

Basis Data Internet

Rengga Asmara, S.Kom. OCA

Pokok Bahasan

- World Wide Web
- Pendahuluan HTML
- Basis Data dan Web
- Arsitektur Application Server dan Server-Side Java
- Pendahuluan XML
- XML : DTDs
- Spesifikasi Domain pada DTDs
- Melakukan Query Data XML (XML-QL)
- Model Data Semi-terstruktur

Tujuan Belajar

- Setelah mempelajari materi dalam bab ini, mahasiswa diharapkan mampu:
 - Memahami konsep world wide web dan html pada database berbasis web
 - Memahami arsitektur pada database web
 - Memahami XML
 - Memahami dan mengimplementasikan model data semi-terstruktur

World Wide Web

- Web memungkinkan mengakses suatu file dimanapun pada internet.
- Suatu file diidentifikasi dengan *universal resource locator* (URL), contohnya :
 - <http://www.eepis-its.edu/~database/index.html>
- URL diatas mengidentifikasi suatu file bernama index.html, disimpan pada direktory ~database pada mesin www.eepis-its.edu.
- File ini adalah suatu dokumen yang mempunyai format *Hypertext Markup Language* (HTML) dan berisi beberapa link ke file lain (yang diidentifikasi melalui URL-nya).

World Wide Web – cont'd

- Perintah diterjemahkan oleh *Web browser* seperti Microsoft Internet Explorer, Netscape Navigator atau Opera untuk menampilkan dokumen secara atraktif, dan pemakai kemudian dapat menavigasi ke dokumen yang berhubungan dengan memilih link.
- Suatu koleksi beberapa dokumen disebut *web site* dan diatur menggunakan program yang disebut *web server*, yang menerima URL dan menampilkan dokumen yang diperlukan.
- Beberapa organisasi saat ini menangani *web site*.
- *World Wide Web* atau *Web*, adalah kumpulan *web site* yang dapat diakses di internet.

World Wide Web – cont'd

- HTML menghubungkan isi URL, yang mengidentifikasi tempat yang berisi file yang berhubungan.
- Jika seorang pemakai meng-click pada suatu link, Web browser menghubungkan ke Web server tujuan menggunakan protokol koneksi HTTP dan mengirim link URL.
- Jika browser menerima file dari web server, akan diperiksa tipe file dengan melihat extension dari nama file.
- Web browser menampilkan file berdasarkan tipe file, dan jika perlu memanggil program aplikasi untuk menangani file.

World Wide Web – cont'd

- Sebagai contoh, sebuah file yang diakhiri .txt merupakan file teks tak berformat, web browser menampilkan dengan menterjemahkan karakter ASCII dalam file.
- Struktur dokumen yang lebih tepat dapat dikodekan dalam HTML, yang menjadi cara standart struktur web page.
- Contoh lain, file .doc merupakan dokumen Microsoft Word dan Web browser menampilkan file dengan menggunakan Microsoft Word.

Hipertext Markup Language (HTML)

- *Hipertext Markup Language* (HTML) merupakan bahasa pemrograman berbentuk skrip yang sederhana.
- Teks diawali dan diakhiri dengan perintah yang disebut tags, biasanya terdiri dari start tag dan end tag.

HTML – cont'd

- Contoh skrip HTML untuk melihat daftar buku :

```
<HTML><BODY>
Fiction:
<UL><LI>Author: Milan Kundera</LI>
      <LI>Title: Identity</LI>
      <LI>Published: 1998</LI>
</UL>
Science:
<UL><LI>Author: Richard Feynman</LI>
      <LI>Title: The Character of Physical Law</LI>
      <LI>Hardcover</LI>
</UL></BODY></HTML>
```

Basis Data dan Web

- Halaman web berisi hasil dari query database.
- Untuk membangkitkan halaman web, dilakukan hal-hal sebagai berikut :
 - Web server membuat proses baru dari program yang berinteraksi dengan basis data.
 - Web server berkomunikasi dengan program tersebut melalui CGI (*Common gateway interface*).
 - Program membangkitkan halaman hasil dengan isinya berasal dari basis data.
 - Protokol lain selain CGI adalah ISAPI (*Microsoft Internet Server API*) atau NSAPI (*Netscape Server API*).

