

## Teknologi Web Semantik Untuk Bibliografi Perpustakaan

R. Rhoedy Setiawan<sup>1</sup>, Mukhamad Nurkamid<sup>2</sup>

<sup>1,2</sup>Fakultas Teknik, Universitas Muria Kudus, Kudus59327

E-mail:rhoedy\_05@yahoo.co.id, nurkamid@gmail.com

### ABSTRAK

Sumber data yang terbesar dan sangat berharga untuk setiap pengguna karena didalam web kumpulan dokumen-dokumen saling terhubung dan dapat diakses melalui koneksi internet. Setiap individu didunia bebas menambahkan konten dalam web-nya. Makin banyak informasi yang ditambahkan, makin besar ukuran web dan semakin sulit pula untuk mencari informasi yang benar-benar yang diinginkan. Tujuan perpustakaan, pengarsipan, yang ada di institusi adalah untuk memelihara dokumen yang memiliki harga keilmuan yang dapat diakses dimasa depan sama mudahnya diakses pada saat ini. Dengan semakin kecilnya media data, penyimpanan menggunakan media digital menjadi sangat menarik. Keuntungan penyimpanan menggunakan media digital adalah pencarian dapat dilakukan tidak hanya di antara masukan katalog elektronik, tetapi dapat menampilkan sepanjang keseluruhan isi dokumen. Peneliti dan pengembang web mengusulkan tambahan informasi bagi konten web yang sering diistilahkan dengan metadata. Istilah metadata ini sering dipakai dalam semantic web, yang menggambarkan pendekatan dalam menangani dan menyimpan dokumen. Saat ini dokumen mulai disajikan dalam format XML (*Extensible Markup Language*). Bahasa lain yang digunakan dalam mendukung visi Semantic Web selain XML adalah RDF/OWL, Selanjutnya RDF/OWL dengan kemampuan dan fasilitas yang dimilikinya mampu digunakan untuk merepresentasikan makna istilah dalam web hingga dapat diproses dalam mesin sehingga dapat dikelola dan dicari.

**Kata kunci** : Arsip, Proses, Cari

### 1. PENDAHULUAN

#### 1.1. Latar belakang

Perkembangan teknologi saat ini kian hari kian pesat, hampir semua bidang teknologi sekarang berbasis internet. Salah satu teknologi yang saat ini kita rasakan manfaatnya adalah *world wide web* atau dikenal dengan nama *web*. Sumber data yang terbesar dan sangat berharga untuk setiap pengguna karena didalam *web* kumpulan dokumen-dokumen saling terhubung dan dapat diakses melalui koneksi internet. Setiap individu didunia bebas menambahkan konten dalam *web*-nya. Makin banyak informasi yang ditambahkan, makin besar ukuran *web* dan semakin sulit pula untuk mencari informasi yang benar-benar yang diinginkan di *web*, oleh karena itu dibutuhkan suatu teknologi untuk mendapatkan informasi yang benar-benar diinginkan tersebut. Hal tersebut juga dialami para pengguna informasi dalam memanfaatkan teknologi internet, salah satunya dalam *web digital library*. Semakin berkembang karena dipandang sebagai peluang potensial bagi dunia akademik untuk menyimpan koleksi yang ada diperpustakaan seperti buku, jurnal ilmiah, majalah, tugas akhir, skripsi dan tesis maupun disertasi. Teknologi yang tepat untuk mengembangkan, membangun dan mengelola koleksi dokumen perpustakaan yang benar-benar diinginkan pengguna merupakan tujuan kedepan semua pihak. Tujuan perpustakaan, pengarsipan, yang ada di institusi adalah untuk memelihara dokumen yang memiliki harga keilmuan yang dapat diakses dimasa depan sama mudahnya diakses pada saat ini. Dengan semakin kecilnya media data, penyimpanan menggunakan media digital menjadi sangat menarik. Keuntungan penyimpanan menggunakan media digital adalah pencarian dapat dilakukan tidak hanya di antara masukan katalog elektronik, tetapi dapat menampilkan sepanjang keseluruhan isi dokumen. Permasalahan baru muncul disebabkan *floppy disk*, *hard disk* dan CD cepat ketinggalan jaman. Perkembangan sistem komputer dan komponennya lebih cepat dari pada perkembangan media digital, sehingga menyebabkan masalah *kompatibilitas*(kesenjangan percepatan hardware yang tidak diikuti percepatan software). Peneliti dan pengembang *web* mengusulkan tambahan informasi bagi konten web yang sering diistilahkan dengan metadata. Istilah metadata ini sering dipakai dalam *semantic web*, yang menggambarkan pendekatan dalam menangani dan menyimpan dokumen. Saat ini dokumen mulai disajikan dalam format XML (*Extensible Markup Language*). Bahasa lain yang digunakan dalam mendukung visi *Semantic Web* selain XML adalah RDF/OWL, Selanjutnya RDF/OWL dengan kemampuan dan fasilitas yang dimilikinya mampu digunakan untuk merepresentasikan makna istilah dalam *web* hingga dapat diproses dalam mesin.

