



EDITORIAL BOARD

Alamat Redaksi

Gedung A, Fakultas Ilmu Sosial dan Ilmu Politik

Universitas Airlangga

Jl. Dharmawangsa Dalam Selatan No. 4-6

Surabaya, 60286

Telp. (031) 503 4015

Fax. (031) 501 2442

E-mail: palimpsest@unair.ac.id

Website: <http://palimpsest.fisip.unair.ac.id>



Table of Contents

No	Title	Page
1	ANALISIS KEMAMPUAN SEARCH ENGINE GOOGLE, YAHOO DAN ALTAVISTA	1 - 19



Vol. 5 - No. 1 / 2014-08

TOC : 1, and page : 1 - 19

ANALISIS KEMAMPUAN SEARCH ENGINE GOOGLE, YAHOO DAN ALTAVISTA

ANALISIS KEMAMPUAN SEARCH ENGINE GOOGLE, YAHOO DAN ALTAVISTA

Author :

Yunus Abdul Halim | yunus.halim@unair.ac.id

Fakultas Ilmu Sosial dan Ilmu Politik

Abstract

Information retrieval system is a measure of the success of a case of information search engines like Google, Yahoo and AltaVista. Through information retrieval system that produced a measurable level of the search engines. Measure of ability can be seen from the value of recall, precision and NIAP (Non-Interpolated Average Precision). The third measure can show the effectiveness and capabilities of search engines. Comparison test conducted by the ability of search engines using 30 subject libraries as a keyword by taking a sample of the first page of results returned by search engines. All data that appear at first analyzed by using the formula recall, precision and NIAP. Experiment for google is give results (recall: 1, precision: 0.66, NIAP: 0.56), yahoo (recall: 1, precision: 0.61, NIAP: 0.48) and AltaVista (recall: 1, precision: 12:40, NIAP: 0:31) . These results indicate that the ability to recall all the same search engines while precision capabilities and NIAP, google search engine better when compared with a yahoo and altavista. Based on the results of these experiment can be concluded that information retrieval system owned by the search engines Google, Yahoo and AltaVista have characteristics of each. Google is more likely to yield better precision when compared to Yahoo and AltaVista. While Yahoo and AltaVista still require further investigation.

Keyword : Information, retrieval, system, recall, precision, Non-Interpolated, Average, Precision, ,

Daftar Pustaka :

1. **Chowdhury, G.G, (2004)**. Introduction to Modern Information Retrieval. London : Library Association
2. **Dwipendi, Nripendra, (2013)**. Statistical Analysis of Search Engines. London : International Journal of Computer Theory and Engineering

Copy alamat URL di bawah ini untuk download fullpaper :

journal.unair.ac.id/filerPDF/palimd50b6d7c5cfull.pdf

ANALISIS KEMAMPUAN SEARCH ENGINE GOOGLE, YAHOO DAN ALTAVISTA

Yunus Abdul Halim¹

Abstract

Information retrieval system is a measure of the success of a case of information search engines like Google, Yahoo and AltaVista. Through information retrieval system that produced a measurable level of the search engines. Measure of ability can be seen from the value of recall, precision and NIAP (Non-Interpolated Average Precision). The third measure can show the effectiveness and capabilities of search engines.

Comparison test conducted by the ability of search engines using 30 subject libraries as a keyword by taking a sample of the first page of results returned by search engines. All data that appear at first analyzed by using the formula recall, precision and NIAP.

Experiment for google is give results (recall: 1, precision: 0.66, NIAP: 0.56), yahoo (recall: 1, precision: 0.61, NIAP: 0.48) and AltaVista (recall: 1, precision: 12:40, NIAP: 0:31). These results indicate that the ability to recall all the same search engines while precision capabilities and NIAP, google search engine better when compared with a yahoo and altavista.

Based on the results of these experiment can be concluded that information retrieval system owned by the search engines Google, Yahoo and AltaVista have characteristics of each. Google is more likely to yield better precision when compared to Yahoo and AltaVista. While Yahoo and AltaVista still require further investigation.

Keywords : *Information retrieval system, recall, precision, Non-Interpolated Average Precision.*

Latar Belakang

Sudah menjadi fakta umum bahwa mencari informasi di internet bukanlah sesuatu yang baru lagi. Setiap orang dapat mengakses informasi yang dibutuhkannya lewat internet. Dalam pencarian informasi di internet, pengguna internet kerap kali menggunakan *search engine* yang dinilai sebagai cara mudah yaitu cukup dengan menuliskan kata kunci di kotak *search*. Informasi dari internet tersebut bisa berbentuk ilmiah atau non ilmiah, bisa berasal dari blog atau dari mana pun yang terjaring sesuai dengan kata kunci. Dari sekumpulan informasi yang ada di internet, pengguna internet bisa saja mendapatkan informasi yang diinginkannya, namun tak jarang juga malah makin tersesat di belantara informasi. Untuk mengantisipasi berbagai masalah yang muncul dalam penelusuran informasi di internet, salah satu solusi yaitu, pengguna internet perlu memperhatikan *search engine* mana yang cocok digunakan. Misalnya seorang pengguna internet sedang mencari literatur ilmiah untuk penelitian yang sedang dikerjakannya, dalam hal ini, tentu akan lebih efektif jika pengguna internet tersebut

¹ Korespondensi: Yunus Abdul Halim. Departemen Informasi dan Perpustakaan, FISIP, Universitas Airlangga. Jl. Airlangga 4-6 Surabaya, 60286, Indonesia. Telp. (031) 5011744. E-Mail: yunus.halim@unair.ac.id

memilih *search engine* yang dapat menghubungkannya dengan database yang merupakan literatur atau informasi ilmiah pula.

Kebutuhan akan informasi adalah pengakuan tentang adanya ketidakpastian dalam diri seseorang yang mendorong seseorang untuk mencari informasi. Jika seseorang sudah yakin bahwa sesuatu informasi benar-benar diinginkan, maka keinginan informasi akan berubah menjadi permintaan informasi (*information demands*), (Krikelas, 1983). Dari kebutuhan informasi yang besar ini, maka muncullah berbagai perilaku dari para pemakai teknologi dalam pencarian informasi seperti *search engine google* melalui internet. Permasalahan yang timbul adalah bahwa perkembangan internet yang sangat cepat telah mengakibatkan terjadinya banjir data dan informasi (*information overloaded*) sehingga sangat memudahkan pengguna (*user*) dalam mencari data dan informasi secara tepat sesuai yang diinginkan. Tapi bisa dirasakan juga bagaimana jika pencarian informasi yang sesuai kebutuhan pengguna tanpa ada media *search engine*. Kesulitan mencari informasi akan sama seperti ketika pengguna mencari informasi melalui buku. Pencarian yang menghabiskan waktu, mengeluarkan banyak biaya dan terkadang hasil yang di dapat kurang sesuai. Pencarian jadi kurang mudah dan cepat jika tidak ada *search engine*, harus mengetahui situs-situs atau web dari sebuah penelitian dan itu tidak mudah jika harus di telusuri secara manual, itu sebabnya *search engine* di internet berperan penting dan sangat membantu pengguna dalam mencari informasi. Tetapi dengan adanya internet para pengguna tidak perlu bersusah-susah bila ingin mengetahui segala sesuatu tentang perkembangan informasi baik informasi dalam negeri dan luar negeri seperti jurnal atau artikel.

Kemungkinan besar internet bisa memberikan kenyamanan pada penggunaannya, meliputi kecepatan akses, kemutakhiran dan keberagaman informasi yang ada (Sisson dan Pontau dalam 1-Saleh, 2004). Sebagai akibat dari perkembangan teknologi, internet sebagai pencari informasi mengakibatkan mulailah bermunculan perusahaan-perusahaan semacam *Altavista.com*, *Excite.com*, *Yahoo.com*, *AskJeeves.com*, dan lain sebagainya berlomba-lomba untuk membuat mesin pencari (*search engine*) yang terbaik. Secara prinsip, tujuan dari sebuah program *searching engine* adalah menemukan dokumen atau arsip elektronik di internet yang sesuai dengan kebutuhan atau permintaan pengguna dalam waktu yang sesingkat-singkatnya. Kedua hal inilah, yaitu kualitas hasil temuan dan waktu pencarian, yang kemudian menjadi pengukur baik tidaknya kinerja sebuah *searching engine*. Mesin pencari (*search engine*) yang dapat membantu di dalam penelusuran suatu artikel, file, ataupun data base.

