

## **PENGEMBANGAN GARUDA (GARBA RUJUKAN DIGITAL) SEBAGAI SUMBER RUJUKAN KARYA ILMIAH DI INDONESIA**

**Rizal Fathoni Aji**

*Fakultas Ilmu Komputer Universitas Indonesia*

*Kampus UI Depok*

*Email: rizal@cs.ui.ac.id*

### **ABSTRAKS**

*Kondisi geografis Indonesia yang terpisah-pisah menyebabkan kemampuan untuk mengakses informasi dan ilmu pengetahuan di setiap daerah sangat bervariasi. Perpustakaan sebagai salah satu institusi penyedia informasi memiliki peran sangat penting dalam mewujudkan peningkatan akses terhadap informasi dan ilmu pengetahuan tersebut. Portal Garuda (Garba Rujukan Digital) merupakan salah satu model sistem yang dikembangkan untuk menyatukan informasi dari perpustakaan yang selama ini disekat-sekat oleh perbedaan institusi dan ego masing-masing. Dalam implementasinya, terdapat berbagai kendala yang ditemukan, namun kendala tersebut tidak menjadi penghambat dalam mewujudkan sistem ini.*

*Kata Kunci: Garba Rujukan Digital*

### **1. LATAR BELAKANG**

Perpustakaan, dengan fungsinya sebagai penyedia informasi memiliki peranan yang besar dalam pendidikan. Setiap pengguna perpustakaan diberikan kemudahan dan kebebasan dalam mencari informasi yang dibutuhkannya. Pemakai dapat menemukan beragam informasi yang dapat dipergunakan untuk lebih mengembangkan dirinya. Peran tersebut membuat perpustakaan menjadi salah satu elemen yang penting dalam penyelenggaraan pendidikan yang berkualitas.

Perpustakaan digital merupakan salah satu bentuk perpustakaan dimana informasi disimpan dalam bentuk elektronik sehingga dapat dipertukarkan dengan mudah melalui jaringan komputer. Pengguna tidak harus datang ke perpustakaan secara fisik, melainkan dapat melakukan pencarian dari tempat yang memiliki akses ke komputer yang tersambung ke perpustakaan tersebut, baik lewat internet maupun intranet. Setelah itu mereka dapat mencari koleksi yang ada di perpustakaan tersebut sesuai dengan kebutuhan. Koleksi elektronik ini mudah dipertukarkan, sehingga terdapat peluang yang dapat dimanfaatkan sebagai salah satu sarana untuk pemerataan pendidikan yang berkualitas. Pertukaran data antar perpustakaan dapat menjadi salah satu cara untuk mempersempit jurang-jurang informasi antar daerah di Indonesia. Terbukanya akses informasi ke daerah, secara tidak langsung dapat meningkatkan wawasan masyarakat di daerah tersebut.

Garuda (Garba Rujukan Digital) dikembangkan sebagai portal yang mengintegrasikan data karya ilmiah dari perpustakaan-perpustakaan di

Indonesia. Garuda berfungsi sebagai sebuah portal pencari yang dapat menghubungkan pengguna dengan sumber-sumber karya ilmiah yang ada di Indonesia. Jenis konten yang ada di dalam Garuda antara lain skripsi, tesis, disertasi, laporan penelitian, makalah jurnal dan berbagai sumber informasi karya ilmiah lainnya. Publik dan masyarakat umum yang melakukan penelusuran suatu karya ilmiah melalui Garuda dapat dengan segera mendapatkan informasi tentang keberadaan karya ilmiah yang dicari.

Namun, untuk mewujudkan hal tersebut bukanlah hal yang mudah. Banyak kendala yang ditemukan, terutama terkait dengan perbedaan standar yang digunakan dalam implementasi sistem perpustakaan, terutama dalam hal metadata dan protokol pertukaran data. Perpustakaan-perpustakaan yang berbeda dalam metadata dan juga protokol komunikasi akan sulit untuk saling bertukar data. Kendala lainnya adalah beragamnya spesifikasi komputer dan bandwidth jaringan yang dipakai dalam proses komunikasi tersebut. Keberagaman spesifikasi dan bandwidth jaringan tersebut dapat menyebabkan terjadinya bottleneck dalam proses komunikasi apabila kita salah dalam memilih perpustakaan digital yang akan diajak berkomunikasi.

