

PENGEMBANGAN PORTAL BERITA MENGUNAKAN TEKNOLOGI WEB SEMANTIC PADA BANTEN TV

AGUS IRAWAN

Jurusan Teknik Informatika
Fakultas Teknologi Informas, Universitas Serang Raya
Email : agus.irawan@unsera.ac.id

Abstrak - *Web semantic merupakan pengembangan teknologi web modern. Web semantic saat ini sudah didefinisikan dengan berbagai macam pengertian yang mengacu kepada kemampuan web tersebut, salah satunya adalah kemampuan untuk memahami tentang hal-hal yang di mengerti oleh manusia. Kebutuhan dalam menemukan informasi yang tepat telah mengantarkan pemikiran tim barneers lee tentang web masa depan, dimana web memiliki kemampuan untuk memahami segala hal yang dipahami oleh manusia tantang definisi suatu hal. Teknologi yang dimanfaatkan dalam semantic web adalah pengembangan dari potensi dari XML seperti RDF, XML dan OWL yang dapat dijadikan media penyimpanan data dan dapat di berhubungan dengan bahasa pemrograman dengan menggunakan bahasa SPARQL.*

Keywords : OWL, RDF, XML, SPARQL

I. PENDAHULUAN

Peranan informasi dalam media informasi sangat penting, informasi tidak hanya menyediakan fakta tapi juga merupakan suatu media dalam menambah pengetahuan bagi yang mengkonsumsinya. Sebuah berita mampu menyediakan informasi yang beragam yang diambil dari berbagai sumber berdasarkan waktu kejadian sehingga memiliki nilai informasi yang tinggi. Dengan adanya portal berita di harapkan dapat memudahkan dalam pencarian informasi berita. Untuk mengembangkan Portal berita menggunakan teknologi website.

Sistem pemberitaan pada banten tv masih menggunakan media penyiaran televisi. Untuk mendapatkan berita atau informasi masih harus mencari dan mengumpulkan informasi dari berbagai sumber, setelah itu berita yang terkumpul akan diseleksi berdasarkan sumber berita, sehingga memakan waktu yang lama dikarenakan sulitnya untuk menentukan sumber berita terpercaya. Selain itu tentu saja membutuhkan biaya operasional yang besar untuk mendapatkan kejadian atau fakta yang dapat di olah menjadi berita.

Web semantik merupakan web yang menyediakan informasi dari sumber-sumber yang luas di internet. Hal yang penting dari web semantik adalah bagaimana mesin dapat mengerti dan memahami kata-kata yang berasal dari sebuah dokumen web. Sehingga manusia dan mesin dapat mengerti dan memahami serta mempunyai kemampuan untuk mengakses informasi secara tepat. Hal ini akan mempermudah dalam pencarian berita pada sebuah dokumen web yang berbeda-beda .

II. LANDASAN TEORI

a. Web semantic

Web semantik merupakan pengembangan dari WWW dimana content yang ditampilkan tidak hanya dalam bahasa manusia yang umum tetapi juga dalam format yang dapat dibaca oleh mesin (nico ibrahim,2007:29). Web semantik memiliki informasi yang dimiliki oleh mesin yang memiliki kecerdasan buatan sehingga mampu menemukan dan mengintegrasikan informasi dengan mudah. Tujuan dari web semantik adalah mengatur informasi dan prosedur. Fundamental dalam pembangunan *semantic web* adalah kreasi dan *semantic metadata*.

Metadata Terdiri dari 2 bagian:

- 1) Penggambaran sebuah dokumen. Contohnya adalah halaman web atau bagian dari suatu paragraf .
- 2) Penggambaran entitas didalam suatu dokumen. Seseorang atau suatu objek.

Pada saat membangun sebuah aplikasi semantik sebenarnya ada dua variabel yang terbentuk secara bersamaan . yang pertama adalah web yang terdiri dari *protocol* komunikasi dan format web. Terdapat standar web semantik yang direkomendasikan oleh W3C seperti RDF, OWL, dan SPARQL. Variabel lainnya adalah semantik merepresentasikan makna dari suatu web.

b. Komponen web semantic :

1. XML dan XML Schema
2. RDF dan RDF Schema
3. OWL (*ontology web language*)

c. Ontologi

Ontologi adalah istilah yang dipinjam dari filosofi yang mengacu kepada ilmu yang menggambarkan jenis-jenis entitas didunia dan bagaimana mereka saling berhubungan. pengertian ontologi yaitu sebuah ontologi adalah struktur hirarki dari istilah untuk menjelaskan sebuah domain yang dapat digunakan sebagai landasan untuk sebuah *knowledge base* (nuriana ayunintias, 2009:23)

Pengertian lain mengemukakan ontologi adalah uraian formal yang menjelaskan tentang sebuah konsep dalam satu domain tertentu, properti dari setiap konsep yang menjelaskan bermacam-macam fitur dan atribut sebuah konsep. Dan batasan pada slots. Sebuah ontologi bersama dengan seperangkat *instance* dari *class* membentuk sebuah *knowledge base*

Keuntungan menggunakan ontology

Terdapat beberapa alasan menggunakan Ontologi.

