

Membangun Aplikasi Kamus Bahasa Arab Online Sederhana

Anny Yuniarti
anny@its-sby.edu

Lisensi Dokumen:

Copyright © 2004 IlmuKomputer.Com

Sehuruh dokumen di IlmuKomputer.Com dapat digunakan, dimodifikasi dan disebarkan secara bebas untuk tujuan bukan komersial (nonprofit), dengan syarat tidak menghapus atau merubah atribut penulis dan pernyataan copyright yang disertakan dalam setiap dokumen. Tidak diperbolehkan melakukan penulisan ulang, kecuali mendapatkan ijin terlebih dahulu dari IlmuKomputer.Com.

Bahasa Arab merupakan bahasa umat Islam. Dalam menjalankan ibadahnya umat Islam banyak menggunakan bahasa Arab, misalnya dalam shalat atau dalam membaca Al-Qur'an. Namun adakalanya ibadah yang dilakukan umat Islam terasa hambar tanpa makna dikarenakan umat Islam tidak memahami yang dibacanya ketika shalat, pun ketika membaca Al-Qur'an. Oleh karena itu diperlukan pemahaman bahwa mempelajari bahasa Arab sangat diperlukan bagi umat Islam yang ingin lebih sempurna pelaksanaan ibadahnya, dan penting bagi yang ingin lebih memahami agamanya.

Ilmu bahasa bukanlah ilmu yang sulit. Seorang bayi ketika baru lahir belum mengenal bahasa apapun. Kemudian mulailah ia mengenal kata demi kata dari suara yang didengar dari sekitarnya. Dari kata-kata itu ia rekam sehingga dapat digunakan di kemudian hari ketika sudah mulai berbicara. Begitu pula apabila seseorang ingin menguasai suatu bahasa. Ia harus terlebih dahulu banyak mendengar kosakata bahasa yang ingin dikuasai. Setelah banyak kata dikuasai, barulah ia mulai belajar tata bahasa yang benar dari bahasa tersebut. Oleh karenanya dalam mempelajari bahasa Arab, seseorang haruslah banyak mengenal kosakata bahasa Arab, dan tentu saja mengetahui artinya. Itulah sebabnya adanya kamus merupakan keharusan bagi seseorang yang ingin menguasai bahasa tertentu.

Penggunaan kamus yang mudah dan dapat diakses secara cepat dari tempat mana saja melatarbelakangi pembuatan kamus bahasa Arab-Indonesia dengan berbasis *web*. Dengan begitu, aplikasi dapat dimanfaatkan umat muslim Indonesia secara luas tanpa harus terlebih dahulu meng-*install* program aplikasi tertentu, atau sistem operasi tertentu, misalnya *Microsoft Windows 98 Arabic Enable*.

Permasalahan Yang Ada

Secara sederhana, aplikasi kamus bertugas menyimpan data kamus, yakni data kata sumber dan artinya. Permasalahannya, bagaimana dengan kamus bahasa Arab yang notabene harus menyimpan data karakter Arab? Itulah yang akan kita bahas pada kesempatan kali ini.

Perancangan

Sebelum membuat aplikasi webnya, kita rancang terlebih dahulu struktur basisdatanya. Data yang disimpan adalah kata bahasa Arab, arti kata dalam bahasa Indonesia, dan keterangan tambahan lainnya. (Bisa dikembangkan sesuai kebutuhan, kali ini kita coba yang sederhana).

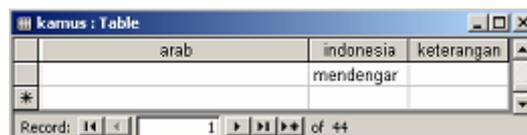
Tabel Kamus		
FIELD	TYPE	KETERANGAN
Id	Long integer	Kode kata
arab	Text[100]	Kata dasar bahasa Arab
indonesia	Text[50]	Arti kata
keterangan	Text[50]	Keterangan tambahan

Cara Penyimpanan Kata Bahasa Arab

Perhatikan Gambar 1 berikut ini!