Kendala yang bersifat non-teknis terutama terkait dengan ego dan kebebasan yang diinginkan oleh setiap pengelola perpustakaan. Memaksakan setiap perpustakaan untuk menggunakan platform dan standar yang sama bukanlah hal yang bijaksana dan sulit dilakukan karena mengharuskan mereka melakukan investasi ulang terhadap aplikasi yang mereka miliki termasuk proses migrasi data ke

aplikasi yang baru.

Untuk mengatasi permasalahan tersebut, tim pengembang Garuda menggunakan spesifikasi metadata dan protokol yang telah umum digunakan di perpustakaan. Untuk mengatasi masalah perbedaan implementasi metadata dan protokol, tim Garuda menggunakan sebagian arsitektur yang dirancang oleh (Kurniawan et al. 2006, Aji et al. 2006) pada penelitian sebelumnya. Dengan menggunakan arsitektur ini, setiap kontributor (pemilik data) dapat mengelola perpustakaan digital mereka tanpa harus merubah sistem yang sudah mereka miliki. Arsitektur ini digunakan dalam Garuda sebagai *tools* untuk menggabungkan dokumen-dokumen dari berbagai sumber menjadi satu kesatuan informasi.

## 2. METADATA

Secara umum, metadata mempunyai pengertian sebagai data yang berisi informasi tentang data lain (Caplan 2003). Kata metadata digunakan pertama kali dalam bidang ilmu komputer pada tahun 1998, namun dunia perpustakaan sebenarnya sudah menggunakan metadata dari sejak lama dalam bentuk kartu katalog (Caplan 2003). Selanjutnya kata metadata digunakan secara umum untuk mendeskripsikan sesuatu yang bentuknya tercetak, maupun berbentuk digital.

Dalam ilmu komputer, ada format standar yang dapat digunakan untuk mendeskripsikan sesuatu. Format yang umum digunakan, terutama dalam pertukaran data adalah XML (Website XML). XML atau *eXtensible Markup Language* merupakan format data yang sering digunakan dalam dunia *world wide web*. XML terdiri atas sekumpulan *tag* yang terdiri dari data. Satu set data dalam XML dimulai dengan *tag* pembuka dan diakhiri dengan *tag* penutup. Format XML diadopsi oleh standar metadata *Dublin Core* (Weibel 1998).

Metadata *Dublin Core* adalah cara terstandarisasi (NISO Standard Z39.85-2001) untuk mendeskripsikan suatu informasi. Metadata *Dublin Core* dapat digunakan untuk mendefinisikan dokumen fisik maupun digital. Ada dua level metadata *Dublin Core*, yang pertama adalah *simple* dan kedua adalah *qualified* yang menambahkan *encoding scheme*, enumerasi nilai dan informasi lainnya (Weibel 1998). Dalam format XML, metadata *Dublin Core* akan berbentuk seperti di bawah ini (*Dublin Core* 2007):

```
<dc:creator>Karl  
Mustermann</dc:creator>  
<dc:title>Algebra</dc:title>  
<dc:subject>mathematics</dc:subject>  
<dc:date>2000-01-23</dc:date>  
<dc:language>EN</dc:language>
```

```
<dc:description>An introduction to  
algebra</dc:description>
```

Selain *Dublin Core*, ada metadata yang umum digunakan oleh perpustakaan. Metadata tersebut adalah MARC (*Machine Readable Cataloging*) yang digunakan sebagai standar pertukaran data bibliografis dan data lain antar perpustakaan. MARC mendefinisikan format data yang muncul dari inisiatif *Library of Congress* sejak tahun 70-an (*Website LOC*, *Website MARC*). Format data MARC terdiri dari serangkaian *tag* dan *field* yang mendeskripsikan suatu koleksi.

## 3. PROTOKOL PERTUKARAN DATA

Protokol digunakan oleh lapisan aplikasi agar dapat menjalin komunikasi dengan aplikasi lain. Untuk meneruskan data, protokol menggunakan lapisan-lapisan di bawahnya, termasuk lapisan perangkat keras, agar antar protokol dapat saling berkomunikasi. Dengan cara ini, akan terjadi proses komunikasi secara *virtual* antara protokol satu dengan protokol yang lainnya.

Saat ini, *web service* merupakan protokol yang banyak digunakan untuk komunikasi data dalam internet (*Web Service @ W3C*). Menurut website tentang *web service* di W3C, yang disebut sebagai *web service* adalah suatu antar muka terprogram yang tersedia untuk memungkinkan terjadinya proses komunikasi antar aplikasi dengan media *world wide web*. *Web service* makin populer seiring dengan meningkatnya kebutuhan untuk melakukan komunikasi antar aplikasi di internet.