Arsitektur Application Server dan Server-Side Java

- Pada CGI, setiap halaman merupakan hasil pembuatan sebuah proses baru, sehingga sangat tidak efisien.
- Oleh karena itu, dikembangkan arsitektur *application server*, yaitu suatu perangkat lunak antara web server dan aplikasi.

Arsitektur Application Server dan Server-Side Java – cont'd

- Fungsi dari *application server* adalah
 - Mengendalikan thread untuk performansi
 - Tempat koneksi basis data (dan menggunakan kembali koneksi yang tersedia)
 - Integrasi antara sumber data yang heterogen
 - Manajemen transaksi melibatkan beberapa sumber data
 - Manajemen session

Arsitektur Application Server dan Server-Side Java – cont'd

- Pemrosesan pada sisi server dapat menggunakan :
 - Java Servlet : yaitu program java yang berjalan pada server dan berinteraksi dengan server melalui API yang didefinisikan dengan baik (*well-defined*)
 - JavaBeans : yaitu komponen perangkat lunak yang dapat digunakan kembali (*reusable*) yang ditulis dalam bahasa pemrograman Java
 - Java Server Page dan Active Server Page : yaitu kode yang berada di dalam halaman web yang diterjemahkan oleh web server.

Pendahuluan XML

- *Extensible HTML* (XML) merupakan bahasa skrip versi lanjutan dari HTML.
- XML merupakan penggabungan antara SGML dan HTML, dimana kekuatan SGML digabungkan dengan kesederhanaan HTML.
- XML menggunakan definisi *markup language* baru yang disebut *document type declaration* (DTDs).

Pendahuluan XML – cont'd

- Elemen yang terdapat pada XML meliputi :
 - blok utama yang dibangun secara terstruktur dengan XML
 - tag mulai (*start tag*) dan akhir (*end tag*)
 - harus secara tepat berulang

Pendahuluan XML – cont'd

- Elemen pada XML dapat mempunyai atribut yang merupakan informasi tambahan mengenai elemen tersebut.
- Entiti pada XML serupa dengan makro yang berupa teks biasa.
- Pada XML juga tersedia perintah untuk komentar dan elemen yang terpenting adalah *document type declaration* (DTDs).

Pendahuluan XML – cont'd

- Contoh skrip XML untuk melihat daftar buku adalah sebagai berikut :

```
<?XML version="1.0" standalone="yes"?>
<!DOCTYPE BOOKLIST SYSTEM "booklist.dtd">
<BOOKLIST>
<BOOK genre="Fiction">
  <AUTHOR>
    <FIRST>Milan</FIRST><LAST>Kundera</LAST>
  </AUTHOR>
  <TITLE>Identity</TITLE>
  <PUBLISHED>1998</PUBLISHED>
<BOOK genre="Science" format="Hardcover">
  <AUTHOR>
    <FIRST>Richard</FIRST><LAST>Feynman</LAST>
  </AUTHOR>
  <TITLE>The Character of Physical Law</TITLE>
</BOOK></BOOKLIST>
```

XML : DTDs

- Sebuah *document type declaration* (DTDs) adalah himpunan aturan yang mendefinisikan elemen, atribut dan entiti yang diperbolehkan pada dokumen.
- Dokumen pada XML mempunyai bentuk yang baik jika tidak mempunyai asosiasi DTD, tetapi digunakan berulang dengan tepat.
- Dokumen XML valid jika mempunyai DTD dan dokumen mengikuti aturan dalam DTD.