Yaitu :

- a. *Menjelaskan suatu domain secara eksplisit Memberikan struktur hirarki dari istilah untuk menjelaskan sebuah domain dan bagaimana mereka berhubungan.*

- b. *Barbagai pemahaman dari informasi yang terstruktur. Beberapa website yang menyediakan informasi yang sama dengan struktur ontologi yang sama maka perangkat lunak dapat mengumpulkan informasi dari website yang berbeda tersebut untuk menjawab permintaan user atau sebagai data input untuk aplikasi lainnya.*
 - c. *Penggunaan ulang domain pengetahuan. Apabila ingin membangun ontologi yang luas dapat mengintegrasikan dengan beberapa ontologi yang sudah ada.*
2. **Komponen ontology**
Ontologi memiliki beberapa komponen untuk menjelaskan ontologi tersebut :
- A. *Instance atau individual digunakan untuk merepresentasikan elemen pada suatu domain*
 - B. *Class Merupakan titik pusat dari ontologi. Class menjelaskan sebuah konsep dalam suatu domain yang terdiri dari beberapa instance . class memiliki nama lainya yaitu concept, object dan categories*
 - C. *Properties atau slot. Terdiri atas object properties dan datatype property. Objek properties akan menghubungkan antara instance dengan instance lainnya, sedangkan datatype property akan menghubungkan instance dengan data type value seperti teks string atau number*
3. **Protégé**

PROTEGE merupan tool ontology dengan platform Open source untuk membangun domain model dan aplikasi knowledge based. Protege mengimplementasikan struktur pemodelan knowledge dan dapat memvisualisasikan hasil ontologi dalam berbagai format (nuriana ayunintias, 2009:23). Sebuah ontologi menggambarkan konsep dan hubungan yang penting dalam domain khusus, yang menyediakan kosakata dalam domain tersebut. Dalam beberapa tahun terakhir, ontologi telah diadopsi dalam bisnis atau komunitas ilmiah seperti scientific knowledge portal, manajemen informasi, integrasi informasi, e-commerce dan semantic web service.

pada protege terdapat 2 cara dalam pemodelan ontologi, yaitu :

A. **Protege frame Editor**

User dapat membangun ontologi dalam frame-based dengan kesepakatan open knowledge base connectivity Protocol (OKBC). Pada model ini, sebuah ontologi terdiri dari seperangkat class yang terorganisir pada suatu hirarki yang merepresentasikan sebuah domain, seperangkat slot yang berhubungan dengan class serta instance dari tiap class tersebut.

B. **Protégé OWL**

Protege OWL editor merupakan kelanjutan dari protege yang mendukung OWL. OWL merupakan pengembangan mutakhir standar bahasa ontologi yang di rekomendasikan oleh W3C untuk mempopulerkan semantic web vision. Protege-OWL Editor memungkinkan User untuk :

- a) Mengambil dan menyimpan OWL dan RDF ontologi.
- b) Mengubah dan memvisualisasikan class, properties, dan semantic web rule language (SWRL)

- c) Menjabarkan karakteristik class secara logis sebagai ekspresi OWL.
 - d) Mengeksekusi penalaran seperti description logic classifier.
 - e) Mengubah OWL individual untuk web semantik
- Protege-OWL berhubungan dengan jena dan mempunyai open source java API untuk pengembangan semantic web service. Jena merupakan framework berbasis java untuk mengkonstruksi aplikasi semantik web. *Framework* ini menyediakan lingkungan pemrograman *RDF, RDF Schema, OWL dan SPARQL*.

Protege-OWL Application Programing Interface (API) adalah sebuah java library open source untuk OWL dan RDF. API menyediakan metode untuk mengambil dan menyimpan file OWL, menanyakan OWL data model, menjalankan penalaran berdasarkan *Description Logic engines* sebagai *graphical user interface*. Api didesain untuk dapat digunakan dalam 2 konsep yakni :

- a) Pengembangan komponen yang dieksekusi dalam protege-OWL editor user interface.
- b) Pengembangan aplikasi stand-alone.

