

TRANSFORMASI XML MENJADI RESOURCE DESCRIPTION FRAMEWORK (RDF) BERBASISKAN TEKS DAN GRAFIS PADA SEMANTIC WEB (PADA KASUS DAFTAR BUKU PERPUSTAKAAN)

Yanes Arfian¹, Yanuar Firdaus A.w.², Arie Ardiyanti Suryani³

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Mesin pencari menjadi solusi untuk mendapatkan suatu informasi, tetapi hasil yang didapatkan masih belum sesuai dengan keinginan user. Hal ini disebabkan karena mesin belum bisa mengerti bahasa yang digunakan sebagai masukan dalam mencari informasi yakni bahasa yang sehari-hari digunakan manusia. Untuk menangani masalah ini, Semantic Web muncul untuk memperbaiki kekurangan mesin dalam mengerti informasi yang berbentuk sebuah kalimat. Dalam hal ini XML yang berfungsi untuk mendeskripsikan struktur suatu informasi, tidak bisa menjelaskan hubungan antar informasi yang dibentuk pada suatu kalimat.

Untuk itu, Semantic Web ini menggunakan Resource Description Framework (RDF) untuk menjelaskan hubungan antar informasi dengan menggunakan suatu konsep subject, predicate dan object. Penggunaan vocabulary pada sintaks RDF dapat menyederhanakan pemodelan graph terhadap informasi yang didefinisikan. Vocabulary tersebut dapat dibedakan menjadi Existing vocabulary dan User vocabulary. Pemodelan RDF graph dapat semakin sederhana jika nested tag yang digunakan semakin sedikit.

Kata Kunci : mesin pencari, Semantic web, XML, RDF, vocabulary, nested tag

Abstract

Search engine have become a solution to get some information, but the results obtained are still not in accordance with the wishes of the user. This is because the machine can not understand the language used as input in the search for information that used by humans everyday. To address this issue, the Semantic Web appears to correct deficiencies in the machine to understand the information in the form of a sentence. In this case the XML which is used to describe an information structure, can not explain the relationship between information that was formed in a sentence.

Therefore, it uses Semantic Web Resource Description Framework (RDF) to describe the relationship between the information by using a concept of subject, predicate and object. The use of vocabulary in RDF syntax graph can simplify the modeling of information that is defined. Vocabulary can be divided into Existing vocabulary and User vocabulary. RDF graph modeling can be more simple if the nested tags that are used less.

Keywords : search engine, Semantic web, XML, RDF, vocabulary, nested tag

1. Pendahuluan

1.1 Latar belakang masalah

Mesin pencari menjadi salah satu solusi untuk dapat mendapatkan informasi yang diinginkan secara tepat, tapi pada umumnya peristiwa yang terjadi adalah informasi yang disajikan oleh mesin pencari masih belum bisa memberikan hasil yang benar-benar sesuai dengan keinginan *user*. Dikutip dari Wikipedia, bahwa dalam perkembangannya teknologi *Web* mencapai pada tahap yang disebut dengan *Web 3.0*, dimana salah satu bagian dari *Web 3.0* ini adalah *Semantic Web*.

Semantic Web merupakan *Web* yang mampu untuk memberikan deskripsi yang dapat dimengerti oleh komputer mengenai suatu informasi[20]. Penemu dari World Wide Web dan juga pimpinan dari World Wide Web Consortium, Tim Berners Lee, mendefinisikan *Semantic Web* sebagai *Web* dari data yang bisa diproses langsung ataupun tidak langsung oleh mesin. Dari definisi tersebut maka *Semantic Web* termasuk ke dalam golongan *Web 3.0* dimana *Web 3.0* yang merupakan generasi terbaru dari perkembangan *Web 2.0* yang ditambahkan dengan teknologi *Semantic Web* dan Linked Data[6].

Semantic Web memperbaiki kekurangan teknologi sebelumnya dimana mesin tidak dapat mengerti informasi yang berbentuk dalam kalimat, misalnya “*Semantic Web* is a part of *Web 3.0*” atau “Jakarta is the capital of Indonesia”. Kalimat seperti itu tentu dapat langsung dimengerti oleh manusia, tetapi bagaimana kalimat tersebut dapat dimengerti oleh mesin? Karena hal itu maka *Semantic Web* mempunyai sebuah cara agar informasi yang berupa kalimat yang dapat dimengerti oleh manusia juga bisa dimengerti oleh mesin.