Kemampuan semua *search engine* seperti google, yahoo dan altavista bisa diukur dari sistem temu kembali informasi yang dihasilkan. Untuk mengukur efektifitas sebuah sistem temu kembali informasi dapat dilakukan dengan perhitungan terhadap nilai *recall* (perolehan) dan nilai *precision* (ketepatan). Semakin tinggi nilai *precision* (ketepatan) suatu penelusuran maka semakin efektif penelusuran yang dilakukan. Suatu sistem temu kembali informasi dinyatakan efektif apabila hasil penelusuran mampu menunjukkan ketepatan (*precision*) yang tinggi sekalipun perolehannya rendah, (Rowley dalam Hasugian, 2003: 5). Suatu dokumen dalam penelusuran dinyatakan relevan, jika kata kunci (*query*) yang digunakan dalam penelusuran terdapat di bagian *abstrac* isi dari suatu dokumen, atau suatu dokumen yang di temukan berhubungan dengan *field* judul, *field* abstrak, *field* pengarang. Dokumen juga dinyatakan relevan jika isi dokumen secara keseluruhan sesuai dengan kata kunci (*query*). Melalui sistem temu kembali informasi yang dimiliki oleh setiap *search engine*, maka peneliti tertarik untuk

menganalisis kemampuan search engine dari yahoo, google dan altavista berdasarkan efektifitas yang dihasilkan dari sistem temu kembali informasinya.

Permasalahan

Saat ini banyak search engine yang digunakan ketika melakukan aktifitas *browsing di internet*. Beberapa search engine yang terkenal dan biasa dipakai antara lain google, yahoo dan altavista. Ketiga *search engine* tersebut memiliki karakteristik dan saling bersaing untuk menarik minat pengguna internet, agar digunakan ketika melakukan pencarian informasi di internet. Secara prinsip, semua *search engine* pasti memiliki sistem temu kembali informasi yang ditanam pada program aplikasinya. Kemampuan sistem temu kembali informasi tersebut sangat menentukan kemampuan telusur yang dimiliki oleh setiap *search engine*. Melalui parameter yang dimiliki oleh sistem temu kembali informasi, peneliti tertarik untuk menganalisis lebih lanjut kemampuan ketiga *search engine* tersebut.

Batasan Masalah

Penelitian yang berjudul analisis kemampuan *search engine* google, yahoo dan altavista menggunakan batasan masalah sebagai berikut:

- a. Uji perbandingan dilakukan untuk tiga *search engine* yaitu google, yahoo dan alta vista.
- b. Ukuran yang digunakan yaitu uji sistem temu kembali informasi
- c. Uji dilakukan dengan menggunakan 30 subject perpustakaan (*libraries*).
- d. Koneksi yang digunakan untuk pengujian menggunakan jalur internet yang sama.

Tujuan

Penelitian analisis kemampuan *search engine* pada google, yahoo dan altavista ini mempunyai tujuan :

1. Mengkaji lebih lanjut *search engine* berdasarkan sistem temu kembali informasi yang digunakan oleh masing masing *search engine*.
2. Membandingkan kemampuan *search engine* melalui uji sistem temu kembali informasi yang dihasilkan berdasarkan *recall*, *precision* dan NIAP.

Metode Penelitian

Penelitian analisa kemampuan *search engine* google, yahoo dan altavista menggunakan pendekatan kuantitatif, melalui uji internet di laboratorium.

- a. Uji coba dilakukan dengan menggunakan 30 subjek *libraries* (Tabel 1).

Tabel 1. Daftar 30 subjek *libraries*

No	Subject	Keyword
1	Digital Libraries	Digital Libraries pdf
2	Classification	Classification pdf
3	Library Automation	Library Automation pdf
4	Integrated Library Systems	Integrated Library Systems pdf
5	Catalogs	Catalogs pdf
6	Library Circulation and Loans	Library Circulation and Loans pdf
7	Collection Management	Collection Management pdf
8	Collection Development	Collection Development pdf
9	Library Information Network	Library Information Network pdf
10	Library Decoration	Library Decoration pdf
11	Duplicate Book	Duplicate Book pdf
12	Bibliographical Libraries	Bibliographical Libraries pdf
13	Library Resources	Library Resources pdf
14	Library Administration	Library Administration pdf
15	Library Science	Library Science pdf
16	Public Libraries	Public Libraries pdf
17	Proprietary Libraries	Proprietary Libraries pdf
18	Private Libraries	Private Libraries pdf
19	Rental Libraries	Rental Libraries pdf
20	Research Libraries	Research Libraries pdf
21	School Libraries	School Libraries pdf
22	Small Libraries	Small Libraries pdf
23	Acquisitions Libraries	Acquisitions Libraries pdf
24	Book list	Book list pdf
25	Library Circulation	Library Circulation pdf
26	Data Processing	Data Processing pdf
27	Inventories	Inventories pdf
28	Library site	Library site pdf
29	Merger Libraries	Merger Libraries pdf
30	Shelving	Shelving pdf

b. Hasil Uji coba akan dianalisa dengan menggunakan pendekatan sistem temu kembali informasi yang dihasilkan melalui analisa *recall*, *precision* dan NIAP yang didapat. Skala *precision* yang digunakan yaitu 0 s/d 1, dengan ketentuan seperti pada tabel 2.

Tabel 2. Range skala precision

Bobot	Keterangan
0	Title tidak Sesuai, isi tidak sesuai
0.5	Title Sesuai, isi kurang sesuai Title tidak Sesuai, isi kurang sesuai
1	Title Sesuai, isi sesuai Title tidak sesuai, isi sesuai

Search Engine

Menurut Akhmad Sudrajat (2005) Search engine adalah sebuah program yang dapat diakses melalui internet yang fungsinya adalah membantu pengguna komputer mencari berbagai hal yang ingin diketahuinya. Di internet, terdapat ratusan (bahkan ribuan) *search engine* yang dapat diakses secara cuma-cuma. Karena fungsinya mirip pintu gerbang bagi para pengguna sebelum memasuki situs atau /website /tertentu, banyak orang yang menyebut *search engine* sebagai portal (bandingkan dengan fungsi portal sesungguhnya yang kerap kita temui di pintu masuk ke sebuah kompleks perumahan). Dari begitu banyak search engine yang ada, yang paling populer diantaranya adalah *Google*, *Yahoo*, dan *Altavista*.

Di beberapa kasus kita juga akan mempergunakan mesin pencari lain, terutama saat membahas fasilitas-fasilitas pencarian yang tidak tersedia di *Google*. Perlu dicatat bahwa sebagian besar contoh-contoh di dalam fasilitas Google Buku menggunakan bahasa Indonesia demi mempermudah pemahaman dari para pengguna yang baru belajar menggunakan *search engine*. Dalam kenyataan sehari-hari, penulis menyarankan agar pengguna internet menggunakan pula bahasa Inggris sebagai kata kunci karena mayoritas situs-situs di internet menggunakan bahasa internasional tersebut. Disamping itu, database yang berbasis bahasa Indonesia masih terhitung sedikit sehingga tidak jarang pencari (surfer) hanya akan mendapatkan sedikit alternatif hasil pencarian atau bahkan tidak memperolehnya sama sekali. Hal itu sangat berbeda dengan pencarian menggunakan bahasa Inggris yang jarang tidak membuahkan hasil. Meskipun demikian, penulis yakin bahwa seiring dengan berjalannya waktu, akan semakin banyak situs-situs berbasis bahasa Indonesia di internet, sehingga mereka yang membutuhkan informasi dalam bahasa Indonesia akan terbantu.

Google

Sistem pencarian *google* dilakukan oleh *googlebots* (robot google) yg namanya *spider*, robot ini menghubungi seluruh web server didunia untuk mendapatkan document. Pencarian ini tentu saja tidak membabi buta mencari tetapi juga atas ijin dari web server atau webmaster. Suatu web baru, akan lebih cepat dicari oleh robot *google* jika memberitahukan alamat web yang telah dibuat ke pihak *Google* (dengan cara mengirimkan daftar alamat atau file sitemap). Dari hasil pencarian, *Google* memberikan nomer dari setiap dokumen yang diketemukannya. Langkah berikutnya (setelah mencari halaman web) adalah membuat peng-*index*-an. Peng-*index*-an dilakukan dengan cara melakukan pencarian/scan kata dari setiap dokumen/halaman web, dari pencarian ini akan dibuat sebuah daftar kata yang berisi penjelasan bahwa kata tersebut terdapat

didokumen mana. Contoh kata “BELAJAR” terjadi dalam dokumen 3, 8, 22, 56, 68, and 92 dan kata “SEO“ terjadi dalam dokumen 2, 8, 15, 22, 68, and 77.