Implementasi dengan pemanfaatan teknologi *semantic web* yang penulis lakukan yaitu pada perpustakaan dengan menampilkan bibliografi, baik untuk majalah, buku, jurnal, dan tesis. Pengelolaan buku, jurnal, tesis yang berfungsi sebagai literatur yang disusun secara sistematis ini sering dikenal dengan istilah bibliografi (Neil, 2004). Ada berbagai macam pelayanan yang disediakan didalam perpustakaan salah satunya yaitu pencarian untuk koleksi dokumen yang merupakan studi kasus pada penelitian ini.

## 1.2. Permasalahan

Bagaimana mengimplementasikan aplikasi pencarian menggunakan JSP (*Java Server Page*), Jena API dan SPARQL sebagai bahasa *query* RDF/OWL

## 1.3. Metode Penelitian

Dalam penelitian ini penulis menggunakan metode pemodelan *Spiral* (Pressman,1982), diantaranya meliputi:

### a. Studi Literatur

Studi literatur mengenai *semantic web*, RDF/OWL dan situs-situs web yang relevan, studi kasus pencarian pada beberapa situs-situs edukasi melalui *digital library* dan mempelajari jenis-jenis properti untuk selanjutnya menjadi pertimbangan dalam pengembangan ontologi bibliografi pada sistem perpustakaan.

### b. Perencanaan (*Planning*)

Menggambarkan mengenai sumber daya (*resource*), alokasi waktu (*timeline*) yang diperlukan untuk menyelesaikan pembuatan sistem dan informasi-informasi lainnya terkait dengan pengembangan sistem

### c. Risk Analysis

Pada bagian ini merupakan analisa terhadap kekurangan-kekurangan yang terjadi pada sistem dengan tujuan untuk memperoleh masukan dan perbaikan sistem kedepan

### d. Implementasi (*Engineering*)

Implementasi model ontologi dibangun dengan tool *protege\_3.4*, kemudian aplikasi pencariannya diimplementasikan menggunakan *Java Server Pages* sebagai antar muka, *Jena API 2.5.6* untuk menghubungkan bahasa *Java* dengan basis data RDF/OWL dan SPARQL sebagai bahasa *query* RDF/OWL.

### e. Construction and Release

Pengujian dilakukan untuk validasi apakah sistem yang dibangun sudah berjalan sesuai yang diharapkan dan memberikan hasil yang relevan.