Semantic Web menjelaskan informasi dengan cara yang dapat dimengerti oleh mesin, yaitu dengan menjelaskan hubungan antara informasi tersebut seperti “A is a part of B”, “Y is a member of Z” dan properti dari informasi tersebut seperti ukuran, harga, umur, berat[20]. Penggunaan XML dibutuhkan pada saat ingin mendeskripsikan sebuah informasi, tetapi tidak dapat menangani bagaimana menjelaskan hubungan antar informasi tersebut. Pada search engine, kekurangan ini dapat menghasilkan hasil pencarian yang tidak sesuai dengan *keyword* yang dimasukan oleh user, sehingga user harus memilih lagi konten yang sesuai keinginannya dari hasil pencarian.

Oleh karena itu untuk menangani kekurangan ini maka *Semantic Web* menggunakan Resource Description Framework (RDF) untuk dapat menjelaskan hubungan antara informasi ini. RDF merupakan sebuah *framework* yang digunakan untuk merepresentasikan informasi yang berupa *metadata* untuk dapat dipertukarkan di *Web*. RDF sendiri ada yang berbasis teks dimana direpresentasikan dalam bentuk sintaks XML, oleh karena itu biasanya disebut dengan RDF/XML. Walaupun dibuat berdasarkan XML, terdapat perbedaan antara dokumen XML dengan dokumen RDF dimana dokumen RDF itu lebih kompleks daripada dokumen XML karena di dalam RDF itu terdapat sintaks yang

bisa dimengerti oleh mesin. Selain dalam bentuk sintaks, RDF juga dapat dimodelkan dalam bentuk yang berbasis grafis yang bisa dibuat dari sintaks RDF yang ada. RDF yang berbasis grafis dalam hal ini merupakan grafis pemodelan terhadap sintaks RDF seperti halnya ERD atau tabel relasi pembuatan database ataupun UML dalam pemodelan pada pemrograman Object Oriented.

Pada dasarnya RDF bertujuan untuk memberikan informasi terhadap sebuah informasi (*metadata*), dimana *metadata* tersebut digunakan oleh mesin pencari untuk melakukan tugasnya dalam menemukan informasi yang diinginkan oleh *user*. RDF mendukung bahasa yang dapat dimengerti oleh mesin sehingga pertukaran informasi dapat berjalan dengan lebih baik. RDF memperluas struktur link dari *Web* sebagai penamaan hubungan antara informasi yang ada di *Web*. Hubungan antar informasi tersebut dimuat ke dalam bentuk kalimat yang mempunyai pola *subject, predicate, object* yang disebut sebagai *statement* dimana di dalam RDF ketiga bagiannya dituliskan dalam bentuk sebuah *Web identifiers (URI)*. Konsep ini disebut dengan *Triple*.

Dengan adanya konsep ini mesin dapat lebih mengerti hubungan antar informasi yang ada di dalam *Web* dan memudahkan user untuk mendapatkan informasi yang sesuai dengan keinginan mereka, misalnya seseorang dapat menemukan informasi yang cukup berhubungan dengan *keyword* yang dia masukan ke dalam mesin pencari sehingga dia tidak perlu memilah lagi konten hasil dari pencarian tersebut.

Dengan melihat keadaan dimana beberapa informasi lebih banyak tersedia pada XML dibandingkan RDF, maka tugas akhir ini membahas bagaimana transformasi dokumen XML menjadi dokumen RDF sebagai usaha pembuatan *Semantic Web* sehingga informasi yang tersedia pada XML juga dapat tersedia pada RDF dan ketentuan apa saja yang perlu diperhatikan dalam proses transformasi yang diterapkan pada kasus daftar buku perpustakaan di IT Telkom. Alasan penulis mengambil kasus terhadap daftar buku perpustakaan ini adalah bahwa di dalam perpustakaan terdapat aplikasi seperti Search Engine yang dapat membantu mahasiswa dalam mencari judul buku yang diinginkan. Sementara itu Search Engine sangat berkaitan dengan permasalahan yang ada di dalam penelitian ini dimana RDF dapat membantu user untuk mendapatkan informasi yang tepat.

1.2 Perumusan masalah

Dalam pembahasan mengenai judul Tugas Akhir ini beberapa permasalahan yang akan diteliti dalam pengerjaan Tugas Akhir ini adalah sebagai berikut:

- a. Bagaimana membuat informasi berupa kalimat yang berisi hubungan antar suatu benda yang dapat dimengerti oleh mesin?
- b. Bagaimana menggambarkan hubungan antar informasi ke dalam sebuah bentuk graph?