Google dalam melakukan pencarian akan melakukan 3 langkah yang dikenal dengan istilah CIS (Crawling, Indexing, Serving, yang akan menjelaskan secara lebih detail tentang ketiga langkah tersebut.

Crawling adalah sebuah proses dimana Googlebot menemukan situs baru atau situs yang telah diupdate dan ditambahkan pada indeks *Google*. *Google* menggunakan seperangkat besar komputer dalam melakukan *crawling* terhadap milyaran halaman situs. Program yang dibuat untuk melakukan tugas ini disebut *Googlebot* (kadang ada yang menyebutnya robot, bot atau spider). Saat *Googlebot* mengunjungi sebuah situs, *Googlebot* akan mengunjungi setiap link yang terdapat pada situs tersebut (hal ini akan berhubungan dengan sitemap-topik yang akan saya bahas pada kesempatan yang lain) dan menempatkan tiap halaman pada daftar halaman yang akan di-*crawl*. Menambahkan situs baru, mengganti situs yang terdahulu dan broken link akan dicatat pada langkah ini. Google tidak menerima pembayaran untuk melakukan *crawling* situs, dan google secara profesional akan memisahkan hasil pencarian dengan iklan yang ada pada *AdWord*.

Googlebot akan memproses setiap halaman yang telah di *crawl* untuk disusun menjadi satu indeks yang sangat besar menurut kata kunci yang ditemukannya dan berdasarkan lokasi dari tiap halaman. Sebagai tambahan, *Google* memproses informasi kata kunci yang terkandung pada meta tags, title tags, dan atribut ALT yang terdapat pada img tags. *Googlebot* dapat memproses banyak, tapi tidak semua, jenis content. Sebagai contoh, *Google* tidak dapat memproses file media dan juga halaman dinamis (menggunakan bahasa pemrograman php,asp,java dan lainnya). Jadi, kata kunci pada media tersebut tidak dapat di proses oleh *Google*.

Ketika user memasukkan kata kunci, mesin *Google* akan mencari halaman yang tepat pada databasenya sesuai dengan permintaan user dan akan menyajikan hasil yang paling relevan kepada user. Tingkat relevansi ini ditentukan oleh lebih dari 200 faktor, yang salah satunya adalah PageRank. PageRank adalah alat ukur seberapa baik halaman website Anda dengan berdasarkan pada link dari situs lain. Mudahnya, setiap link dari situs lain yang terhubung pada halaman situs Anda akan menambah *PageRank* situs Anda. Tidak semua link bernilai sama, *Google* sampai sekarang masih bekerja keras untuk memperbaiki penilaiannya dengan cara mengidentifikasi link spam dan tindakan negatif lain yang berdampak pada hasil pencarian. Jenis link yang paling baik adalah yang memeberikan link berdasarkan kualitas dari situs Anda.

Yahoo

Pada awalnya *Yahoo!* Search merupakan *search engine* yang bergantung pada *Google* untuk menampilkan hasilnya. Baru pada awal 2003, *Yahoo!* Mulai menjadi search engine yang berdiri sendiri. Pada Oktober 2007, *Yahoo!* Search menambahkan fitur Search Assist yang mana membantu memberikan ide kepada pgunanya sesuai dengan yang telah ketik. Lalu pada July 2008, *Yahoo* mengumumkan produk baru yang disebut “Build Your Own Search Services” (BOSS). Produk ini memberikan kesempatan kepada para pengembang untuk membuat search engine mereka sendiri dengan menggunakan system di *Yahoo!*Search. Cara *Yahoo* Search bekerja adalah membandingkan hasil ketikan di textbox search lalu mencarikan link ke halaman dengan format HTML dan juga beberapa format yang umum antara PDF, Excelspreadsheet, PowerPoint, dokumen Word, RSS/XML dan lain sebagainya. Dan

untuk anggota dari pada *Yahoo!* diberikan fasilitas tambahan yaitu *Advanced Search Interface*.

Altavista

Altavista menggunakan teori *neural network* untuk mesin pencariinya. Mesin pencari Altavista dinamakan *ranknet*. Bagi yang pernah belajar *Artificial Intelligence* (Intelegensi Buatan), mungkin akan mengerti dasar dari teori ini. Secara sederhana dapat dijelaskan demikian, bahwa suatu mesin yang menggunakan teori ini diharapkan dapat belajar dengan memperbandingkan variabel-variabel yang tersimpan didalamnya, semakin lama suatu mesin tersebut dipakai maka variabel-variabel yang tersimpan akan semakin banyak, karena variabel semakin banyak diharapkan mesin tersebut akan semakin pandai. Oleh sebab itu mesin pencari di alta vista disebut juga *learning machine* (mesin yang dapat belajar sendiri). Mungkin sedikit kurang jelas dengan apa yang dimaksud dengan variabel, variabel itu bisa bermacam-macam tergantung tujuan dari mesin tersebut, misalnya untuk mesin pencari web ini : penempatan kata kunci, jumlah pengunjung web, dll dapat digunakan sebagai variabel. Mesin pencari ini dalam belajar bukan berarti tidak tentu arah, tetapi tergantung dari algoritma-algoritma yang ada didalamnya yang memberikan prioritas-prioritas terhadap variable yang tersimpan didalamnya yang mempengaruhi hasil dan ranking dari yang dicarinya. Perlu diketahui bahwa mesin pencari alta vista mempunyai 569 algoritma (mungkin saat ini sudah lebih) yang terkandung didalamnya yang tentunya mempunyai fungsi yang bermacam-macam.

Sistem Temu Kembali Informasi

Sistem temu balik informasi merupakan kegiatan yang bertujuan untuk menyediakan dan memasok informasi bagi pemakai sebagai jawaban atas permintaan atau berdasarkan kebutuhan pemakai. Hasugian (2006: 2) menyatakan bahwa pada dasarnya sistem temu balik informasi adalah suatu proses untuk mengidentifikasi, kemudian memanggil (*retrieve*) suatu dokumen dari suatu simpanan (*file*), sebagai jawaban atas permintaan informasi. Lancaster dalam Hardi (2006: 1) menyatakan bahwa “temu kembali informasi sebagai suatu proses pencarian dokumen dengan menggunakan istilah luas untuk mengidentifikasi dokumen yang berhubungan dengan subjek tertentu”. Artinya dalam proses penemuan informasi perlu digunakan istilah-istilah tertentu. Tague-Sutcliffe dalam Hasugian (2006: 3) menyatakan bahwa, “Tujuan utama sistem temu kembali informasi adalah untuk menemukan dokumen yang sesuai dengan kebutuhan informasi pengguna secara efektif dan efisien, sehingga dapat memberikan kepuasan baginya”. Ada beberapa fungsi utama sistem temu balik informasi seperti yang dinyatakan Chowdhury (1999: 3) bahwa ada tujuh fungsi utama sistem temu balik informasi yang antara lain adalah:

1. Untuk mengidentifikasi informasi (sumber informasi) yang relevan dengan bidang-bidang yang sesuai dengan minat dan tujuan komunitas pemakai;
2. Untuk menganalisis isi dari sumber informasi (dokumen);
3. Untuk merepresentasikan isi dan sumber informasi yang telah dianalisis dengan cara yang sesuai untuk kemudian menyesuaikannya dengan permintaan pemakai;
4. Untuk menganalisis permintaan-permintaan pemakai dan merepresentasikannya ke dalam bentuk yang disesuaikan, untuk disesuaikan dengan database;
5. Untuk menyesuaikan pernyataan penelusuran dengan database;

6. Untuk menemukan informasi yang relevan;
7. Untuk membuat penyesuaian kebutuhan pada dasar sistem arus balik dari pemakai.

Dari uraian diatas disimpulkan bahwa sistem temu balik informasi merupakan proses pencarian kembali informasi sesuai dengan kebutuhan pencari informasi. Dalam proses perolehan informasi pencari merumuskan pertanyaan (*query*) atau menggunakan istilah-istilah berkaitan dengan informasi yang dibutuhkan.