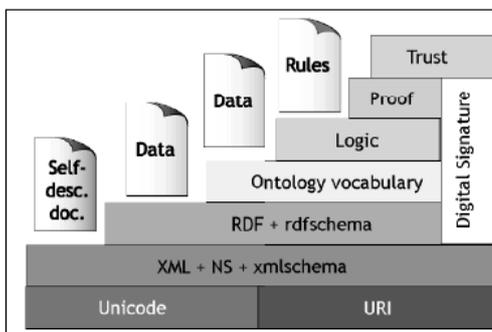
## 1.4. Pengertian *Semantic Web*

W3C (*World Wide Web Consortium*) memberikan suatu visi dari *semantic web* yaitu gagasan untuk memiliki data di *web* yang didefinisikan dan dihubungkan sedemikian rupa sehingga bisa digunakan oleh mesin, bukan hanya untuk ditampilkan saja tetapi juga untuk tujuan automasi, integrasi dan penggunaan kembali data antar berbagai aplikasi (W3C, 2001). Mesin disini mengacu pada komputer atau program komputer yang memiliki fungsi di *web*, misalnya seperti *software agents*, *softbot* dan aplikasi yang lain.

Sementara pencetus ide *web* pertama kali, Lee, dkk (2001) mengatakan bahwa dengan *semantic web* merupakan perluasan dari *web* saat ini, dimana informasi memiliki arti yang terdefinisi secara lebih baik dengan mengupayakan persamaan persepsi antara konsep-konsep yang ada, sehingga memungkinkan manusia dan komputer untuk bekerjasama secara lebih optimal.

## 1.5. Arsitektur *Semantic Web*

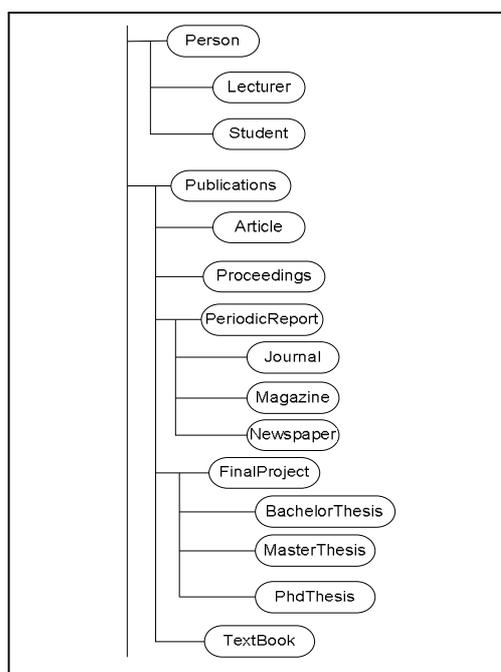
W3C (*World Wide Web Consortium*) merekomendasikan didalam *semantic web activity* terdapat beberapa *layer* arsitektur dari *semantic web* yang ditunjukkan oleh gambar 1. Paling bawah adalah *Unicode* dan *URI (Uniform Resource Identifiers)* yang memastikan penggunaan sekumpulan karakter yang telah disepakati secara internasional dan menyediakan alat untuk mengidentifikasi obyek di *semantic web*. Jenis *URI* yang paling terkenal adalah *URL (Uniform Resource Locator)* yang akan memberitahu komputer dimana letak suatu *resource* tertentu. Di atas *Unicode dan URI (Uniform Resource Identifier)* adalah *XML, namespace dan xmlschema*. *Layer* inilah yang mengintegrasikan definisi *Semantic web* dengan dokumen *XML* lain yang sesuai standar. *XML* merupakan format standar untuk dokumen terstruktur dan sebagai cara paling fleksibel untuk menciptakan standar bagi format informasi dan kemudian menyediakan format tersebut beserta datanya di *web*. *XML Schema* menggambarkan struktur dan batasan dari isi dokumen *XML* sedangkan *namespace* merupakan standar yang digunakan untuk menentukan label unik kepada sekumpulan nama elemen yang didefinisikan oleh *XML Schema*.