III. Cara Penelitian

1. **Alat dan bahan penelitian**

Tahapan pertama adalah pengumpulan data dengan memanfaatkan RSS dari sumber berita yang sudah di verifikasi. Beberapa portal berita yang dijadikan rujukan adalah detik, antara dan viva. Setiap portal berita tersebut menyediakan layanan RSS dalam format XML yang dapat diserap secara otomatis dan berkala memanfaatkan fitur server *cron* dan memanfaatkan web server yang dibuat menggunakan bahasa pemrograman PHP. Menggunakan komputer dengan prosesor amd turion 1,6 ghz, ram 1,5 GB dengan platform linux mint. Untuk perangkat lunak menggunakan apache sebagai web server, bahasa pemrograman PHP, protege untuk membangun ontologi, mysql sebagai database dan geany sebagai text editor.

IV. Analisa

1. **Kebutuhan data**

- a. *Resources*, merupakan metadata berita yang berhasil di kumpulkan oleh *domain expert* (admin) dan *aggregator* (web servis) kemudian di simpan ke dalam database.
- b. Admin, sebagai pengelola *resources* juga disebut domain expert
- c. Kriteria, merupakan kategori dari berita
- d. *Aggregator*, data yang diextract oleh *aggregator* yang nanti dikonversi dalam bentuk RDF yang sesuai dengan skema.
- e. *Keyword*, kata kunci pencarian

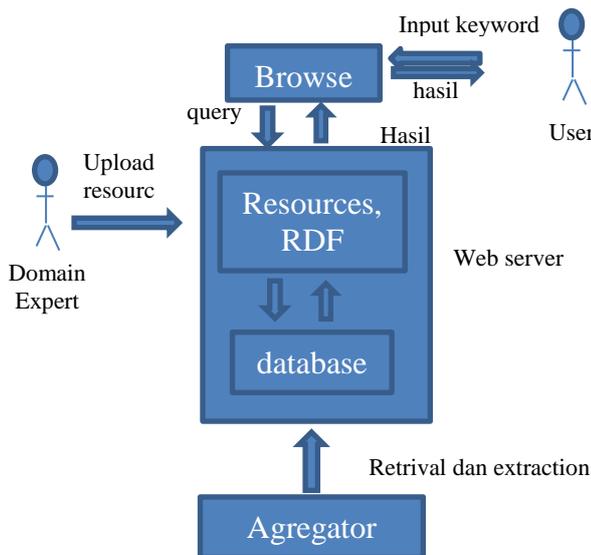
2. **Kebutuhan proses**

- a. Pencarian
- b. Login
- c. Pengelolaan *resources*
- d. *Retrieval, extraction* dan konversi metadata

V. Jalannya Penelitian

Setelah alat dan bahan tersedia, maka langkah selanjutnya adalah membuat perancangan system. Dalam perancangan di buat pemodelan dengan langkah-langkah

sebagai berikut : komponen untuk merancang ontologi, komponen interface, dan komponen portal.



Arsitektur rancangan *system web portal Semantic*

VI. HASIL DAN PEMBAHASAN

1. Membuat ontology domain berita

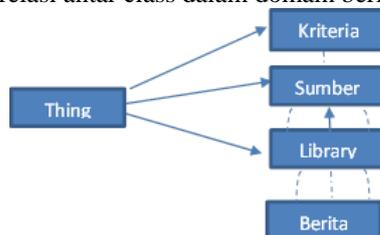
Untuk membuat ontology berita digunakan protégé versi 4.3. protégé merupakan software yang digunakan untuk membangun RDF, RDFS dan OWL, hasilnya merupakan file XML dengan ekstensi RDF, OWL, XML dan N3.

Langkah-langkah membangun ontology adalah sebagai berikut :

- a. Membuat Domain, domain di perlukan untuk identifikasi dimana setiap objek akan diidentifikasi melalui domain tersebut melalui IRI
- b. Membuat *class*, class merupakan representasi dari entitas, class yang dibuat antara lain, berita, kriteria, *resources* dan *library*.
- c. Membuat *object properties*, digunakan untuk menghubungkan antara class satu dan lainnya melalui properti. Object properties yang dibuat antara lain, hasKriteria yang menghubungkan antara class berita dan kriteria, hasSumber menghubungkan antara berita dengan sumber, dan isSumberOf merupakan kebalikan dari hasSumber.
- d. Data properties, digunakan untuk membangun relasi antara Object dengan data literal (Plain data). Data properties diinisiasi dengan kata yang bersifat predikat. Data properties yang dibuat antara lain, hasDate, hasSinopsis, hasTitle, hasUrl, namaSumber, homePageSumber, hasName, hasSinonim.
- e. Individual, individual merupakan *instance* dari sebuah class. Dalam sebuah class terdiri dari banyak property yang merupakan representasi dari class itu sendiri.

