal (KK03)	10	10	0	10
Pengrusakan OR Pembunuhan (KK04)	2	2	0	2

Adapun performansi sistem temu balik dapat dilihat hasil pengujiannya pada Tabel 2. Nilai tertinggi dari *recall* dan *precision* adalah 1 dan nilai terendahnyan adalah 0. Dari pengujian diatas dapat dilihat bahwa sistem dapat menemukan semua dokumen yang relevan sehingga mendapatkan nilai *recall* dan *precision* yang tinggi, hal ini disebabkan karena Metode *Boolean* hanya mampu menemukan dokumen yang relevan dan menolak dokumen yang tidak relevan. Tidak ada perhitungan kepada dokumen yang mirip.

Tabel 2. Performansi Sistem Temu Kembali

Pengujian	*Recall*	*Precision*
KK01	1	1
KK02	1	1
KK03	1	1
KK04	1	1

4.2 Perbandingan Waktu Pengindeksan

Proses pengindeksan dilakukan saat proses penyimpanan dokumen. Pengujian pada proses pengindeksan dilakukan pada 3 laptop. Dokumen yang disimpan berekstensi .docx. Tabel 3 adalah tabel evaluasi proses pengindeksan.

Tabel 3. Perbandingan Waktu Pengindeksan

Nama dok.	File size	Waktu Pengindeksan (milidetik)			Rerata / dok
		K1	K2	K3	
Fulbertus	36 KB	231	969	681	627
Idris	61 KB	923	4.240	2.160	2.441
Alexander Wanggu	21 KB	993	4.160	2.680	2.611
Heli	30 KB	1.520	6.730	5.090	4.446,67
BAB IV	2.106 KB	594	2.230	1.540	1.454,67
Rata-Rata		852,2	3.665,8	2.430,2	2.316,068

4.3 Perbandingan Waktu Pencarian

Proses ini melakukan pengujian pencarian dengan memasukkan beberapa *query* dan dicatat waktu pencariannya, serta dilakukan juga perbandingan perhitungan bobot manual dan bobot pada sistem.

Tabel 4. Perbandingan Waktu Pencarian

Query	Dokumen ditemukan dari n tersimpan	Waktu Pencarian (ms)		
		K1	K2	K3
Tersangka AND Fulbertus	1 dari 5 tersimpan	24	93	56
Tersangka AND Fulbertus	1 dari 10 tersimpan	28	95	59
Pencurian OR Penganiayaan	10 dari 20 tersimpan	227	845	634
Iping AND Pencurian	1 dari 25 tersimpan	45	167	112
Tersangka OR Pasal	30 dari 30 tersimpan	259	1038	984
Rata-Rata Waktu Pencarian		116,6	447,6	369

Pencarian dilakukan pada 3 komputer dengan *query* dan dokumen yang sama. Hasil perbandingan waktu pencarian dapat dilihat pada Tabel 4.

5. KESIMPULAN

Kesimpulan yang diperoleh berdasarkan hasil dan analisis sebagai berikut:

a. Sistem temu kembali yang dibuat mampu membantu petugas kepolisian di Polsek Satarmese melakukan pencarian dokumen pelanggaran hukum yang diperlukan dengan cepat dan tepat dengan rata-rata waktu pencarian adalah 311,08 milidetik.
b. Kinerja sistem sangat dipengaruhi oleh spesifikasi komputer yang digunakan.
c. Kecepatan pengindeksan dan pencarian sebuah dokumen teks tergantung dari jumlah dokumen yang disimpan, ukuran *file* yang disimpan, kompleksitas isi dokumen dan spesifikasi komputer yang digunakan.
d. Secara kinerja, sistem temu kembali yang dikembangkan memiliki nilai *recall* dan *precision* yang tinggi yaitu 1.

REFERENSI

[1] Nadirman, F., 2006. *Sistem Temu Kembali Informasi Dengan Metode Vector Space Model Pada Pencarian File Dokumen Berbasis Teks,* Tugas Akhir, Unpublished, Universitas Gajah Mada Yogyakarta.

[2] Luis, D., 2015. *Perancangan Information Retrieval System Dengan Metode Extended Boolean Dan Savoy*, Tugas Akhir, Unpublished, Sekolah Tinggi Manajemen Informatika Dan Komputer Medan.

[3] Lake, L.K., 2015, *Information Retrieval System Dokumen Teks Abstrak Skripsi Dengan Metode Pembobotan Tf-Idf Dan Pengukuran Similarity Menggunakan Vector Space Model*, Tugas Akhir, Unpublished, Universitas Katolik Widya Mandira Kupang.

[4] Purwandari, S.S.S., 2012, *Rancang Bangun Search Engine Tafsir Ayat- Ayat Al- Quran yang Sesuai dengan Dokumen Teks Berbahasa Indonesia Menggunakan Metode Jaccard Similarity,* Tugas Akhir, Unpublished, Universitas Islam Negeri Malang.

[5] Rijsbergen, C.J., 1979, *Information Retrieval, Second Edition*, Butterworths, London

[6] Greengrass, E., 2000, *Information Retrieval: A Survey*, Online, URL = https://www.csee.umbc.edu/csee/research/cadip/readings/IR.report.120600.book.pdf diakses pada tanggal 22 September 2017

[7] Baeza-Yates, R., Riberio-Neto, B., 1999, *Modern Information Retrieval,* ACM. Press, New York: ANSI.

[8] Manning, C.D., Raghavan, P., dan Schutze, H., 2009, *Introduction to Information Retrieval,* Cambridge University Press.

[9] Ingwersen, P., 1992, *Information Retrieval Interaction*, London, Taylor Graham Publishing, URL= http://www.db.dk/pi/iri diakses pada tanggal 9 April 2017.

[10] Robertson, S.E., 2005, *Understanding Inverse Document Frequency: On theoretical arguments for IDF*, Journal of Documentation, Vol. 60, pp. 502–520.

[11] Lu, G., 1999, *Multimedia Database Management Systems,* Artech House Inc. Canton Street.

[12] Kadir, A., 2014, *Pengenalan Sistem Informasi Edisi Revisi*, Andi Offset, Yogyakarta.

[13] Pressman, R.S., 2001, *Software Engineering : A Practitioner's Approach 5th Edition,* McGraw Hill, New York

STIKOM UYELINDO KUPANG

PROGRAM STUDI :
SISTEM INFORMASI (S1) TERAKREDITASI
TEKNIK INFORMATIKA (S1) TERAKREDITASI
TEKNIK INFORMATIKA (D3) TERAKREDITASI

Jalan Perintis Kemerdekaan I -Kayu Putih Kupang-NTT
Telp; 0380-8554500, 85554499, Fax.0380-8554502
Website: http://www.uyelindo.ac.id
Website: http://www.semmau.uyelindo.ac.id
Email: stikom@uyelindo.ac.id, semmau@uyelindo.ac.id



ISBN
978-602-73628-0-2