Gambar 1: Arsitektur *Semantic Web* (Lee, dkk, 2001)

### 1.6. Definisi Kelas dan Hirarki Kelas

Representasi dari hirarki kelas adalah mengelompokkan kelas-kelas dengan karakteristik yang sama. Suatu kelas dapat diturunkan dari kelas yang lain, dimana atribut dari kelas semula dapat diwariskan ke kelas yang baru. Pada aplikasi bibliografi ini dikembangkan dengan menggunakan pendekatan metode *top-down*, sebagai mana yang diungkapkan oleh Uschold dan Gruninger (1996) dalam makalah "*Ontologies: Principles, Methods and Applications*" bahwa ada beberapa pendekatan metode yang dapat digunakan untuk membangun struktur hirarki kelas diantaranya yaitu metode *top-down*, *bottom-up*, *combination*. Dengan metode *top-down*, kelas-kelas didefinisikan dari mulai konsep yang paling umum sampai konsep yang lebih spesifik. Dari metode *top-down* ini didapatkan struktur hirarki kelas untuk bibliografi yang dapat dilihat pada gambar 2



Gambar 2: Rancangan Struktur dan *Hirarki Kelas Bibliografi*

Barclay dan Savage (2004) menyatakan bahwa karakteristik dari pendekatan sistem berorientasi obyek pada dasarnya kumpulan dari obyek yang saling berinteraksi. Kumpulan obyek-obyek yang sama ini dinamakan sebagai Kelas (*Class*). Didalam ontologi bibliografi pada perpustakaan didefinisikan dua kelas utama yaitu kelas *Person* dan kelas *Publications* yang kemudian dapat diperinci lagi sesuai dengan kebutuhan.

1. *Person*, kelas ini merepresentasikan informasi tentang pencipta dari suatu bibliografi. Kelas *Person* dapat dibagi kedalam subkelas yang lebih spesifik dengan nama kelas *Lecturer* dan kelas *Student*.
2. *Publications*, kelas ini merepresentasikan jenis-jenis referensi yang ada pada umumnya. Kelas ini adalah *superkelas* dari kelas *Article*, *Proceedings*, *PeriodicReport*, *TextBook* dan kelas *FinalProject*. Didalam kelas *PeriodicReport*

dapat diuraikan lebih rinci kedalam subkelas *Journal*, *Magazine* dan *Newspaper*. Sementara untuk kelas *FinalProject* dapat diperinci menjadi beberapa subkelas dengan nama kelas *BachelorThesis*, kelas *MastersThesis* dan kelas *PhdThesis*.

### 1.7. Relasi Diagram Kelas (Class Diagram)

Suatu relasi adalah hubungan yang saling berkaitan antar kelas. Relasi mengijinkan sebuah kelas mengetahui atribut, operasi dan hubungan dengan kelas lainnya. Ada beberapa relasi didalam pendekatan desain berorientasi obyek salah satunya adalah relasi generalisasi. Generalisasi digunakan untuk memperlihatkan relasi pewarisan sifat antar kelas, sehingga pewarisan memperbolehkan suatu kelas untuk mewarisi semua atau sebagian atribut, operasi, relasi dan properti kelas lain. Diagram kelas atau *class diagram* menunjukkan interaksi antar kelas didalam sistem, sehingga dengan diagram kelas penyajian informasi yang dimiliki oleh setiap kelas dapat terlihat jelas. Kelas didefinisikan sebagai sebuah kumpulan dari obyek-obyek yang memiliki karakteristik yang sama dan menggambarkan perilaku dan atribut untuk semua obyek dari kelas (Garido, 2003).

### 1.8. Definisi Properti Kelas (Slot)

Setelah kelas diciptakan langkah selanjutnya adalah mendefinisikan properti kelas. Sebuah kelas jika berdiri sendiri tidak akan memberikan informasi yang cukup tanpa adanya properti yang melekat didalamnya. Dengan properti inilah sebuah kelas akan mempunyai nilai tambah dalam hal ini adalah informasi. Menurut Bernet (2001) bahwa properti dapat didefinisikan sebagai elemen-elemen data yang dimiliki oleh setiap kelas.