2. Bagan ontologi

Bagan relasi antar class dalam domain berita :



Gambar 2

Bagan ontologi domain berita

3. Retrieval dan extraction data

Untuk mengumpulkan resources berita, teknik yang digunakan adalah dengan mengambil informasi yang disediakan oleh website media pemberitaan dalam bentuk RSS (really simple syndication). RSS yang diambil merupakan data terbaru dari sebuah website media pemberitaan. File XML RSS diambil dengan menggunakan bahasa pemrograman PHP kemudian diextract dengan menggunakan tool XML Parse, hasil dari parse data tersebut kemudian ditulis dengan bahasa pemrograman PHP dalam format RDF/OWL yang sudah dibuat skemanya menggunakan tool protégé. Kemudian file tersebut diload dengan menggunakan fungsi dari *framework* ARC untuk dimasukan kedalam database.

Untuk membuat proses ini dapat bekerja secara simultan dan berkala digunakan *tool cron tab* pada hosting. *Crontab* merupakan *tool* untuk menjadwalkan sebuah *task* untuk dijalankan.

4. Pembahasan

Web semantic merupakan web yang menitikberatkan pemanfaatan metadata untuk menemukan informasi yang bersesuaian. Untuk menemukan informasi yang sesuai dengan kebutuhan pengguna maka ditanamkan pengetahuan pada aplikasi untuk mampu memahami makna tentang segala hal dan relevansinya dengan hal-hal lainnya. Kemampuan sebuah aplikasi untuk memahami makna “*thing*” sesuai dengan pemahaman manusia, merupakan sebuah langkah untuk menciptakan system atau aplikasi yang memiliki kecerdasan semantic dan merupakan aplikasi yang memiliki perbedaan dengan lainnya. Perbedaan yang jelas antara web semantic dengan web biasa adalah :

- a. Web semantic menitikberatkan pada metadata bukan pada data.
- b. Web semantic melakukan query akan tetapi bukan query relational database akan tetapi query untuk relational metadata dalam file XML.
- c. kata kunci pada web semantic tidak hanya mencocokkan kata tapi juga makna dari kata tersebut.

VII. PENUTUP

1. Kesimpulan

Pengembangan web semantic pada banten TV ini menggunakan framework ARC dengan bahasa pemrograman PHP, sedangkan untuk membangun ontologi digunakan tool protégé sehingga menghasilkan file xml dalam format OWL yang kemudian akan di query menggunakan SPARQL.

2. Saran

- a. Perlunya memperkaya sumber berita, agar informasi yang disajikan akan memperkaya pengetahuan bagi masyarakat yang menggunkannya.
- b. Web semantic merupakan teknologi baru, maka diharapkan para peneliti local dapat mengembangkan dan lebih melakukan eksplorasi agar dapat memberikan kontribusi terhadap arah

perkembangan teknologi web dimasa yang akan datang.

DAFTAR PUSTAKA

1. Ibrahim niko, (2007) pengembangan aplikasi semantic web untuk membangun web yang lebih cerdas, Universitas Kristen Maranatha, Bandung.
2. Nugroho Atmoko, (2009) Implementasi semantic web studi kasus penelusuran informasi sarana-sarana kesehatan di kabupaten bantul Yogyakarta, Universitas Semarang, Semarang
3. ARC homepage, diakses 15 september 2013, dari github, <https://github.com/Arc2/Wiki/>
4. Protégé homepage, diakses 20 agustus 2013, dari Stanford university, <http://protege.stanford.edu>
5. cerdas, Universitas Kristen Maranatha, Bandung
6. Wicaksono Aji Setyo, Novrido, dan Jayadianti Herlina, (2010) penerapan teknologi semantik web untuk menentukan jalur bis Trans Jogja, UPN "veteran" Yogyakarta, Yogyakarta.
7. Ayuningtyas Nuriana, (2009) Implementasi Ontologi web dan aplikasi semantic untuk system sitasi jurnal elektronik Indonesia, Universitas Indonesia, Jakarta.