Komponen Sistem Temu Balik Informasi

Sistem temu balik informasi terdiri dari komponen-komponen yang saling berkaitan satu sama lain. Chowdury (1999) menyatakan bahwa pada intinya dalam sistem temu balik informasi terdapat tiga komponen utama yang saling mempengaruhi, yaitu (1) kumpulan dokumen; (2) kebutuhan informasi pengguna, dan (3) proses pencocokan (*matching*) antara keduanya. Pernyataan yang sama juga diuraikan Singhal (2000) bahwa Secara garis besar komponen sistem temu balik informasi terdiri dari pemakai (*user*), dokumen, dan *matcher-machine*.

Adapun komponen-komponen sistem temu balik informasi menurut Singhal (2000) antara lain, (1) Pengguna; (2) Query; (3) Dokumen; (4) Indeks Dokumen dan (5) Pencocokan/ *Matcher Function*.

a. Pengguna

Pengguna mesin pencari adalah orang yang menggunakan atau memanfaatkan mesin pencarian dalam rangka kegiatan pengelolaan dan pencarian informasi. Berdasarkan perannya, pengguna mesin pencari dibedakan atas 2 (dua) kelompok yaitu pengguna (*user*) dan pengguna akhir (*end user*).

b. Query

Query adalah format bahasa permintaan yang di *input* (dimasukan) oleh pengguna ke dalam mesin pencari. Dalam *interface* (antar muka) mesin pencari selalu disediakan kolom/ruas sebagai tempat bagi pengguna untuk mengetikkan (menuliskan) *query* nya.

c. Dokumen

Dokumen adalah istilah yang digunakan untuk seluruh bahan pustaka, apakah itu artikel, buku, laporan penelitian dan sebagainya. Seluruh bahan pustaka dapat disebut sebagai dokumen.

d. Indeks Dokumen

Indeks adalah daftar istilah atau kata (*list of terms*). Dokumen yang dimasukkan/disimpan dalam *database* diwakili oleh indeks, Indeks itu disebut indeks dokumen.

e. Pencocokkan (*Matcher Function*)

Pencocokkan istilah (*query*) yang dimasukkan oleh pengguna dengan indeks dokumen yang tersimpan dalam *database* adalah dilakukan oleh mesin komputer. Komputerlah yang melakukan proses pencocokkan itu dalam waktu yang sangat singkat sesuai dengan kecepatan *memory* dan *processing* yang dimiliki oleh komputer itu. Dari beberapa uraian di atas disimpulkan bahwa sistem temu balik informasi memiliki komponen-komponen penyusun yang paling sedikit terdiri dari tiga bagian yaitu dokumen, pencari informasi dan proses pencocokan atau penghubung antara dokumen dan pencari informasi. Dan lebih rincinya sistem temu balik informasi terdiri atas lima komponen yaitu pengguna, query, dokumen, indeks dokumen dan pencocokan.

Pengukuran Efektivitas Temu Balik Informasi

Efektif yang merupakan kata dasar efektivitas dalam bahasa Inggris *effective* didefinisikan *producing the result that is wanted or intended* dan definisi sederhananya *coming into use* maksudnya yaitu terpakai atau digunakan dalam mencapai suatu tujuan (Oxford Learner's Pocket Dictionary, 2008: 143). Kamus Besar Bahasa Indonesia (2002: 584) mendefinisikan efektif dengan ada efeknya (akibatnya, pengaruhnya, kesannya) atau dapat membawa hasil, berhasil guna (usaha, tindakan) dan efektivitas diartikan keadaan berpengaruh, berkesan atau keberhasilan (usaha, tindakan). The Liang Gie dalam Ensiklopedi Administrasi (1989:108) mendefinisikan, Efektivitas sebagai suatu keadaan yang mengandung pengertian mengenai terjadinya efek atau akibat yang dikehendaki. Jika seseorang melakukan suatu perbuatan dengan maksud tertentu yang memang dikehendaki, maka orang itu dikatakan efektif kalau memang menimbulkan akibat dari yang dikehendakinya. Sederhananya, efektivitas merujuk pada kemampuan untuk memiliki tujuan yang tepat atau mencapai tujuan yang telah ditetapkan. Efektivitas juga berhubungan dengan masalah bagaimana pencapaian tujuan atau hasil yang diperoleh, kegunaan atau manfaat dari hasil yang diperoleh, tingkat daya fungsi unsur atau komponen, serta masalah tingkat kepuasan pengguna/*client*. Berkaitan dengan temu balik informasi, efektivitas temu balik informasi merupakan kemampuan dari sistem untuk memanggil berbagai dokumen dari suatu basis data sesuai dengan permintaan pengguna.

Efektivitas temu balik informasi tersebut dapat diukur dengan menghitung rasio atau perbandingan dari *recall* (perolehan) dan *precision* (ketepatan) (Pao, 1989: 225). *Recall* (perolehan) berhubungan dengan kemampuan sistem untuk memanggil dokumen yang relevan. *Precision* (ketepatan) berkaitan dengan kemampuan sistem untuk tidak memanggil dokumen yang tidak relevan. Untuk menghitung nilai *recall* (perolehan) dan *precision* (ketepatan) digunakan rumus NIAP (*non interpolated average precision*) (Pao, 1989: 226). Kondisi ideal dari keefektifan suatu sistem temu kembali informasi adalah apabila rasio *Recall* dan *Precision* sama besarnya (1:1). Dengan demikian, dapat disimpulkan bahwa bagian terpenting dalam proses temu balik informasi adalah ketika kebutuhan informasi pengguna tercapai dikarenakan *precision* (ketepatan) yang dihasilkan dalam penelusuran tinggi. Ukuran yang digunakan untuk menghitung ketepatan sistem temu balik informasi yaitu *recall*, *precision* dan NIAP.

Recall menemukan seluruh dokumen yang relevan dalam koleksi. *Recall* dapat dihitung dengan rumus:

$$Recall = \left(\frac{\text{Jumlah dokumen relevan ditemukan}}{\text{Jumlah Dokumen relevan dalam koleksi}} \right)$$

Dimana nilai *recall* tertinggi adalah 1, yang berarti seluruh dokumen dalam koleksi berhasil ditemukan. *Precision* diartikan menemukan hanya dokumen yang relevan saja dalam koleksi. Nilai *precision* tertinggi adalah 1, yang berarti seluruh dokumen yang ditemukan adalah relevan, dengan rumus :

$$Precision = \left(\frac{\text{Jumlah dokumen relevan ditemukan}}{\text{Jumlah Dokumen ditemukan}} \right)$$

Sedangkan untuk NIAP (*Non Interpolated Average Precision*) adalah penggabungan dari *recall* dan *precision*, yang dapat dihitung dengan rumus:

$$NIAP = \sum_{t=1}^n \frac{\text{precision pada dokumen ke } i}{\text{jumlah dokumen relevan dalam koleksi}}$$

Di mana n menunjukkan jumlah dokumen yang dicari hingga seluruh dokumen

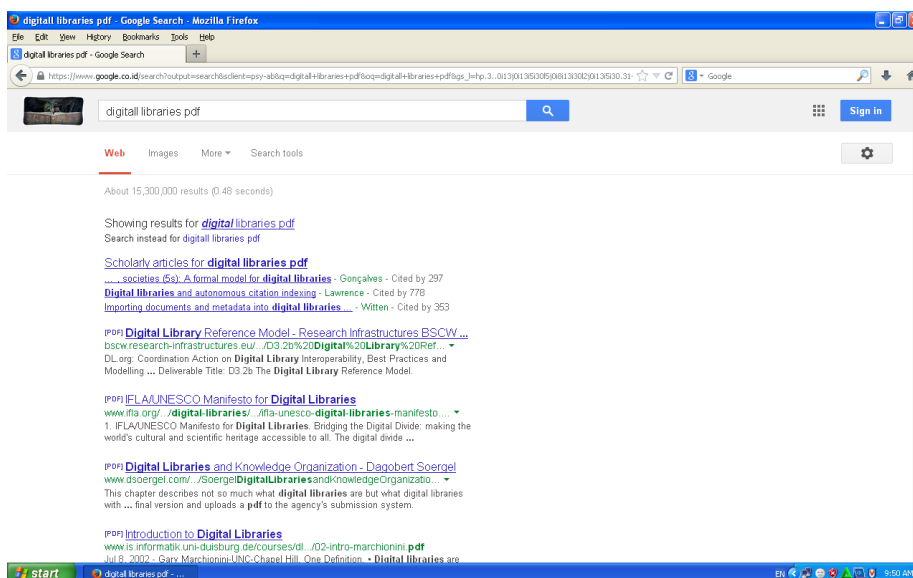
relevan ditemukan. Nilai NIAP tertinggi adalah 1, yang berarti seluruh dokumen relevan berhasil ditemukan dengan seluruh dokumen relevan tersebut ditempatkan pada urutan teratas dalam hasil pencarian. Nilai NIAP akan digunakan untuk mengecek kebenaran hasil pencarian dari perangkat lunak yang dibangun.