Untuk kelas *Lecturer* misalnya, properti yang dimiliki seperti *employeeId*, *email* dan *fullName*. Adapun properti-properti yang dimiliki kelas yang lainnya dapat disajikan seperti pada tabel 1 dan tabel 2.

Tabel 1: Properti Kelas *Lecturer*

Nama	Tipe	Kardinalitas	Keterangan
<i>employeeId</i>	<i>Int</i>	<i>Single</i>	Identitas dosen
<i>email</i>	<i>String</i>	<i>Multiple</i>	Email dosen
<i>fullName</i>	<i>String</i>	<i>Single</i>	Nama lengkap dosen

Kelas *Student* digunakan untuk menyimpan data-data mahasiswa. Tabel 2 menunjukkan beberapa properti yang dimiliki oleh kelas *Student* diantaranya, *studentNumber*, *year*, *email* dan *fullName*.

Tabel 2: Properti Kelas *Student*

Nama	Tipe	Kardinalitas	Keterangan
<i>studentNumber</i>	<i>Int</i>	<i>Single</i>	Identitas mahasiswa
<i>year</i>	<i>Int</i>	<i>Single</i>	Tahun masuk mahasiswa
<i>email</i>	<i>String</i>	<i>Multiple</i>	Email mahasiswa
<i>fullName</i>	<i>String</i>	<i>Single</i>	Nama lengkap mahasiswa

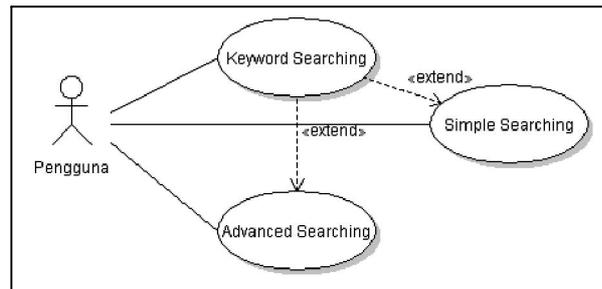
## 2. SIMULASI DAN ANALISIS

### 2.1. Gambaran Umum Aplikasi

Aplikasi yang dibangun dalam penelitian ini adalah aplikasi pencarian untuk koleksi perpustakaan yang merupakan salah satu bagian dari layanan sistem informasi dipergustakaan dengan menggunakan pendekatan *thesaurus*. Metode *thesaurus* adalah sebuah skema klasifikasi untuk istilah-istilah yang saling berkait membentuk struktur bahasa sehingga sebuah kata dapat dipahami dengan kata lainnya (Pendit, 2008). Perancangan aplikasi pencarian berbasis teknologi *semantic web* ini menggunakan pendekatan ontologi, dimana didalam ontologi tersebut digunakan sebagai penyimpanan dokumen bibliografi. Selanjutnya untuk menemukan informasi yang terdapat didalam ontologi dapat digunakan SPARQL sebagai bahasa *query* RDF/OWL. Jena API berfungsi sebagai *framework* yang menghubungkan antara ontologi dengan bahasa Java.