Pertukaran data dengan menggunakan *web service* bersifat interaktif. Untuk mengakses layanan dalam *web service*, *client* mengirim *request* ke sebuah alamat yang menyediakan *web service*, selanjutnya layanan tersebut akan memproses pesan yang diterima dan mengembalikan hasilnya. Semua pesan dalam *web service* menggunakan format XML dan mengalir menggunakan protokol FTP, HTTP maupun SMTP (Galbraith et al. 2002). Agar *client* dapat menggunakan *service* yang disediakan, digunakanlah WSDL (*Web Service Description Language*) sebagai cara agar *server* dapat memberikan informasi tentang layanan yang disediakan.

Protokol lain adalah OAI-PMH (*Open Archive Initiative Protocol for Metadata Harvesting*) (Lagoze dan Van de Sompel 2001). Sesuai dengan namanya, dalam protokol ini pengumpul data (*service provider*) mengumpulkan data dari *data provider* dalam interval waktu tertentu. Hasil yang didapatkan pengguna bukan merupakan hasil *real time* yang terdapat pada *data provider*.

OAI-PMH atau *Open Archive Initiative-Protocol for Metadata Harvesting* merupakan

protokol yang banyak digunakan untuk pertukaran data perpustakaan digital. Protokol ini berbasis teknologi standar seperti HTTP, XML dan menggunakan metadata *Dublin Core*. Dalam OAI-PMH, ada dua objek yang saling berinteraksi, yaitu *data provider* dan *service provider*. *Data Provider* bertugas untuk menyediakan data yang dimilikinya agar dapat diakses oleh *service provider*. Selanjutnya, *service provider* bertugas mengambil data yang ada pada *data provider* secara reguler dan menyimpan data ke dalam repositori lokal. Untuk mengambil data dari *data provider*, *service provider* menggunakan *verb-verb* atau perintah standar dalam OAI-PMH. *Verb* yang umum digunakan untuk pengambilan adalah *ListRecord* dan *GetRecord*. *ListRecord* digunakan untuk mengambil sekumpulan dokumen sesuai parameter yang diberikan, sementara *GetRecord* digunakan untuk mengambil dokumen tunggal sesuai dengan *identifier* yang diminta (Lagoze dan Van de Sompel 2001).

Berbeda dengan OAI-PMH, Z39.50 merupakan protokol yang bersifat interaktif. Z39.50 adalah protokol standar berbasis *client-server* yang memungkinkan komputer *client* untuk mencari dan mendapatkan informasi ke server data. Terdapat sembilan operasi yang dapat dilakukan dalam Z39.50, yaitu: Init, Search, Present, Delete, Scan, Sort, Resource-report, Extended-services dan Duplicate Detection. *Client* mengirim perintah operasi ke *server* dan hasil dari operasi tersebut akan dikirim ke *client* (NISO 2000). Implementasi Z39.50 banyak digunakan untuk bertukar data bibliografis bertipe MARC (Machine-Readable Catalogue). MARC umum digunakan sebagai metadata standar untuk pertukaran data bibliografis antar perpustakaan (Website MARC).

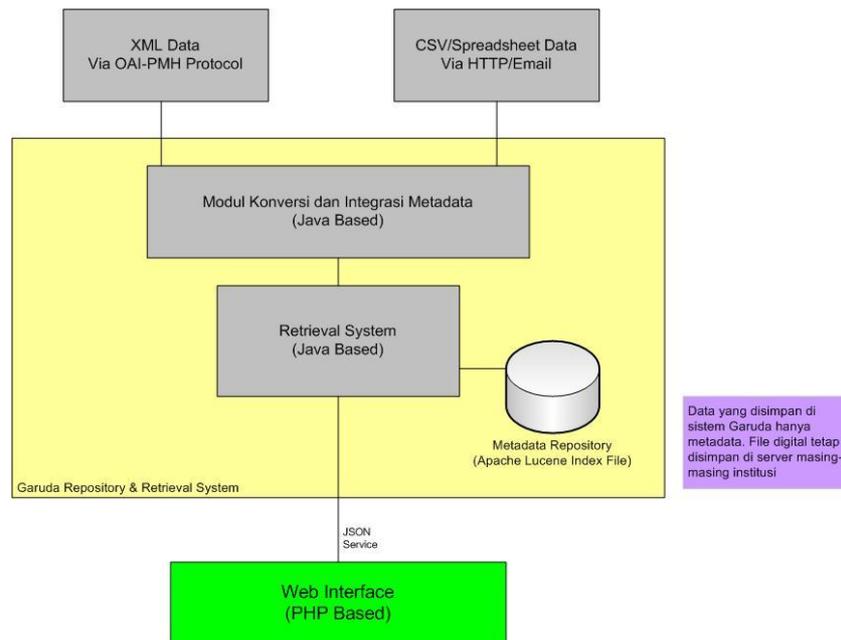
#### 4. PERANCANGAN ARSITEKTUR DAN IMPLEMENTASI