Hasil Percobaan

a. Google

Pengujian sistem temu kembali pada google.co.id, dilakukan dengan menggunakan keyword seperti pada tabel 1. File yang dicari yaitu file berjenis pdf. Page yang dianalisa yaitu page pertama, dengan asumsi bahwa page pertama merupakan page yang berisikan site dengan tingkat kunjungan tertinggi. Contoh hasil penelusuran dapat dilihat pada gambar 1.

Metode pencarian yang digunakan goggle.co.id hampir sama dengan dengan semua metode pencarian di yahoo.co.id dan altavista.com, sehingga perbandingannya bersifat setara. Metode pencarian dilakukan dengan cara biasa bukan *advance search*. Pola frase text pencarian yang dilakukan oleh google.co.id, yaitu text+pdf kemudian pdf. Dengan demikian pengambilan sampel dengan menggunakan halaman pertama saja bisa mewakili sampel site dengan kunjungan paling banyak.



Gambar 1 Hasil Penelusuran melalui google.co.id

Berdasarkan hasil pengujian dengan menggunakan 30 *keyword subject libraries*, diperoleh hasil untuk *recall*, *precision* dan *NIAP*. (Tabel 3). Sedangkan untuk rincian NIAP dapat dilihat pada tabel 4.

Tabel 3.
Data Recall, Precision dan NIAP untuk google.co.id

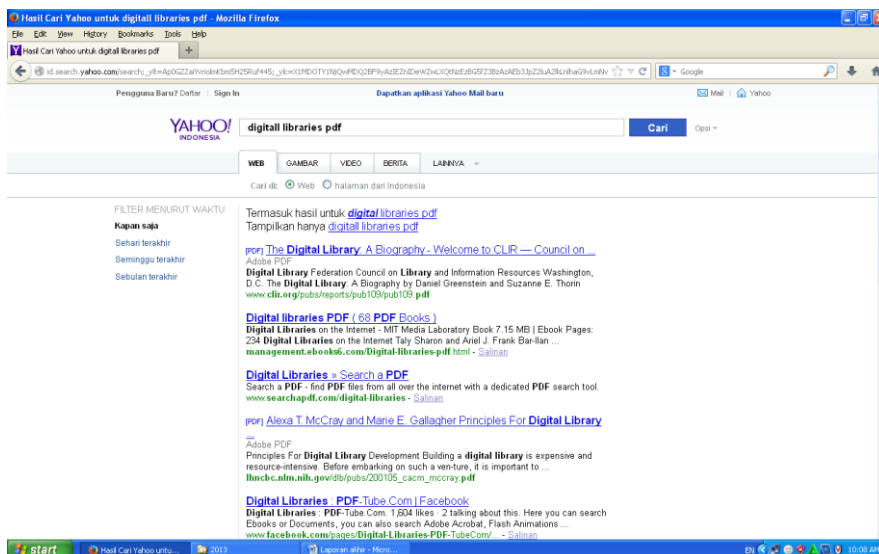
No	Subject	Keyword	Google.co.id		
			Recall	Precision	NIAP
1	Digital Libraries	Digital Libraries pdf	1	0.80	0.65
2	Classification	Classification pdf	1	0.80	0.65
3	Library Automation	Library Automation pdf	1	0.90	0.90
4	Integrated Library Systems	Integrated Library Systems pdf	1	0.60	0.55
5	Catalogs	Catalogs pdf	1	0.00	0.00
6	Library Circulation and Loans	Library Circulation and Loans pdf	1	0.60	0.50
7	Collection Management	Collection Management pdf	1	0.90	0.80
8	Collection Development	Collection Development pdf	1	0.90	0.90
9	Library Information Network	Library Information Network pdf	1	0.70	0.70
10	Library Decoration	Library Decoration pdf	1	0.3	0.30
11	Duplicate Book	Duplicate Book pdf	1	0.4	0.40
12	Bibliographical Libraries	Bibliographical Libraries pdf	1	0.7	0.70
13	Library Resources	Library Resources pdf	1	0.6	0.60
14	Library Administration	Library Administration pdf	1	0.5	0.45
15	Library Science	Library Science pdf	1	0.5	0.40
16	Public Libraries	Public Libraries pdf	1	0.9	0.90
17	Proprietary Libraries	Proprietary Libraries pdf	1	0.7	0.60
18	Private Libraries	Private Libraries pdf	1	0.8	0.65
19	Rental Libraries	Rental Libraries pdf	1	0.5	0.30
20	Research Libraries	Research Libraries pdf	1	0.8	0.50
21	School Libraries	School Libraries pdf	1	1	0.95
22	Small Libraries	Small Libraries pdf	1	0.5	0.35
23	Acquisitions Libraries	Acquisitions Libraries pdf	1	0.7	0.45
24	Book list	Book list pdf	1	0.8	0.65
25	Library Circulation	Library Circulation pdf	1	0.6	0.50
26	Data Processing	Data Processing pdf	1	1	0.75
27	Inventories	Inventories pdf	1	0.9	0.90
28	Library site	Library site pdf	1	0.5	0.25
29	Merger Libraries	Merger Libraries pdf	1	0	0.00
30	Shelving	Shelving pdf	1	0.8	0.60
Rerata			1	0.66	0.56

Tabel 4
Data NIAP untuk google.co.id

No	NIAP	precision / dokumen (scale 0,0.5,1)									
		1	2	3	4	5	6	7	8	9	10
1	0.65	1.00	1.00	1.00	1.00	1.00	0.50	0.50	0.50	0.00	0.00
2	0.65	1.00	1.00	1.00	1.00	1.00	0.50	0.50	0.50	0.00	0.00
3	0.90	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00
4	0.55	1.00	1.00	1.00	1.00	1.00	0.50	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.50	1.00	1.00	1.00	0.00	0.00	0.50	0.00	0.50	1.00	0.00
7	0.80	1.00	1.00	1.00	1.00	1.00	0.50	0.50	1.00	0.00	1.00
8	0.90	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00
9	0.70	1.00	1.00	1.00	1.00	1.00	0.00	1.00	1.00	0.00	0.00
10	0.30	0.00	0.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
11	0.40	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.70	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	0.00
13	0.60	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
14	0.45	1.00	0.50	1.00	0.00	1.00	0.00	1.00	0.00	0.00	0.00
15	0.40	1.00	1.00	1.00	0.00	0.00	0.50	0.00	0.50	0.00	0.00
16	0.90	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00
17	0.60	1.00	1.00	1.00	1.00	1.00	0.50	0.50	0.00	0.00	0.00
18	0.65	1.00	1.00	1.00	1.00	1.00	0.50	0.50	0.50	0.00	0.00
19	0.30	1.00	0.50	0.50	0.50	0.50	0.00	0.00	0.00	0.00	0.00
20	0.50	1.00	1.00	0.50	0.50	0.50	0.50	0.50	0.50	0.00	0.00
21	0.95	1.00	1.00	0.50	1.00	1.00	1.00	1.00	1.00	1.00	1.00
22	0.35	1.00	1.00	0.50	0.00	0.00	0.50	0.00	0.50	0.00	0.00
23	0.45	1.00	1.00	0.50	0.50	0.00	0.50	0.00	0.50	0.50	0.00
24	0.65	1.00	1.00	1.00	1.00	1.00	0.50	0.50	0.50	0.00	0.00
25	0.50	1.00	1.00	1.00	1.00	0.00	0.00	0.50	0.50	0.00	0.00
26	0.75	1.00	1.00	1.00	1.00	1.00	0.50	0.50	0.50	0.50	0.50
27	0.90	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
28	0.25	0.50	0.50	0.50	0.00	0.00	0.00	0.50	0.50	0.00	0.00
29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	0.60	1.00	1.00	1.00	1.00	0.00	0.50	0.50	0.50	0.50	0.00
Rerata	0.56	0.88	0.85	0.83	0.72	0.63	0.47	0.42	0.45	0.25	0.12

b. Yahoo

Sedikit berbeda dengan google.co.id, hasil sistem temu kembali yang dimunculkan oleh yahoo.co.id tidak memunculkan total dokumen dan waktu pencarian. Fitur yang dimunculkan justru langsung fitur *advance search* yang terletak di pojok kiri atas mesin yahoo.co.id. Metode pencarian yang diujikan sama dengan perlakuan pada google.co.id, yaitu dengan menggunakan subject pengujian tabel 1. Secara lengkap hasil pengujian bisa dilihat pada tabel 4 dan 5