## 2.2. Diagram Use Case Pencarian Bibliografi

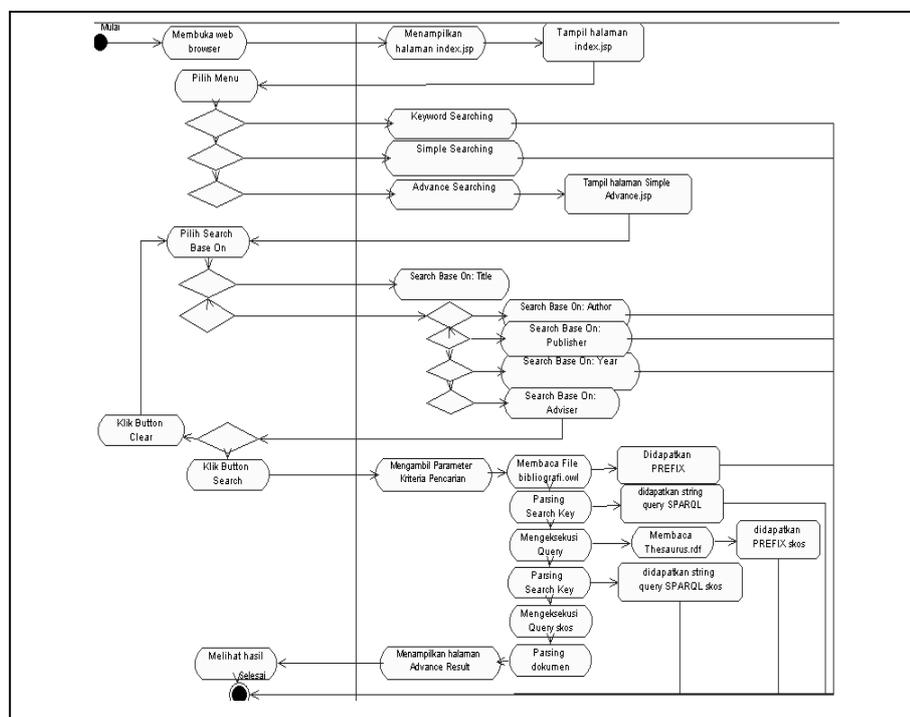
Use Case adalah deskripsi fungsi dari sebuah sistem yang mendefinisikan apa yang dilakukan oleh sistem dan elemen-elemennya. Bahwa pengguna dimungkinkan dapat memilih beberapa menu pencarian yang ada, seperti *keyword searching*, *simple searching* dan *advance serching*. Notasi “*extend*” memberikan alternatif apabila dengan *keyword searching* tidak memberikan hasil sehingga pengguna dapat memilih menu yang lain.



Gambar 3 : Diagram Use Case Pencarian Bibliografi

## 2.3. Aktifitas Advanced Searching

Pencarian selanjutnya adalah *advanced searching*, yang menyediakan lebih banyak kriteria pencarian yang dapat dipilih oleh pengguna, yaitu memilih kriteria berdasarkan penulis (*author*), penerbit (*publisher*), tahun (*year*) dan pembimbing (*adviser*).



Gambar 4 : Diagram Aktifitas Advanced Searching

Pencarian *advance searching* pada prinsipnya bertujuan menyediakan solusi dari kekurangan-kekurangan pencarian sebelumnya yang terdapat di *keyword searching* dan *simple searching*.

## 2.4. Perancangan Antarmuka Pengguna (GUI)

Antarmuka *user* atau lebih dikenal dengan GUI (*Graphical User Interface*) adalah bagian komponen dari sistem yang tidak dapat dipisahkan dengan *user*, karena dari GUI inilah *user* dapat berinteraksi dengan sistem dengan beberapa layanan yang telah disediakan. Rancangan antar muka bibliografi ini *user* dikategorikan menjadi tiga bagian, yaitu antarmuka untuk *keyword searching*, *simple searching* dan *advanced searching*. Antar muka *keyword searching* adalah antarmuka yang akan

ditampilkan ketika pengguna menjalankan aplikasi dan terlihat di *web browser*. Di halaman *web* inilah pengguna dapat melakukan pencarian melalui beberapa menu yang telah disediakan oleh sistem. Antarmuka *keyword searching* menawarkan pencarian berdasarkan judul (*title*).