Gambar 1
Form Tambah Data Kamus



Gambar 2
Isi Tabel Kamus

Jika Anda menyimpan form diatas kedalam *database* dengan metode seperti biasa Anda menyimpan data, data kata bahasa Arab tidak akan tersimpan. Hal ini disebabkan program/sistem yang tidak mengenal karakter Arab. Untuk itu diperlukan suatu pengkodean dari karakter Unicode menjadi simbol-simbol yang memudahkan.

Sekilas Tentang Unicode

Pada pertengahan pertama tahun 1980-an Xerox Corporation berhasil mengimplementasikan hasil penelitiannya mengenai sistem pengkodean karakter multilingual dan berhasil mengeluarkan sebuah proyek yang disebut Unification Code atau Unicode, yang memiliki tujuan utama menyatukan semua karakter yang ada di dunia menjadi sebuah kumpulan karakter. (Searle, 1999). Unicode adalah standard internasional untuk mengkodekan semua bahasa di dunia pada komputer. Unicode dapat mencegah ambiguitas yang muncul ketika menampilkan skrip-skrip kompleks seperti bahasa Jepang, Arab atau Cina pada sistem komputer. Unicode dapat berjalan lebih cepat di Windows NT, 2000 dan XP. (Gerhards, 2001)

Karakter-karakter yang biasa dipakai (seperti alphabet ANSI) berdasarkan karakter 8 bit yang disebut byte. Satu byte dapat merepresentasikan lebih dari 256 nilai yang berbeda. Skrip yang biasa digunakan dalam bahasa Inggris, Perancis atau Jerman sudah cukup dengan menggunakan representasi ini. Tetapi untuk bahasa-bahasa yang kompleks seperti bahasa Jepang atau Korea, 256 karakter yang berbeda ini tidak akan mencukupi.

Sehingga user dari bahasa-bahasa ini telah membangun double byte character sets, yang disebut DBCS. Pada DBCS, masing-masing karakter direpresentasikan oleh satu atau lebih byte. Sayangnya, terdapat banyak encoding DBCS untuk bahasa yang diberikan. Sistem operasi dan bahasa pemrograman yang berbeda menggunakan encoding DBCS yang berbeda pula. Dan lagi pemrograman relatif lebih kompleks karena operasi-operasi parsing byte.

Tujuan Unicode adalah untuk memecahkan masalah ini dengan menggunakan lebih dari satu byte untuk setiap karakter. Pada implementasi tertentu, digunakan 2 byte, dan mampu merepresentasikan 65.564 karakter berbeda. Ini cukup untuk menyimpan hampir semua karakter di dunia. Jadi dengan Unicode, semua karakter yang berbeda dapat disimpan dalam satu string. Dan karena semua karakter memiliki lebar yang tetap, kompleksitas pemrograman banyak tereduksi.

Aplikasi yang menggunakan Unicode mampu menyimpan dan memproses semua karakter di dunia. Untuk dapat mengirimkan string Unicode sebagai data ke sebuah program yang tidak 'mengetahui Unicode' dan hanya mendukung encoding yang berbeda, gunakan "escapes". Selain dengan menggunakan

“escapes”, dapat juga dengan mentransformasikan semua data pada field menjadi bentuk hex, misal dalam UTF-16 menjadi "03B10020226400200033" atau dalam UTF-8 menjadi "CEB120E289A42033". Akan tetapi dengan cara ini membutuhkan space lebih dan kurang enak dibaca. Cara yang lain adalah dengan mentransformasi menjadi legacy encoding. Hal ini akan menyebabkan korupsi pada data yang tidak dapat diekspresikan dalam legacy encoding. Sebagian besar bahasa pemrograman saat ini telah memiliki tipe data spesial untuk karakter Unicode/ ISO 10646-1, misalnya Java, Perl, C# dan lain sebagainya.