Gambar 2. Hasil penelusuran melalui yahoo.co.id

Tabel 4
Data Recall, Precision dan NIAP untuk yahoo.co.id

No	Subject	Keyword	Yahoo.co.id		
			Recall	Precision	NIAP
1	Digital Libraries	Digital Libraries pdf	1	0.50	0.35
2	Classification	Classification pdf	1	0.70	0.50
3	Library Automation	Library Automation pdf	1	0.60	0.60
4	Integrated Library Systems	Integrated Library Systems pdf	1	0.60	0.45
5	Catalogs	Catalogs pdf	1	0.50	0.35
6	Library Circulation and Loans	Library Circulation and Loans pdf	1	0.60	0.40
7	Collection Management	Collection Management pdf	1	0.80	0.65
8	Collection Development	Collection Development pdf	1	0.90	0.80
9	Library Information Network	Library Information Network pdf	1	0.60	0.60
10	Library Decoration	Library Decoration pdf	1	0.50	0.30
11	Duplicate Book	Duplicate Book pdf	1	0.50	0.30
12	Bibliographical Libraries	Bibliographical Libraries pdf	1	0.60	0.60
13	Library Resources	Library Resources pdf	1	0.60	0.50
14	Library Administration	Library Administration pdf	1	0.40	0.35
15	Library Science	Library Science pdf	1	0.40	0.30
16	Public Libraries	Public Libraries pdf	1	0.80	0.80
17	Proprietary Libraries	Proprietary Libraries pdf	1	0.60	0.50
18	Private Libraries	Private Libraries pdf	1	0.70	0.45
19	Rental Libraries	Rental Libraries pdf	1	0.60	0.30
20	Research Libraries	Research Libraries pdf	1	0.70	0.45
21	School Libraries	School Libraries pdf	1	0.80	0.80
22	Small Libraries	Small Libraries pdf	1	0.60	0.50

No	Subject	Keyword	Yahoo.co.id		
			Recall	Precision	NIAP
23	Acquisitions Libraries	Acquisitions Libraries pdf	1	0.80	0.40
24	Book list	Book list pdf	1	0.60	0.45
25	Library Circulation	Library Circulation pdf	1	0.60	0.30
26	Data Processing	Data Processing pdf	1	0.90	0.60
27	Inventories	Inventories pdf	1	0.70	0.70
28	Library site	Library site pdf	1	0.60	0.30
29	Merger Libraries	Merger Libraries pdf	1	0.00	0.00
30	Shelving	Shelving pdf	1	0.60	0.70
Rerata			1	0.61	0.48

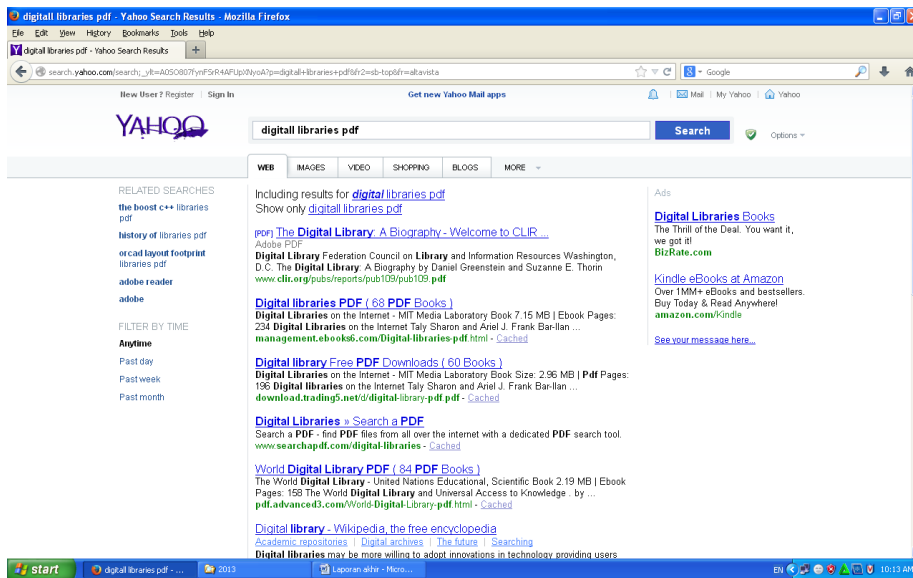
Tabel 5
Data NIAP untuk yahoo.co.id

No	NIAP	precision / dokumen (scale 0,0.5,1)									
		1	2	3	4	5	6	7	8	9	10
1	0.35	0.00	0.00	1.00	0.00	1.00	0.50	0.50	0.50	0.00	0.00
2	0.50	0.00	1.00	0.50	1.00	1.00	0.50	0.50	0.50	0.00	0.00
3	0.60	0.00	0.00	1.00	0.00	1.00	1.00	1.00	1.00	1.00	0.00
4	0.45	0.00	0.00	1.00	0.50	1.00	0.50	0.00	0.00	1.00	0.50
5	0.35	0.00	1.00	0.50	0.00	0.00	1.00	0.50	0.00	0.00	0.50
6	0.40	0.00	1.00	0.50	0.00	0.00	0.50	0.50	0.50	1.00	0.00
7	0.65	0.00	1.00	1.00	0.50	1.00	0.50	0.50	1.00	0.00	1.00
8	0.80	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00
9	0.60	0.00	1.00	1.00	1.00	1.00	0.00	1.00	1.00	0.00	0.00
10	0.30	0.00	0.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
11	0.30	0.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.60	0.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	0.00
13	0.50	0.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
14	0.35	0.00	0.50	1.00	0.00	1.00	0.00	1.00	0.00	0.00	0.00
15	0.30	0.00	1.00	1.00	0.00	0.00	0.50	0.00	0.50	0.00	0.00
16	0.80	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00
17	0.50	0.00	1.00	1.00	1.00	1.00	0.50	0.50	0.00	0.00	0.00
18	0.45	0.00	1.00	0.50	0.50	1.00	0.50	0.50	0.50	0.00	0.00
19	0.30	0.00	0.50	0.50	0.50	0.50	0.00	0.00	0.50	0.50	0.00
20	0.45	0.00	1.00	1.00	0.50	0.50	0.50	0.50	0.50	0.00	0.00
21	0.80	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00
22	0.50	0.00	1.00	1.00	0.00	1.00	0.50	0.00	0.50	0.00	1.00
23	0.40	0.50	0.50	0.50	0.50	0.00	0.50	0.50	0.50	0.50	0.00
24	0.45	0.00	0.00	1.00	1.00	1.00	0.50	0.50	0.50	0.00	0.00
25	0.30	0.00	0.00	0.50	0.50	0.50	0.50	0.50	0.50	0.00	0.00
26	0.60	1.00	1.00	0.50	0.00	1.00	0.50	0.50	0.50	0.50	0.50
27	0.70	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00

No	NIAP	precision / dokumen (scale 0,0.5,1)									
		1	2	3	4	5	6	7	8	9	10
28	0.30	0.50	0.50	0.50	0.50	0.00	0.00	0.50	0.50	0.00	0.00
29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	0.70	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.50	0.50
Rerata	0.48	0.10	0.63	0.78	0.53	0.72	0.53	0.48	0.48	0.30	0.20

c. Altavista

Altavista.com merupakan mesin pencarian yang telah diakuisisi oleh yahoo. Meskipun telah diakuisisi, tampilan dengan yahoo.co.id secara umum berbeda. Altavista masih mencatumkan ciri kasnya yaitu fitur *related searches* yang tidak muncul pada mesin pencarian yahoo.co.id. Secara cermat bisa dilihat perbandingannya antara gambar 2 dan 3.



Gambar 3. Hasil penelusuran melalui altavista.com

Berdasarkan pengujian yang sama, dengan subject pada tabel 4.1 maka altavista memperlihatkan hasil yang jauh menurun bila dibandingkan dengan yahoo.co.id. Walaupun begitu, altavista mempunyai satu kelebihan yaitu menjadi mesin pencari link site yang lebih banyak bila dibandingkan dengan google.co.id dan yahoo.co.id. Secara lengkap hasil ujicoba bisa dilihat pada tabel 6 dan 7.