XML : DTDs – cont'd

- Contoh DTDs untuk melihat daftar buku adalah sebagai berikut :

```
<!DOCTYPE BOOKLIST [  
  <!ELEMENT BOOKLIST (BOOK)*>  
  <!ELEMENT BOOK (AUTHOR, TITLE, PUBLISHED?)>  
  <!ELEMENT AUTHOR (FIRST, LAST)>  
    <!ELEMENT FIRST (#PCDATA)>  
    <!ELEMENT LAST (#PCDATA)>  
  <!ELEMENT TITLE (#PCDATA)>  
  <!ELEMENT PUBLISHED (#PCDATA)>  
  <!ATTLIST BOOK genre (Science|Fiction) #REQUIRED>  
  <!ATTLIST BOOK format (Paperback|Hardcover) "Paperback">  
>
```

Spesifikasi Domain pada DTDs(1)

- Untuk menangani sumber-sumber data yang berlainan, dikembangkan standarisasi DTDs untuk memungkinkan domain dapat menukar data diantara sumber-sumber yang heterogen.
- Contohnya adalah domain pada DTDs untuk mengkodekan material matematika pada web dengan menggunakan *Mathematical Markup Language* (MathML).

Spesifikasi Domain pada DTDs(2)

- Perbedaan HTML dengan MathML dapat dilihat pada persamaan matematika di bawah ini ;

– Dalam HTML :

```
<IMG SRC="xysq.gif" ALT="(x+y)^2">
```

– Dalam MathML :

```
<apply> <power/>  
  <apply> <plus/> <ci>x</ci> <ci>y</ci> </apply>  
  <cn>2</cn>  
</apply>
```

Melakukan Query Data XML (XML-QL)

- Tujuan melakukan query data XML adalah melakukan manipulasi dokumen XML dengan menggunakan bahasa deklaratif tingkat tinggi.
- Query data XML belum dilakukan standarisasi.

XML-QL – cont'd

- Contoh query dengan XML-QL :

```
WHERE
  <BOOK>
    <NAME><LAST>$1</LAST></NAME>
  </BOOK> in "www.booklist.com/books.xml"
CONSTRUCT <RESULT> $1 </RESULT>
```

XML-QL – cont'd

- Contoh query dengan XML-QL yang lebih kompleks :

```
WHERE <BOOK> $b <BOOK> IN "www.booklist.com/books.xml",
  <AUTHOR> $n </AUTHOR>
  <PUBLISHED> $p </PUBLISHED> in $e
CONSTRUCT
  <RESULT>
    <PUBLISHED> $p </PUBLISHED>
      WHERE <LAST> $l </LAST> IN $n
      CONSTRUCT <LAST> $l </LAST>
  </RESULT>
```