Dalam melakukan proses integrasi data dari para kontributor, beberapa permasalahan muncul mengingat adanya keberagaman *platform* dan aplikasi yang dipergunakan dalam perpustakaan

digital. Dalam penelitian-penelitian terdahulu, (Hidayanto 2002, Hidayanto et al. 2006, Aji dan Wibowo 2007) telah mengembangkan arsitektur untuk integrasi sistem temu kembali informasi dalam lingkungan terdistribusi yang bersifat homogen. Untuk lingkungan yang heterogen, penelitian yang sudah ada dilakukan antara lain oleh (Fuhr et al. 2002, Chernov et al. 2006, Kurniawan et al. 2006, Aji et al. 2006).

Menyeragamkan seluruh aplikasi dan platform perpustakaan di Indonesia bukanlah perkara yang mudah mengingat setiap perpustakaan sudah memiliki sistem perpustakaan digital tersendiri. Oleh sebab itu, yang dilakukan oleh tim Garuda adalah mengembangkan penelitian yang sudah dikerjakan sebelumnya khususnya terkait dengan masalah integrasi system perpustakaan digital dalam lingkungan yang bersifat heterogen. Metode yang digunakan dalam pengembangan arsitektur aplikasi ini mengikuti model yang digunakan oleh (Huang dan Hsu 2005, Kurniawan et al. 2006, Aji et al. 2006) dengan menempatkan sebuah *gateway* sebagai perwakilan dari tiap protokol yang ada. *Gateway* ini bertugas untuk mengumpulkan data dari *client-client* di bawahnya dan menyediakan sebuah *service* agar pengguna yang ingin mencari data di protokol yang bersangkutan, hanya perlu memanggil *service* yang disediakan oleh *gateway* tersebut.

Garuda menggunakan dua buah *gateway*. Satu *gateway* sebagai penghubung kontributor yang menggunakan protokol OAI-PMH, dan satu *gateway* lagi digunakan bagi kontributor yang menggunakan CSV (*Comma Separated Value*). Format CSV digunakan oleh Garuda karena masih banyak perpustakaan dan sumber informasi di Indonesia yang belum mengimplementasikan protokol OAI-PMH, sehingga mereka lebih mudah jika menyediakan data dalam bentuk CSV. Setiap *gateway* akan menjadi pusat informasi untuk melakukan transformasi data yang diterima kedalam bentuk standar metadata yang digunakan oleh Garuda. Bentuk arsitektur yang digunakan terlihat pada gambar berikut.



Gambar 1. Arsitektur Aplikasi Pertukaran Data

*Gateway* dirancang menggunakan bahasa pemrograman java dan dirancang secara modular sehingga memungkinkan penambahan jenis protokol pertukaran data yang lain. Gateway yang digunakan dalam Garuda bertipe *gateway* pengumpul. *Gateway* jenis ini mengambil data secara berkala dari penyedia data dan menyimpannya sehingga untuk melakukan pencarian terhadap data, peminta data tidak langsung mencari ke penyedia data, tetapi melakukan pencarian ke tempat penyimpanan lokalnya.

Keseluruhan data yang dikumpulkan disatukan dan di *index* dalam sebuah *retrieval system*. Selanjutnya, dari *retrieval system* ini dibuatlah *service* yang menyediakan fungsi-fungsi yang dibutuhkan oleh *web interface* untuk menampilkan hasil pencarian.