Tabel 6
Data Recall, Precision dan NIAP untuk altavista.com

No	Subject	Keyword	altavista.com		
			Recall	Precision	NIAP
1	Digital Libraries	Digital Libraries pdf	1	0.30	0.20
2	Classification	Classification pdf	1	0.40	0.25
3	Library Automation	Library Automation pdf	1	0.40	0.40
4	Integrated Library Systems	Integrated Library Systems pdf	1	0.40	0.30
5	Catalogs	Catalogs pdf	1	0.50	0.35
6	Library Circulation and Loans	Library Circulation and Loans pdf	1	0.50	0.30
7	Collection Management	Collection Management pdf	1	0.40	0.25
8	Collection Development	Collection Development pdf	1	0.30	0.30
9	Library Information Network	Library Information Network pdf	1	0.30	0.30
10	Library Decoration	Library Decoration pdf	1	0.50	0.50
11	Duplicate Book	Duplicate Book pdf	1	0.40	0.40
12	Bibliographical Libraries	Bibliographical Libraries pdf	1	0.50	0.50
13	Library Resources	Library Resources pdf	1	0.50	0.40
14	Library Administration	Library Administration pdf	1	0.30	0.25
15	Library Science	Library Science pdf	1	0.40	0.30
16	Public Libraries	Public Libraries pdf	1	0.40	0.40
17	Proprietary Libraries	Proprietary Libraries pdf	1	0.30	0.25
18	Private Libraries	Private Libraries pdf	1	0.40	0.20
19	Rental Libraries	Rental Libraries pdf	1	0.30	0.15
20	Research Libraries	Research Libraries pdf	1	0.40	0.25
21	School Libraries	School Libraries pdf	1	0.60	0.60
22	Small Libraries	Small Libraries pdf	1	0.30	0.25
23	Acquisitions Libraries	Acquisitions Libraries pdf	1	0.40	0.20
24	Book list	Book list pdf	1	0.30	0.25
25	Library Circulation	Library Circulation pdf	1	0.40	0.20
26	Data Processing	Data Processing pdf	1	0.70	0.50
27	Inventories	Inventories pdf	1	0.60	0.60
28	Library site	Library site pdf	1	0.40	0.20
29	Merger Libraries	Merger Libraries pdf	1	0.00	0.00
30	Shelving	Shelving pdf	1	0.30	0.30
Rerata			1	0.40	0.31

Tabel 7
Data NIAP untuk altavista.com

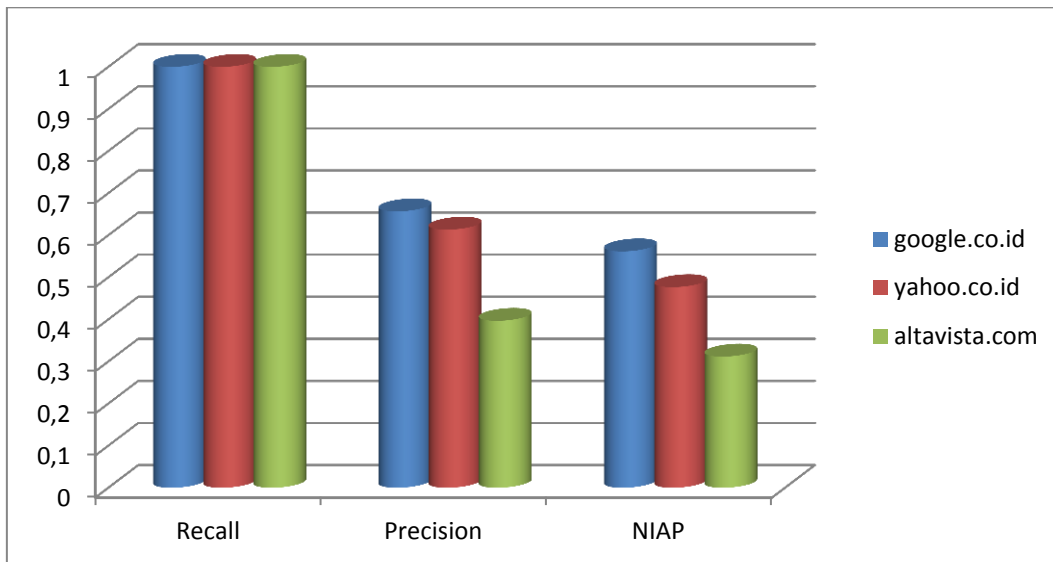
No	NIAP	precision / dokumen (scale 0,0.5,1)									
		1	2	3	4	5	6	7	8	9	10
1	0.20	0.00	0.00	0.00	0.00	1.00	0.00	0.50	0.50	0.00	0.00
2	0.25	0.00	0.00	0.50	0.00	1.00	0.50	0.00	0.50	0.00	0.00
3	0.40	0.00	0.00	1.00	0.00	1.00	0.00	0.00	1.00	1.00	0.00
4	0.30	0.00	0.00	1.00	0.00	0.00	0.50	0.00	0.00	1.00	0.50
5	0.35	0.00	1.00	0.50	0.00	0.00	1.00	0.50	0.00	0.00	0.50
6	0.30	0.00	0.00	0.50	0.00	0.50	0.00	0.50	0.50	1.00	0.00
7	0.25	0.00	1.00	0.00	0.50	0.00	0.50	0.50	0.00	0.00	0.00
8	0.30	0.00	1.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00
9	0.30	0.00	0.00	1.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
10	0.50	0.00	0.00	1.00	1.00	1.00	0.00	1.00	1.00	0.00	0.00
11	0.40	0.00	1.00	1.00	1.00	0.00	1.00	0.00	0.00	0.00	0.00
12	0.50	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	0.00
13	0.40	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
14	0.25	0.00	0.50	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.00
15	0.30	0.00	1.00	1.00	0.00	0.00	0.50	0.00	0.50	0.00	0.00
16	0.40	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.00	1.00	0.00
17	0.25	0.00	0.00	1.00	0.00	1.00	0.00	0.50	0.00	0.00	0.00
18	0.20	0.00	0.00	0.50	0.50	0.00	0.00	0.50	0.50	0.00	0.00
19	0.15	0.00	0.00	0.50	0.00	0.50	0.00	0.00	0.00	0.50	0.00
20	0.25	0.00	0.00	1.00	0.50	0.00	0.50	0.50	0.00	0.00	0.00
21	0.60	1.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00
22	0.25	0.00	0.00	1.00	0.00	1.00	0.50	0.00	0.00	0.00	0.00
23	0.20	0.00	0.50	0.50	0.00	0.00	0.00	0.50	0.00	0.50	0.00
24	0.25	0.00	0.00	1.00	0.00	1.00	0.00	0.50	0.00	0.00	0.00
25	0.20	0.00	0.00	0.50	0.00	0.00	0.50	0.50	0.50	0.00	0.00
26	0.50	1.00	1.00	0.50	0.00	1.00	0.50	0.00	0.00	0.50	0.50
27	0.60	0.00	0.00	1.00	0.00	1.00	1.00	1.00	1.00	0.00	1.00
28	0.20	0.50	0.00	0.50	0.00	0.00	0.00	0.50	0.50	0.00	0.00
29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	0.30	0.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	0.00	0.00
Rerata	0.31	0.08	0.27	0.62	0.25	0.57	0.30	0.42	0.25	0.25	0.12

Pembahasan

Berdasarkan rekapitulasi dari data pengujian *recall* dan *precision* (tabel 8), memperlihatkan bahwa semua *search engine* mempunyai nilai *recall* 1, artinya baik google, yahoo dan altavista mempunyai *recall* yang bagus. Hampir semua site menunjukkan link yang betul dan tidak ada link yang salah ataupun *blank site*. Setiap link ketika ditelusur lebih lanjut memberikan konten yang ada isinya. Jadi untuk *recall* untuk ketiga *search engine* tersebut bernilai sama.

