

# MOBILE BLOG UNTUK CITIZEN JOURNALISM DENGAN PENGKATEGORIAN BERITA MENGGUNAKAN METODE INNER PRODUCT

Yuliana Setiowati, Afrida Helen, Lilik Istianah  
Politeknik Elektronika Negeri Surabaya  
Institut Teknologi Sepuluh Nopember  
Kampus PENS-ITS Sukolilo Surabaya 60111

E-mail : [yuliana@eepis-its.edu](mailto:yuliana@eepis-its.edu); [helen@eepis-its.edu](mailto:helen@eepis-its.edu); [hanaitsi@gmail.com](mailto:hanaitsi@gmail.com)

## ABSTRAK

*Saat ini tidak hanya wartawan yang bisa menulis berita, semua orang bisa membuat suatu berita, yang kemudian disebut sebagai citizen journalism. Sebuah kejadian yang dapat diberitakan, bisa terjadi kapan saja di mana saja, untuk itu diperlukan suatu media pengiriman berita yang cepat dan praktis salah satunya dengan menggunakan telepon genggam. Untuk mendukung kegiatan citizen journalism tersebut dibuat sebuah aplikasi telepon genggam yang bisa mengirimkan berita berupa gambar dan teks kemudian akan ditampilkan pada sebuah blog menurut kronologi waktu dan diklasifikasikan berdasarkan kategori berita (politik, ekonomi, olahraga, kriminal, hiburan).*

*Berita-berita yang dikirimkan, untuk memudahkan pembacaan maka berita akan dikategorikan secara otomatis. Input berita akan melalui proses text mining, yaitu tokenizing, filtering dan stemming, kemudian hasilnya akan diproses dengan kumpulan kata kunci tiap kategori berita menggunakan metode inner product.*

*Hasil uji coba menunjukkan bahwa rata-rata tingkat keakuratan hasil klasifikasi berita dengan metode inner product sekitar 92%. Metode ini tergantung pada banyaknya kata kunci tiap kategori berita, semakin banyak kata kunci tiap kategori berita, semakin baik hasil pengkategorian.*

Kata kunci: berita, blog, text mining, tokenizing, filtering, stemming, inner product

## 1. Pendahuluan

Saat ini tidak hanya wartawan yang bisa menulis berita, dengan media teknologi yang semakin canggih semua orang bisa membuat suatu berita, yang kemudian disebut sebagai citizen journalism, namun tidak banyak redaksi berita yang menerima berita-berita dari citizen journalism dengan pertimbangan aturan jurnalis,

biasanya redaksi berita membuat sebuah blog atau halaman khusus yang menampung berita-berita dari citizen journalism dengan mengusung aturan kebebasan dalam menulis tanpa diatur oleh aturan-aturan pemberitaan.

Sebuah kejadian yang dapat diberitakan bisa terjadi kapan saja di mana saja, untuk itu diperlukan suatu media pengiriman berita yang cepat dan praktis, saat ini dengan bermodalkan telepon genggam seseorang dapat mengambil gambar atau merekam peristiwa untuk dijadikan berita, maka untuk kecepatan penyampaian sebuah berita muncul ide untuk membuat sebuah aplikasi pengiriman berita ke blog dengan media telepon genggam.

Dalam penerimaan berita, editor blog untuk citizen journalism akan mengecek terlebih dahulu kelayakan dari berita yang akan ditampilkan. Sebelumnya pengelompokan berita yang masuk dilakukan secara manual sesuai dengan kategori berita. Pada penelitian ini sebuah aplikasi telepon genggam yang bisa mengirimkan berita berupa gambar dan teks, kemudian berdasarkan teks beritanya, berita-berita yang dikirimkan akan dikategorikan secara otomatis menggunakan metode inner product kemudian ditampilkan berdasarkan kronologis waktu pada sebuah blog.

## 2. Tinjauan Pustaka

### 2.1 BLOG

Blog merupakan singkatan dari "web log" adalah bentuk aplikasi web yang menyerupai tulisan-tulisan pada sebuah halaman web umum. Tulisan-tulisan ini seringkali dimuat dalam urutan terbalik (isi terbaru dahulu baru kemudian diikuti isi yang lebih lama), meskipun tidak selamanya demikian.