Gambar 5: Rancangan Antarmuka *Advanced Searching* Aplikasi Bibliografi

Gambar 6 : Implementasi antar muka *Advanced Searching*

Fasilitas *advanced searching* menawarkan beragam pilihan, pada jenis pencarian *advanced searching* pengguna diminta memasukkan kata kunci berdasarkan beberapa jenis kriteria pencarian, kriteria tersebut diantaranya berdasar *title*, *publisher*, *author*, *year*, dan *adviser*. Kriteria judul (*title*) menegaskan bahwa pencarian didasarkan pada judul bibliografi yang ingin ditemukan selanjutnya diikuti dengan kata kunci yang di dasarkan pada penerbit (*publisher*), penulis (*author*), pembimbing (*adviser*) dan tahun (*year*). Hasil yang ditampilkan merupakan hasil semua judul yang berdasarkan kata kunci yang diberikan dan ditambah dengan *filter* kata kunci yang didasarkan pada menu “*search base on*”.

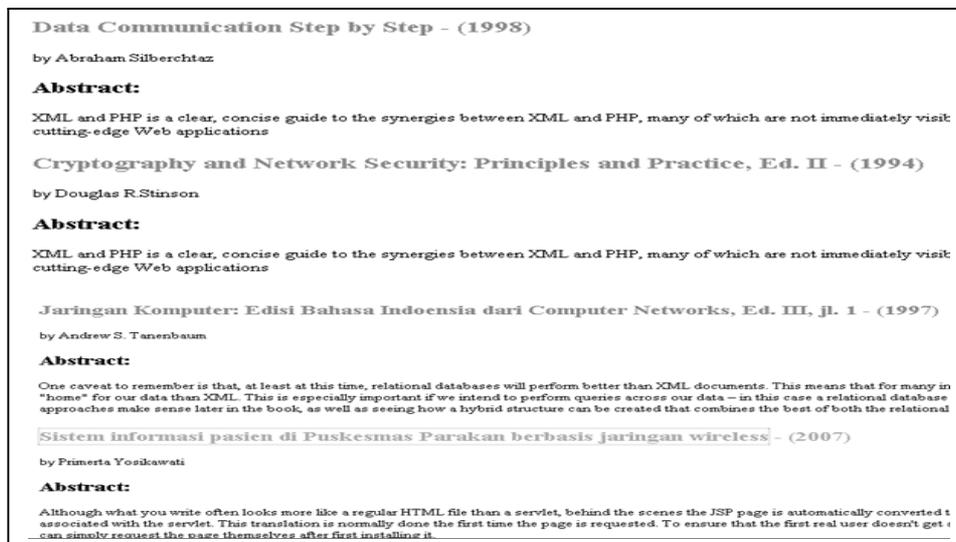
**Aplikasi Short Message Service dan J2ME untuk Sistem Pelelangan Hasil Komoditas Pangan - (2008)**  
by Rosihan  
Adviser: Dr Tech Khabib Mustofa SSI MKom

**Abstract:**  
Data mining usually means the methodologies and tools for the efficient new knowledge discovery from databases. In this paper, a genetic algorithms (GAs) based approach to proposed for extracting the decision rules including the predictors, the corresponding inequality and threshold values simultaneously so as to building a decision-making model accuracy. Early many studies of handling the breast cancer diagnostic problems used the statistical related techniques. As the diagnosis of breast cancer is highly nonlinear in a comprehensive model taking into account all the independent variables using conventional statistical approaches.

Gambar 7: Hasil untuk kata kunci “j2me” dan *author* “rosihan” pada *Advanced Searching*

## 2.5 Implementasi *Keyword Searching* pendekatan *Semantic*

Pada implementasi dengan menggunakan pendekatan *semantic*, pencarian yang dibangun menggunakan bantuan sebuah konsep. Konsep disini artinya menegaskan pendekatan yang digunakan untuk pencarian berdasarkan ‘makna (*semantic*)’, sehingga sesuatu yang diproses oleh komputer dapat dipahami berdasarkan makna bukan kata (*string*). Kata kunci yang dimasukkan (*keyword*) akan dipahami dengan kata yang lainnya yang memiliki makna yang sama. Seperti pada implementasi ini misalnya pengguna memasukkan kata kunci dengan kata “jaringan” maka informasi yang ditampilkan tidak hanya memuat kata atau *string* yang mengandung “jaringan” akan tetapi padanan kata yang lain seperti “*network*”, “*data communication*”, “*Local Area Network*” juga akan ditampilkan. Detail implementasi dengan menggunakan pendekatan *semantic* dapat disajikan melalui gambar 8.