Model Data Semi Terstruktur

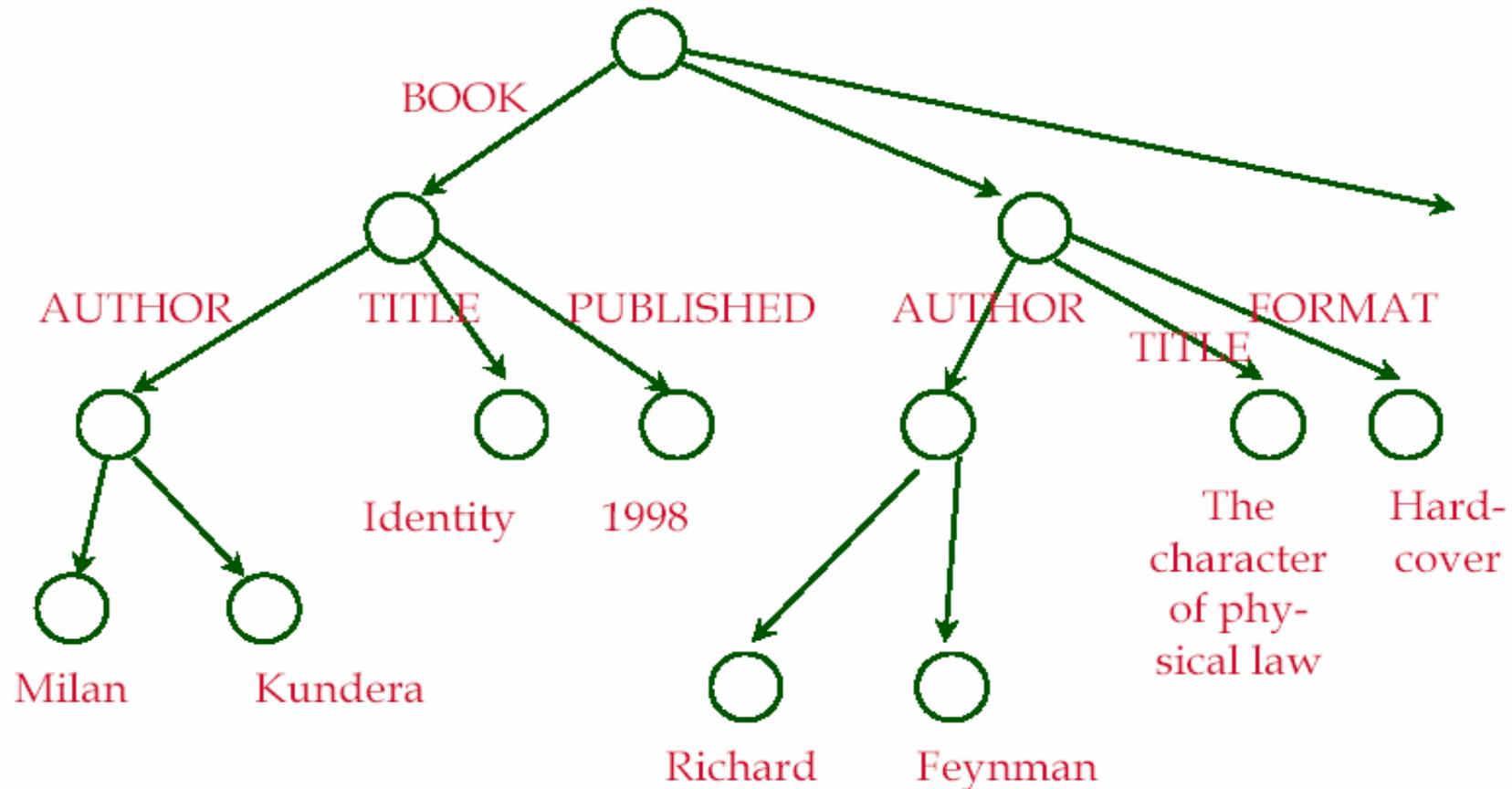
- Data semi terstruktur adalah data dengan struktur parsial.
- Semua model data untuk data semi terstruktur menggunakan beberapa tipe graph berlabel.
- Pada sub bab ini diperkenalkan model data semi terstruktur yang disebut *object exchange model* (OEM).

Model Data Semi Terstruktur – cont'd

- OEM mempunyai karakteristik sebagai berikut :
 - Obyek berbentuk triple (label, tipe, nilai).
 - Obyek yang kompleks di dekomposisi secara hirarki ke dalam obyek yang lebih kecil.

Model Data Semi Terstruktur – cont'd

- Contoh model data daftar buku menggunakan OEM dapat dilihat pada Gambar 8-1



Gambar 8-1: Model data daftar buku dengan menggunakan OEM

Ringkasan

- File pada World Wide Web diidentifikasi melalui universal resource locator (URL).
- Web browser membawa URL ke tempat yang berisi file dan bertanya pada Web server pada tempat tersebut untuk file yang dimaksud.
- Jika kemudian menampilkan file yang tepat, membawa ke tipe file dan instruksi terformat.
- Browser memanggil program aplikasi untuk menangani tipe file tersebut, misalnya memanggil Microsoft Word untuk menangani dokumen Word.