Kembali ke cara penyimpanan kata bahasa Arab. Dalam aplikasi yang akan kita buat ini nanti, satu buah karakter Arab akan dikonversikan menjadi tujuh buah karakter teks untuk kemudian disimpan dalam *database*. Buatlah fungsi JavaScript yang akan mengkonversikan karakter tersebut sebagai berikut:

```
0    function konversi()
1    {        var kata = new Array();
2            var i = 0;
3            var str2 = "";
4
5            for (i = 0; i < document.form1.inputkata.value.split("").length; i++)
6            {        kata[i] = document.form1.inputkata.value.split("")[i];
7                    str2 = str2 + "&#x" + escape(kata[i]).substr(3,3) + ";" ;
8            }
9
10           document.form1.katanya.value = str2;
11           document.form1.submit();
12    }
```

Lihat isi database Anda. Dengan data seperti pada Gambar 1 maka database Anda akan berisi seperti pada Gambar 3 berikut ini:



	arab	indonesia	keterangan
*	سَمِعَ	mendengar	

Gambar 3
Isi Tabel Kamus

Dengan demikian Anda sudah bisa menyimpan kata bahasa Arab. Bagaimana menampilkannya? Tampilkan seperti biasa Anda menampilkan data dari *database*. Pengkonversian yang dilakukan pada fungsi JavaScript sebelumnya dilakukan untuk dapat menampilkan di web juga. Perhatikan baris 7 dari fungsi konversi diatas. Berikut ini hasil pencarian kamus dengan data kata سَمِعَ. Perlu diketahui dalam bahasa Arab juga dikenal pola kata. Kata سَمِعَ termasuk dalam pola fi'il madhi yang artinya kata kerja lampau (sudah dilakukan).

Gambar 4
Contoh Hasil Penggunaan Kamus

Pembuatan Form Masukan

Permasalahan kedua, agar kamus bahasa Arab online Anda dapat digunakan oleh semua pengguna tanpa harus meng-install perangkat lunak atau sistem operasi tertentu, buatlah form masukan yang memudahkan pengguna untuk menulis kata bahasa Arab. Untuk membuatnya terlebih dahulu tambahkan Arabic Language Support pada komputer Anda. Anda dapat men-download di <http://zsigri.tripod.com/fontboard/fbarab.zip>. Contoh form yang dimaksud adalah sebagai berikut:

Gambar 5
Form Masukan Kata Bahasa Arab

Pada gambar diatas ketika pengguna menekan tombol  maka *textbox* akan berisi karakter yang dimaksud **س**. Bagaimana membuatnya? Buatlah gambar-gambar yang memuat huruf-huruf hijaiyah. Susun gambar-gambar tersebut pada halaman web. Kemudian pada tag tambahkan `onClick="form1.txt.value=form1.txt.value + 'Ó';"`

Kode Ó diperoleh dengan menekan tombol s pada keyboard yang telah terinstall *Arabic Language Support*. Itulah sebabnya untuk membuat form masukan ini digunakan Fontboard Arabic Keyboard.

Jadi, jika sudah berhasil membuatnya, jangan lupa lengkapi data kamus Anda, karena aplikasi kamus yang baik adalah aplikasi kamus yang didukung dengan data yang lengkap. Dan jika Anda sudah memahami bahasa Arab, Anda bisa mengembangkan aplikasi Anda dengan membuatnya dapat menerima pola-pola turunan tanpa harus menyimpan seluruh kata. Selamat mencoba!

BIOGRAFI PENULIS



Anny Yuniarti. Lahir di Surabaya, 22 Juni 1981. Lulus dari Teknik Informatika ITS tahun 2003. Sekarang bekerja sebagai staf pengajar/dosen di jurusan dan institut yang sama.

Informasi lebih lanjut tentang penulis ini bisa didapat melalui:
Email: anny@its-sby.edu