## 5. KESIMPULAN DAN PENUTUP

Model arsitektur yang dibuat digunakan untuk menggabungkan data-data perpustakaan dari jenis sumber informasi. Sistem pertukaran data ini telah diimplementasikan dan dapat dilihat di situs garuda (<http://garuda.depdiknas.go.id>). Hasil implementasi memperlihatkan kemajuan yang baik yang ditandai oleh bertambahnya jumlah kontributor dan jumlah koleksi yang ada di Garuda, serta banyaknya hits yang dicatat oleh server.

Dari keseluruhan data yang berhasil dikumpulkan, terlihat beberapa perbedaan yang sulit untuk disatukan. Perbedaan tersebut antara lain, perbedaan *subject heading* yang digunakan, serta format penulisan judul dan pengarang. Namun, seiring berjalannya waktu dan berkembangnya teknologi, di masa depan kita dapat

melihat perbedaan-perbedaan yang ada tersebut dapat disatukan. Untuk pengembangan selanjutnya, dapat ditambahkan *gateway-gateway* baru yang dapat digunakan untuk menghubungkan perpustakaan dan sumber-sumber informasi lain, seperti museum, pusat arsip dan pusat kebudayaan.

## PUSTAKA

- Aji, Rizal Fathoni., Heri Kurniawan, Zainal A. Hasibuan. 2006. *Adaptive System For Library Integration*, Yogyakarta: Proceeding of IIWAS (International Conference on Information Integration and Web-based Application & Services)
- Aji, Rizal Fathoni., Wahyu C. Wibowo. 2007. *Arsitektur Pertukaran Data Perpustakaan di Indonesia*, Depok: Proceeding of NACSIT (National Conference on Computer Science and Information Technology)
- Arms, William. 2000. *Digital Libraries*. MIT Press
- Callan, James, et al. 1998. *Searching Distributed Collections With Inference Network*. Proceedings of the 18<sup>th</sup> Annual International ACM SIGIR Conference
- Caplan, Priscilla. 2003. *Metadata Fundamentals for All Librarians*, American Library Association
- Chernov, S., C. Kohlschutter dan W. Nejd. 2006. *A Plugin Architecture Enabling Federated Search for Digital Library*. Japan: Proceedings of International Conference on Asian Digital Libraries
- Dublin Core. *Dublin Core*. <http://dublincore.org>
- Fuhr, N., P. Klas, A. Schaefer, dan P. Mutschke. 2002. *An Integrated Desktop for Supporting High-Level Search Activities in Federated*

- Digital Libraries*. Proceeding of the 6<sup>th</sup> European Conference on Research and Advanced Technology for Digital Libraries
- Galbraith, Ben. et al. 2002. *Professional Web Services Security*. Birmingham: Wrox Press Ltd.
- Hidayanto, A.N. 2002. *Pengembangan Sistem Temu Kembali Berbahasa Indonesia*. Tesis Magister Ilmu Komputer, Fakultas Ilmu Komputer Universitas Indonesia
- Hidayanto, A.N., Harry B. Santoso, Rizal F. Aji dan S. Bressan. 2006. *Community access point in Indonesia: Improving access to quality information and promoting local potentials*. Bangkok: Proceeding of 5<sup>th</sup> International Conference of E-Business
- Huang, Chung-Min dan Tz-Heng Hsu. 2005. *A Resource Exchange Architecture for Peer-to-Peer File Sharing Applications*. The Computer Journal Vol. 48 No.1, British Computer Society
- Lagoze, C., H. Van de Sompel. 2001. *The open archives initiative: building a low-barrier interoperability framework*, ACM/IEEE Joint Conference on Digital Libraries
- Kurniawan, H., Rizal F. Aji, Zainal A. Hasibuan. 2006. *Information Resource Sharing Based On Multi-Platform Library Systems in Indonesian Higher Education Institutions Network*, Singapore: Proceeding of International Conference on Educational Technology (ICET)
- Kurniawan, H., Rizal F. Aji. 2006. *Otomatisasi Pengelompokan Koleksi Perpustakaan dengan Pengukuran Cosine Similarity dan Edit distance*. Yogyakarta: Prosiding Seminar Nasional dan Aplikasi Teknologi Informasi. Yogyakarta
- NISO (National Information Standards Organization). 2000. *Information Retrieval (Z39.50): Application Services Definition and Protocol Specification (ANSI/NISO 239.50-2003)*. Maryland: NISO Press
- Pengembang Lontar. 2006. *LONTAR Developer Manual*
- Web Service @ W3C. *Web Service @ W3C*. <http://www.w3.org/2002/ws/>