Blog mempunyai fungsi yang sangat beragam, dari sebuah catatan harian, media publikasi dalam sebuah kampanye politik, sampai dengan program-program media dan perusahaan perusahaan. Sebagian blog dipelihara oleh seorang penulis tunggal, sementara sebagian

lainnya oleh beberapa penulis. Banyak juga weblog yang memiliki fasilitas interaksi dengan para pengunjungnya, seperti menggunakan buku tamu dan kolom komentar yang dapat memperkenalkan para pengunjungnya untuk meninggalkan komentar [1]

## 2.2 MOBILE BLOG

Mobile blog adalah salah satu bentuk cara posting pada blog dengan menggunakan media mobile secara langsung, contohnya dengan menggunakan telepon genggam. Mobile blog membantu pengguna blog yang memiliki mobilitas tinggi, sehingga memungkinkan untuk mengirim posting yang berupa gambar atau video, dengan media email atau MMS. Tidak hanya pada sisi pengirim, yaitu perangkat mobile, pada sisi server atau web yang akan menampilkan posting blog ini juga harus memiliki dukungan khusus untuk menerima kiriman posting dari telepon genggam, sehingga bisa saling berhubungan[2].

## 2.3 TEXT MINING

### 2.3.1 Pengertian text mining

Text mining adalah proses menemukan informasi dari beberapa data yang berupa dokumen teks dengan tujuan mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen [3]

Text mining yang digunakan dalam proyek akhir ini terdiri dari beberapa proses, diantaranya adalah tokenizing, filtering, stemming

### 2.3.2 Tokenizing

Tokenizing yaitu proses menghilangkan angka-angka tanda baca ( , . ? ! ) ( - \* \$ # ; : % ^ @ / & + = ' \n \t) dan pemotongan strip input berdasarkan tiap kata yang menyusunnnya

Contoh:

Input: konser hiburan di taman kota menampilkan penyanyi dan pelawak

Output:

konser  
hiburan  
di  
taman  
kota  
menampilkan  
penyanyi  
dan  
pelawak

### 2.3.3 Filtering

Filtering yaitu mengambil kata-kata penting dari hasil tokenizing, menggunakan algoritma stop list, yaitu membuang kata-kata yang kurang penting kemudian dijadikan sebagai kata kunci

untuk proses selanjutnya, dari contoh di atas dapat diambil kata-kata

konser  
hiburan  
taman  
kota  
menampilkan  
penyanyi  
pelawak

kata di , dan dihilangkan

### 2.3.4 Stemming

Stemming yaitu tahap mencari kata dasar dari tiap kata hasil filtering dengan menghilangkan imbuhan-imbuhan pada kata tersebut. Untuk kata-kata dalam bahasa Indonesia digunakan pola kata sebagai berikut:

Algoritma ini didahului dengan pembacaan tiap kata dari file sampel. Sehingga input dari algoritma ini adalah sebuah kata yang kemudian dilakukan :

1. Pemeriksaan semua kemungkinan bentuk kata. Setiap kata diasumsikan memiliki 2 Awalan (prefiks) dan 3 Akhiran (sufiks). Sehingga bentuknya menjadi :

**Prefiks 1 + Prefiks 2 + Kata dasar + Sufiks 3 + Sufiks 2 + Sufiks 1**

Seandainya kata tersebut tidak memiliki imbuhan sebanyak imbuhan di atas, maka imbuhan yang kosong diberi tanda x untuk prefiks dan diberi tanda xx untuk sufiks.