Ringkasan – cont'd

- HTML adalah markup language yang sederhana yang digunakan untuk menggambarkan dokumen.
- Program Java, audio dan video dapat disertakan dalam dokumen HTML.
- Data yang diakses melalui Web disimpan dalam DBMS.
- Web server dapat mengakses data dalam DBMS untuk membentuk halaman yang diminta oleh Web browser.

Ringkasan – cont'd

- Web server terkadang harus mengeksekusi program pada tempat secara berurutan untuk memenuhi permintaan dari Web browser.
 - Contohnya, kita dapat mengakses data dalam DBMS.
- Terdapat dua cara untuk Web server untuk mengeksekusi program : membuat proses baru dan mengkomunikasikan menggunakan protokol CGI, atau membuat thread baru untuk Java Servlet.
- Pendekatan kedua menghindari banyaknya pembentukan proses baru untuk setiap permintaan.

Ringkasan – cont'd

- Application server mengatur beberapa thread dan menyediakan fungsi lain untuk memfasilitasi eksekusi program pada Web server.
- Fungsi tambahan termasuk didalamnya keamanan, manajemen sesi, koordinasi akses ke multipel sumber data.
- JavaBeans dan Java Server Page adalah teknologi berbasis Java yang menuntun pembuatan dan pengaturan program yang didesain untuk digunakan pada Web server.

Ringkasan – cont'd

- XML adalah deskripsi standar dokumen yang menggambarkan isi dan struktur dokumen dalam hal penampilan langsung.
- XML berdasarkan HTML dan SGML, yang merupakan standar dokumen yang banyak digunakan.
- XML didesain cukup sederhana untuk manipulasi yang lebih mudah, berbeda dengan SGML, yang hanya memungkinkan pemakai mengembangkan deskripsi dokumen sendiri, tidak seperti HTML.

Ringkasan – cont'd

- Secara khusus, DTD adalah deskripsi dokumen yang independen dari isi dokumen, seperti halnya skema basis data relasional yang merupakan deskripsi basis data yang independen dari instance basis data aktual.
- Dokumen XML mempunyai struktur yang lebih baik daripada basis data relasional yang disebut semistruktur.

Question?

End of Session

Latihan Soal(1)

1. Definisikan arti dan gambarkan penggunaan
 - a. HTML
 - b. URL
 - c. CGI
 - d. pemrosesan server-side
 - e. Java Servlet
2. Apakah CGI itu ? Apa kerugian arsitektur menggunakan skrip CGI
3. Apa perbedaan antara Web server dan application server ?
4. Bagaimanan dokumen XML dibentuk dengan baik ?
5. Misalnya pada toko buku. Diasumsikan pelanggan juga ingin mencari buku berdasarkan judul. Buatlah dokumen HTML yang memungkinkan pemakai menginputkan judul buku.

Latihan Soal(2)

6. Diketahui katalog Eggface computer mail-order.

- “Eggface menjual hardware dan software.
 - Kita menjual Palm Pilot V baru seharga \$400; nomor perangkat 345.
 - Kita juga menjual IBM ThinkPad 570 seharga \$1999; nomor perangkat 3784.
 - Kita menjual baik software bisnis dan entertainment.
 - Microsoft Office 2000 baru datang dan anda dapat membeli Standard Edition dengan harga \$140, nomor perangkat 974.
 - Software dari Adobe yang berjudul InDesign seharga \$200, nomor perangkat 664.
 - Software Game dari Blizzard berjudul Diablo II seharga \$30, nomor perangkat 12,
 - dan anda dapat membeli Starcraft seharga \$10, nomor perangkat 812”
- a. Desain dokumen HTML yang menampilkan item yang ditawarkan Eggface
 - b. Buatlah dokumen XML yang menggambarkan isi dari katalog Eggface
 - c. Buatlah DTD untuk dokumen XML anda dan pastikan dokumen yang anda buat pada pertanyaan terakhir adalah valid untuk DTD.
 - d. Tuliskan query XML-QL yang menampilkan semua software pada katalog
 - e. Tuliskan query XML-QL yang menampilkan harga semua hardware pada katalog.