- c. Apakah ketentuan yang perlu diperhatikan dalam melakukan transformasi XML menjadi sintaks dan graph RDF?

1.3 Tujuan

Tujuan yang ingin dicapai dalam penyelesaian Tugas Akhir ini diantaranya adalah:

- a. Melakukan Transformasi data dalam dokumen XML menjadi dokumen RDF dan mengetahui bentuk dokumen RDF yang berbasiskan teks (sintaks RDF/XML) hasil dari konversi dokumen XML.
- b. Mengetahui ketentuan apa saja yang perlu diperhatikan dalam transformasi dokumen XML menjadi dokumen RDF yang berbasiskan teks dan grafis.
- c. Menganalisa dan melakukan validasi dokumen RDF hasil transformasi dari dokumen XML.
- d. Membuat aplikasi yang dapat merubah sintaks XML menjadi sintaks RDF/XML.

1.4 Ruang Lingkup & Batasan Masalah

Adapun ada batasan-batasan yang terdapat pada permasalahan Tugas Akhir ini adalah:

- a. Data XML yang digunakan berupa data daftar buku di perpustakaan IT Telkom.
- b. Tool untuk memvalidasi sintaks XML bersifat offline, sedangkan untuk memvalidasi sintaks RDF menggunakan tool yang bersifat online.
- c. Sistem ini tidak menangani permasalahan yang menyangkut tipe data seperti *number* dari nilai suatu atribut, karena pada sistem ini semua nilai atribut diperlakukan sebagai tipe data *string*.

1.5 Metodologi Penyelesaian Masalah

Metode yang akan digunakan dalam penyelesaian masalah dalam pengerjaan Tugas Akhir ini adalah sebagai berikut:

- a. Studi Literatur
Melakukan pencarian terhadap berbagai referensi yang berhubungan dengan *Resource description Framework* untuk mendapatkan materi yang dapat membantu dalam pengerjaan Tugas Akhir ini. Referensi tersebut menjelaskan beberapa teori tentang RDF dan bentuk dari sintaks RDF. Salah satu paper referensi yang berjudul *XML to RDF Conversion: a Generic Approach*. Selain itu juga terdapat dokumentasi mengenai RDF yang dikeluarkan oleh W3.org yang mengeluarkan standar mengenai RDF.
- b. Pengumpulan data
Mengumpulkan data yang berbentuk XML yang nantinya akan digunakan untuk transformasi menjadi dokumen RDF berbasiskan teks dan

berbasiskan grafis yaitu sintaks dan pemodelan grafis. Data XML yang akan digunakan berasal dari data daftar buku perpustakaan IT Telkom.

c. Pengujian dan analisis

Melakukan pengujian terhadap aplikasi untuk merubah dokumen XML menjadi dokumen RDF dengan berbagai bentuk XML yang berbeda. Lalu memvalidasi hasil sintaks RDF hasil dari aplikasi tersebut, setelah itu melakukan analisis terhadap hasil yang akan didapatkan tersebut.

d. Kesimpulan

Dengan melihat data dari hasil pengujian maka dapat diambil kesimpulan mengenai permasalahan yang telah dibahas.

e. Penyusunan Laporan

Melakukan pendokumentasian terhadap proses-proses kerja yang dilakukan selama pengerjaan Tugas Akhir dan hasil yang didapatkan dari pengujian.



5. Kesimpulan dan Saran

5.1 Kesimpulan

Dari hasil analisis, didapatkan beberapa kesimpulan yang menjelaskan pembahasan dan pengujian yang telah dilakukan pada bagian sebelumnya, Kesimpulannya adalah sebagai berikut:

- a. Penggunaan user *vocabulary* mempunyai kelebihan yaitu dapat membuat pemodelan RDF graph menjadi lebih sederhana, tetapi mempunyai kelemahan berupa user *vocabulary* tidak selalu dapat langsung dimengerti oleh aplikasi-aplikasi yang dapat memproses RDF.
- b. Penggunaan existing *vocabulary* mempunyai kelebihan yaitu tidak perlu lagi untuk mendefinisikan *vocabulary* yang digunakan karena sudah dikenal terlebih dahulu secara umum, tetapi mempunyai kelemahan berupa pembuat RDF harus menyesuaikan struktur RDF sesuai dengan *vocabulary* yang sudah ada tersebut.
- c. Pada kasus ini, *vocabulary* yang cocok untuk digunakan merupakan user *vocabulary*, tetapi jika user *vocabulary* tersebut sama dengan *vocabulary* yang ada pada existing *vocabulary* maka lebih baik menggunakan existing *vocabulary* tersebut.
- d. *Nested tag* mempengaruhi jarak yang ditempuh untuk mencapai sebuah node. Kedalaman *nested tag* berbanding lurus dengan panjangnya jarak yang ditempuh sehingga semakin besar kedalaman *nested tag* maka jarak yang ditempuh juga semakin besar.
- e. Untuk kasus ini, struktur XML yang direkomendasikan merupakan XML yang mempunyai *nested tag* dengan jumlah yang lebih sedikit atau tidak mempunyai *nested tag* sama sekali.

5.2 Saran

Saran untuk penelitian selanjutnya mengenai RDF, diantaranya adalah:

- a. Perlu dilakukan perubahan dalam transformasi sintaks XML menjadi RDF/XML agar dapat melakukan transformasi terhadap sintaks XML yang dapat menerima berbagai macam tag XML.
- b. Mencari parameter lain yang dapat mempengaruhi bentuk sintaks RDF/XML dan juga bentuk RDF graph.
- c. Membuat aplikasi untuk transformasi XML ke RDF menjadi aplikasi dinamis yang dapat membaca XML yang memiliki berbagai bentuk struktur tag sehingga studi kasus menjadi lebih luas.

Daftar Pustaka

- [1] Amutiara. *Ontology: Bahasa dan Tools Protégé*. 2007
- [2] Carroll, Jeremy J. *Signing RDF Graphs*. Digital Media Systems Laboratory. 2003
- [3] Decker, Stefan. *The Semantic Web: The Roles of XML and RDF*. Stanford University.
- [4] Deursen, Davy Van. *XML to RDF Conversion*. Department of Electronics and Information Systems – Multimedia Lab.
- [5] Graves, Alvaro. *A Method To Rank Nodes In An RDF Graph*. Department of Computer Science Rensselaer Polytechnic Institute.
- [6] Hendler, Jim. *Web 3.0 Emerging*. Rensselaer Polytechnic Institute. 2009
- [7] Hua, Gu Yun. *Web Resource Description Model Bases on RDF*. Univesity of Information Science and Technology Nanjing, China.
- [8] Ibrahim, Niko. *Pengembangan Aplikasi Semantic Web Untuk Membangun Web Yang Lebih Cerdas*. Fakultas Teknologi Informasi, Universitas Kristen Maranatha.
- [9] Klapsing, Reinhold. *Semantic s in Web Engineering: Applying the Resource Description Framework*. University of Essen, Germany.
- [10] Lee, Young-Koo. *Transforming Valid XML Documents Into RDF Via RDF Schema*. Department of Computer Engineering. Kyung Hee University. Korea.
- [11] Nurkamid, Mukhamad. *Aplikasi Bibliografi Perpustakaan Berbasis Teknologi Semantic Web*. Fakultas Teknik Universitas Muria Kudus.
- [12] Permatasri, Hanum Putri. *Aplikasi Pencarian Karya Ilmiah Berbasis Semantic Web*. Universitas Gunadarma.
- [13] Prapanca, Aditya. *Desain Rules Untuk Transformasi Sintaks RDF (Text) ke RDF Berbasis Grafik Dalam Semantic Web*. FMIPA Universitas Negeri Surabaya.
- [14] RDF Primer diakses pada 18 Februari 2012 di <http://www.w3.org/TR/2004/REC-RDF-primer-20040210/>
- [15] RDF Tutorial diakses pada 21 November 2011 di <http://www.w3schools.com/RDF/default.asp>
- [16] *Semantic Web* diakses pada 23 November 2011 di http://id.wikipedia.org/wiki/Web_semantik
- [17] Siahaan, Daniel. *RDFGraph: New Data Modeling Tool for Semantic Web*. World Academy of Science, Engineering and Technology.
- [18] Suteja, Bernald Renaldy. *Ontology e-Learning Konten berbasis Web Semantic* . Fakultas Teknologi Informasi UK, Maranatha.
- [19] Tauberer, Joshua. *RDF:about*. Diakses pada 23 November 2011 di <http://RDFabout.com/>
- [20] XML Tutorial diakses pada 16 Februari 2012 di <http://www.w3schools.com/XML/default.asp>