Gambar 8: Hasil untuk kata kunci "jaringan" dengan Pendekatan *Semantic*

### 3. KESIMPULAN

Dari hasil implementasi yang diujicobakan dalam aplikasi bibliografi maka dapat disimpulkan:

1. Pendekatan ontologi adalah salah satu alternatif yang layak dipertimbangkan untuk penyimpanan basis data.
2. Pada dasarnya ide *semantic web* bukanlah hanya berbicara mengenai pertukaran data, akan tetapi perlu di lihat bagaimana data didistribusikan dan disatukan, SKOS(*Simple Knowledge Organization System*) adalah salah satu mekanisme data dapat didistribusikan yang didukung bahasa RDF sebagai suatu format bahasa untuk pendistribusian data dilingkungan *web*.
3. Dengan pendekatan SKOS, paradigma pencarian berbasis teks (*text-based search*) dapat digantikan oleh pencarian berbasis konsep (*concept-based search*).

### DAFTAR PUSTAKA

- [1] Baeza Y. R. dan Riberio N. B., 1999, Modern Information Retrieval, ACM Press, New York.
- [2] Isaac, A. dan Summers, E., 2009, SKOS Simple Knowledge Organization System Primer, Recommendation, World Wide Web Consortium (W3C) :  
<http://www.w3.org/TR/2009/WD-skos-primer-20090615/> available:  
<http://www.w3.org/TR/skos-primer>
- [3] Karsanti, H. T., 2006, Penerapan Teknologi Semantic Web pada Sistem Manajemen Perpustakaan Digital, Skripsi, Program Studi Ilmu Komputer FMIPA UGM, Yogyakarta.
- [4] Miles, A., Matthews, B., Beckett, D., Brickley, D., Wilson, M dan Rogers, N., 2005, SKOS : A Language to describe simple knowledge structures for the web, Recommendation, World Wide Web Consortium (W3C) :  
<http://www.w3.org/TR/swbp-skos:a language to describe simple knowledge structures for the web>
- [5] Miles, A. dan Bechhofer, S., 2009, SKOS Simple Knowledge Organization System Reference , Recommendation, World Wide Web Consortium (W3C):  
<http://www.w3.org/TR/2009/PR-skos-reference-20090615/>, Lastest version available :  
<http://www.w3.org/TR/skos-reference>
- [6] Manola, F. dan Miller, E., 2004, RDF Primer, Recommendation, World Wide Web Consortium(W3C) :  
<http://www.w3.org/TR/2004/REC-rdf-primer-20040210/> Lastest version available :  
<http://www.w3.org/TR/rdf-primer/>
- [7] Miles, A. dan Brickley, D., 2005, SKOS Core Guide, Recommendation, World Wide Web Consortium (W3C) :  
<http://www.w3.org/TR/2005/WD-swbp-skos-coreguide-20051102> , available  
<http://www.w3.org/TR/swbp-skos-core-guide>
- [8] Miles, A., 2005, Quick Guide to Publishing a Thesaurus on the Semantic Web, Recommendation, World Wide Web Consortium (W3C) :  
<http://www.w3.org/TR/2005/WD-swbp-thesaurus-pubguide-20050517>Lastestversionavailable:  
<http://www.w3.org/TR/swbp-thesaurus-pubguide>
- [9] Miles, A., Matthews, B., Beckett, D., Brickley, D., Wilson, M dan Rogers, N., 2005, SKOS : A Language to describe simple knowledge structures for the web, Recommendation, World Wide Web Consortium (W3C) :  
<http://www.w3.org/TR/swbp-skos:a language to describe simple knowledge structures for the web>