

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Bahasa adalah alat pemersatu bagi suatu komunitas, suku bangsa, bahkan masyarakat global. Bahasa juga sangat penting untuk mempertahankan identitas budaya yang mengalir dalam darah suatu entitas sosial. Bahasa Indonesia adalah bahasa pemersatu bagi seluruh suku bangsa yang mendiami nusantara, namun dalam konteks kedaerahan, bahasa Sunda adalah bahasa kedua terbesar setelah bahasa Jawa. Kini, Menghadapi era globalisasi dimana kompetisi hadir di seluruh aspek kehidupan, penggunaan bahasa pun tak luput dari hal tersebut. Indonesia yang memang sejak zaman dahulu kala terbuka dengan perubahan internasional telah terpengaruh oleh arus globalisasi dimana penggunaan bahasa asing kini semakin digencarkan. Sementara itu, disisi lain, ternyata penggunaan bahasa daerah di Indonesia semakin tergerus karena sistem pendidikan di banyak sekolah kurang mengadopsi pemeliharaan nilai-nilai budaya lokal dan tradisional dan justru terlena oleh hal-hal yang terlalu berbau internasional tanpa melalui filterisasi.

Hal ini terlihat efeknya pada generasi muda dari suku Sunda, yang lebih sering menggunakan bahasa Indonesia dalam pergaulan sehari-hari [11]. Akibatnya bahasa Sunda pun tidak terjaga kelestariannya. Dengan demikian untuk menjaga kelestarian budaya Sunda, yaitu bahasa diperlukan usaha untuk menjaga bahasa, salah satunya adalah mengumpulkan kata-kata bahasa Sunda dalam bentuk digital. Pengumpulan kata bahasa Sunda pernah dikerjakan oleh Novihana Nurani Setiawan yaitu Pembangunan Basis Data Leksikal Wordnet Bahasa Sunda. Namun pada proyek tersebut pengumpulan kata bahasa sunda masih dilakukan dengan cara memasukkan kata satu per satu sehingga memerlukan waktu yang lama untuk pengumpulan kata dan butuh pengguna yang banyak, sedangkan pengguna yang memasukkan kata tidak dibatasi pada pengguna dari suku Sunda saja [14].

Maka dari itu diperlukan sebuah cara untuk mengumpulkan kata secara cepat, yaitu dengan mencari kata bahasa Sunda pada halaman *web* seperti Wikipedia, Blog atau, *web* berita. Dalam proyek akhir ini akan menggunakan halaman *web* Wikipedia dikarenakan halaman blog memiliki bahasa yang tergantung pada penulis blog tanpa ada translasi ke bahasa lain dan halaman *web* berita hanya memiliki bahasa Indonesia dan bahasa Inggris saja sedangkan wikipedia memiliki translasi bahasa yang beraneka ragam. Disamping itu wikipedia telah memiliki 19.123 artikel.

1.2 Rumusan Masalah

Adapun rumusan masalah dari proyek akhir adalah sebagai berikut:

1. Bagaimana mengumpulkan kata bahasa Sunda yang bersumber dari wikipedia bahasa Sunda?
2. Bagaimana cara mengetahui arti kata bahasa Sunda yang telah dikumpulkan?

1.3 Tujuan

Adapun tujuan dari dari proyek akhir ini sebagai berikut:

1. Membangun aplikasi yang dapat mengumpulkan kata bahasa Sunda dari halaman Wikipedia.
2. Membangun aplikasi yang dapat mengetahui arti kata bahasa Sunda yang telah dikumpulkan dengan membandingkan pada kamus *online* bahasa Sunda.

1.4 Batasan Masalah

Adapun batasan-batasan masalah dalam Proyek Akhir ini adalah sebagai berikut.

1. Konsep crawling yang digunakan adalah dengan cara menuliskan sumber halaman dan melakukan pencarian halaman padanan dengan menggunakan *library* PHP Simple HTML DOM.

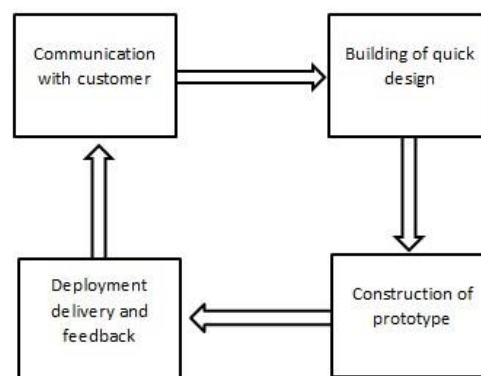
2. Untuk mengetahui arti dari kata yang sudah di kumpulkan, akan di di bandingkan dengan kamus *online* yang sedang di kembangkan pada proyek akhir dengan judul Ekstraksi dan Pengelompokkan Kata Bahasa Sunda Berbasis *WordNet*, oleh Billy Alghifary.
3. Pengumpulan kata pada aplikasi ini hanya untuk bahasa Sunda dan Indonesia.

1.5 Definisi Operasional

Aplikasi Ekstraksi Kata Sunda merupakan sebuah aplikasi yang berguna untuk melakukan pengumpulan kata pada halaman *web* Wikipedia bahasa Sunda dan bahasa Indonesia. Kata yang telah dikumpulkan nantinya akan dibandingkan dengan kamus *online* untuk mengetahui arti dari setiap kata.

1.6 Metode Pengerjaan

Adapun metode pengerjaan yang akan di gunakan pada proyek akhir ini adalah Metode *Prototype* dengan tahapan sebagai berikut :



Gambar 1-1 Metode SDLC *Prototype*

1. *Communication with Customer*

Pada tahap ini mengidentifikasi kebutuhan data yang akan digunakan untuk membangun aplikasi ekstraksi kata bahasa Sunda dengan mencari sumber kata bahasa sunda di internet, misalnya Wikipedia dan Blog.

2. *Building of Quick Design*

Mendesain adalah tahap dimana sebuah proses yang fokus pada desain pembuatan aplikasi ekstraksi kata bahasa Sunda yaitu dengan membuat *data flow diagram* untuk penggambaran fungsionalitas sistem serta *entity relationship diagram* untuk penggambaran relasi antar tabel yang terdapat pada *database*.

3. *Construction of Prototype*

Setelah desain dibuat, aktifitas selanjutnya adalah membuat prototype dari aplikasi ekstraksi kata sunda. *Prototype* adalah *sample* dari keseluruhan aplikasi yang sedang dibuat. *Prototype* ini mengadopsi fungsi-fungsi utama dari aplikasi itu sendiri, namun masih sebatas pengembangan dan dalam tahap pengujian.

4. *Deployment Delivery and Feedback*

Pada tahap ini aplikasi *prototype* diuji, diimplementasikan, dievaluasi dan dimodifikasi berulang-ulang yang bertujuan menemukan kesalahan-kesalahan yang terjadi pada aplikasi yang telah dibangun. Pengujian dilakukan dengan menggunakan *whitebox testing* karena aplikasi sedikit memiliki tampilan untuk mendari masukan dari user

1.7 Jadwal Pengerjaan

Tabel 1-1 Jadwal Pengerjaan

No	Kegiatan	2015			2016														
		Desember	Januari	Februari	Maret	April													
1.	Communication with customer - Mencari sumber kata																		
2.	Building of Quick Design - Membuat DFD - Membuat ERD - Membuat rancangan <i>interface</i>																		
3.	Construction of Prototype - Membuat modul crawling - Membuat modul pencarian																		
4.	<i>Deployment Delivery and Feedback</i> - Implementasi - Pengujian																		