Tabel 8
Rekapitulasi *Recall*, *precision* dan NIAP

Keterangan	Hasil		
	Recall	Precision	NIAP
google.co.id	1	0.66	0.56
yahoo.co.id	1	0.61	0.48
altavista.com	1	0.40	0.31



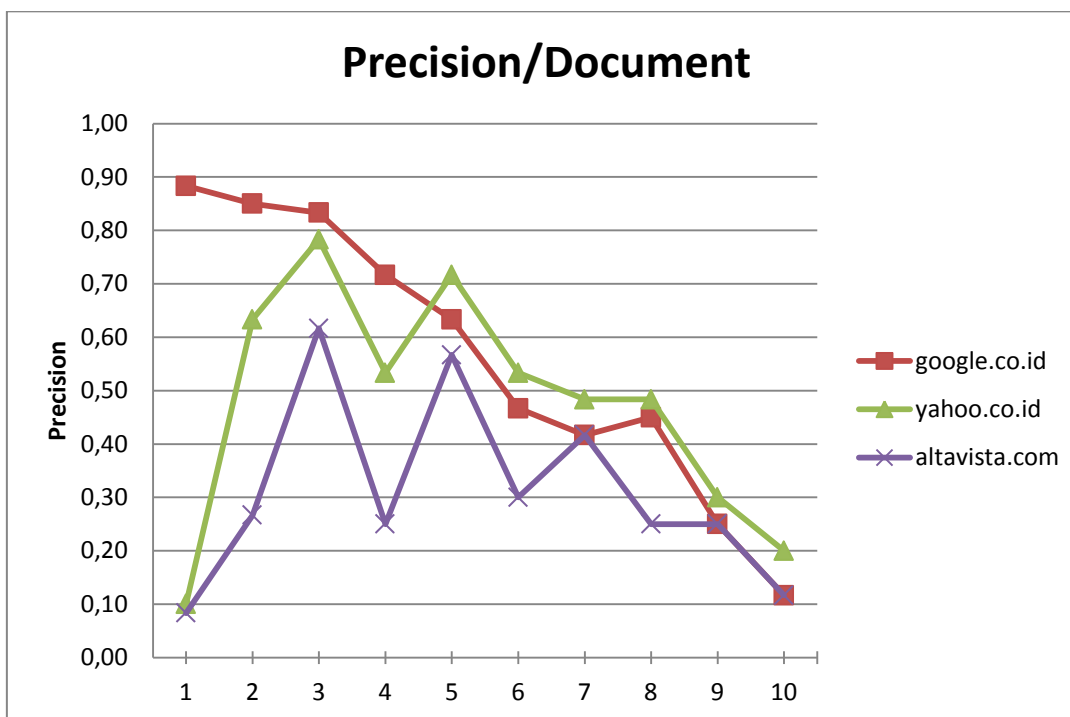
Gambar 4. Grafik Perbandingan google, yahoo dan altavista

Hasil pengujian *precision* menunjukkan hasil yang berbeda, dimana google menempati *precision* tertinggi sebesar 0.66 sedangkan yahoo (0.61) dan googaltavista (0.40). Nilai *precision* ini menunjukkan bahwa google jauh lebih akurat memberikan hasil sesuai dengan *keyword* yang dicari. Beberapa dukungan kepustakaan google seperti *google scholar* dan *google doc* memberikan dampak terhadap pola pencarian yang dilakukan. Dengan *keyword text+pdf*, maka google langsung bisa menunjukkan site dan link dokumen yang dimaksud langsung pada konten pdf yang dimiliki oleh site tersebut. Ini berbeda dengan yahoo dan altavista, yang cenderung menghasilkan link dalam link. Artinya pengguna harus melakukan penelusuran lebih lanjut untuk mendapatkan hasil yang diinginkan, itupun terkadang harus memasukkan *keyword* lagi untuk pencarian. Jadi yahoo dan altavista kurang mendapatkan dukungan data.

Pada awal search engine ini muncul dan ketika belum adanya *google scholar* dan *google doc*, yahoo dan altavista menduduki rating tertinggi. Bahkan altavista menduduki peringkat teratas untuk akurasi dokumen yang ditampilkan. Fakta ini juga yang mengakibatkan yahoo melakukan akuisisi terhadap altavista (Indrajit, 2006). Namun seiring dengan dukungan *google scholar* dan *google doc*, keadaan ini berubah dan lambat laun menempatkan google menduduki peringkat *precision* teratas (Dwivedi, 2013).

Tabel 9
NIAP untuk google, yahoo dan altavista

	precision / dokumen (scale 0,0.5,1)									
Keterangan	1	2	3	4	5	6	7	8	9	10
google.co.id	0.88	0.85	0.83	0.72	0.63	0.47	0.42	0.45	0.25	0.12
yahoo.co.id	0.10	0.63	0.78	0.53	0.72	0.53	0.48	0.48	0.30	0.20
altavista.com	0.08	0.27	0.62	0.25	0.57	0.30	0.42	0.25	0.25	0.12



Gambar 5. Grafik Perbandingan NIAP google, yahoo dan altavista

Sedangkan untuk pengujian NIAP (*Non Interpolated Average Precision*) memperlihatkan hasil bahwa google (0.56), yahoo (0.48) dan altavista (0.31). NIAP merupakan pola pencarian dari sistem temu kembali informasi. NIAP yang bagus harus memperlihatkan hasil bahwa nilai tertinggi harus dimiliki oleh site teratas. Semakin ke bawah maka semakin rendah nilai NIAP. Berdasarkan grafik rerata nilai NIAP, maka google memberikan hasil NIAP dengan pembobotan yang bagus. Google selalu menempatkan dokumen dengan precision tertinggi pada urutan teratas. Yahoo dan altavista tidak selalu memiliki pola tersebut, artinya mesin pencari yahoo dan altavista cenderung untuk tidak memberikan hasil yang akurat pada *site link* untuk urutan yang pertama. Secara lengkap hasil NIAP dapat dilihat pada gambar 5.5.

Pola grafik NIAP bisa dijadikan acuan, seberapa besar akurasi dari mesin pencarian. Artinya, ketika menggunakan mesin pencarian maka hasil yang muncul harusnya berurutan dari hasil yang paling akurat sampai dengan yang terendah akurasi. Jika mesin pencari memiliki kaidah seperti ini, maka mesin pencari tersebut bisa dikatakan mempunyai sistem temu kembali yang bagus. Merujuk konsep ini maka secara peringkat untuk sistem temu kembali google mempunyai nilai sistem temu kembali informasi yang paling baik.

Simpulan

Sistem temu kembali informasi yang dimiliki oleh mesin pencari google, yahoo dan altavista mempunyai karakteristik masing masing. Google lebih cenderung untuk menghasilkan *precision* yang lebih baik bila dibandingkan dengan yahoo dan altavista. Sedangkan yahoo dan altavista masih membutuhkan penelusuran lebih lanjut.

Hasil uji coba memberikan hasil google (*recall*:1, *precision*:0.66, NIAP:0.56), yahoo (*recall*:1, *precision*:0.61, NIAP:0.48) dan altavista (*recall*:1, *precision*:0.40, NIAP:0.31). Hasil tersebut menunjukkan bahwa kemampuan *recall* untuk semua mesin pencari sama sedangkan kemampuan *precision* dan NIAP, mesin pencari google lebih baik bila disbanding dengan yahoo dan altavista.

Daftar Pustaka

- Chowdhury, G.G. 2004. *Introduction to Modern Information Retrieval*. London : Library Association.
- Dwipendi, Nripendra. 2013. *Statistical Analysis of Search Engines*. International Journal of Computer Theory and Engineering vol 5 no 2, 2013.
- Grishman, Ralph .1997. *Information Extraction: Techniques and Challenges* Lecture Notes in Computer Science, Vol. 1299, Springer-Verlag.
- Indrajit, Richardus Eko, dkk. 2006. *Pemanfaatan Search Engine Sebagai Sarana Penunjang Proses Pembelajaran*. Yogyakarta : Andi Offset.
- Krikelas, J. 1983. Information-seeking behavior: Patterns and concepts. *Journal of Documentation*, 19(2):5-20.
- Lavrenko, Victor., and Bruce Croft, W., 2001. *Relevance-Based Language Models*. Center for Intelligent Information Retrieval, Department of Computer Science, University of Massachusetts, United States.
- Mandala, Rila. 2006. *Evaluasi Efektivitas Metode Mechine-Learning pada Search Engine*. Seminar Nasional Aplikasi Teknologi Informasi 2006 (SNATI 2006) ISSN: 1907-5022 Yogyakarta, 17 Juni 2006
- Pao, Miranda Lee. 1989. *Concepts of Information Retrieval*. Englewood, Colorado: Libraries Unlimited.
- Robertson, S.E., van Rijsbergen, C.J., and Porter, M.F., 1981. *Probabilistic Model of Indexing And Searching*. Oddy Etal(eds), Information Retrieval Research, Butterworths.
- Singhal, Amit., 2000. *Modern Information Retrieval: A Brief Overview*. Google, Inc., Sillicon Valley, California.
- Van Rijsbergen, C.J., 1979. *Information Retrieval*. Department of Computing Science, University of Glasgow.