2. Pemotongan dilakukan secara berurutan sebagai berikut :

AW= awalan, AK=akhiran, KD= kata dasar

AW : AW

AK : AK

KD : KD

- a. AW I, hasilnya disimpan pada p1
- b. AW II, hasilnya disimpan pada p2
- c. AK I, hasilnya disimpan pada s1
- d. AK II, hasilnya disimpan pada s2
- e. AK III, hasilnya disimpan pada s3

Pada setiap tahap pemotongan di atas diikuti dengan pemeriksaan di kamus apakah hasil pemotongan itu sudah berada dalam bentuk dasar. Kalau pemeriksaan ini berhasil maka proses dinyatakan selesai dan tidak perlu melanjutkan proses pemotongan imbuhan lainnya. Contoh pemenggalan kata “mempermainkannya”

Langkah 1 : Cek apakah kata ada dalam kamus

Ya : Success

Tidak : lakukan pemotongan AW I

Kata = memainkannya

Langkah 2 : Cek apakah kata ada dalam kamus

Ya : Success

Tidak : lakukan pemotongan AW II

Kata = mainkannya

- Langkah 3 : Cek apakah kata ada dalam kamus  
Ya : Success  
Tidak : lakukan pemotongan AK I  
Kata = mainkan
- Langkah 4 : Cek apakah kata ada dalam kamus  
Ya : Success  
Tidak : lakukan pemotongan AK II  
Kata = main
- Langkah 5 : Cek apakah kata ada dalam kamus  
Ya : Success  
Tidak : lakukan pemotongan AK III. Dalam hal ini AK III tidak ada, sehingga kata tidak diubah.  
Kata = main
- Langkah 6 : Cek apakah kata ada dalam kamus  
Ya : Success  
Tidak : "Kata tidak ditemukan"

3. Jika sampai pada pemotongan AK III, belum juga ditemukan di kamus, maka dilakukan proses kombinasi. KD yang dihasilkan dikombinasikan dengan imbuhan-imbuhan dalam 12 konfigurasi berikut :
- KD
  - KD + AK III
  - KD + AK III + AK II
  - KD + AK III + AK II + AK I
  - AW I + AW II + KD
  - AW I + AW II + KD + AK III
  - AW I + AW II + KD + AK III + AK II
  - AW I + AW II + KD + AK III + AKII + AKI
  - AW II + KD
  - AW II + KD + AK III
  - AW II + KD + AK III + AK II
  - AW II + KD + AK III + AK II + AK I

Sebenarnya kombinasi a, b, c, d, h, dan l sudah diperiksa pada tahap sebelumnya, karena kombinasi ini adalah hasil pemotongan bertahap tersebut. Dengan demikian, kombinasi yang masih perlu dilakukan tinggal 6 yakni pada kombinasi-kombinasi yang belum dilakukan (e, f, g, i, j, dan k). Tentunya bila hasil pemeriksaan suatu kombinasi adalah 'ada', maka pemeriksaan pada kombinasi lainnya sudah tidak diperlukan lagi. Pemeriksaan 12 kombinasi ini diperlukan, karena adanya fenomena overstemming pada algoritma pemotongan imbuhan. Kelemahan ini berakibat pada pemotongan bagian kata yang sebenarnya adalah milik kata dasar itu sendiri yang kebetulan mirip dengan salah satu jenis imbuhan yang ada. Dengan 12 kombinasi itu, pemotongan yang sudah terlanjur tersebut dapat dikembalikan sesuai posisinya [5]

### 2.4 INNER PRODUCT

Inner product adalah merupakan istilah lain dari dot product atau dapat juga disebut sebagai scalar product adalah perkalian antara dua vektor

yang dapat berupa vektor baris atau vektor kolom , sehingga hasilnya berupa scalar dengan rumus sebagai berikut:[4]

$$a.b = \sum_{i=1}^n a_i b_i = a_1 b_1 + a_2 b_2 + \dots + a_n b_n$$

Keterangan:  
a = elemen vektor 1  
b = elemen vektor 2

Jika diilustrasikan dalam kasus ini sebagai berikut:

Kata-kata hasil stemming akan menjadi elemen vektor satu dengan nilai masing-masing 1 , kata-kata yang termasuk dalam kata kunci tiap kategori berita sebagai elemen dari vektor dua, sehingga nilai dari elemen vektor satu akan dikalikan dengan nilai dari elemen vektor dua, di mana nilai elemen pada vektor dua akan memiliki nilai tidak sama dengan nol, apabila kata hasil stemming sama dengan kata kunci tersebut, contoh perkaliannya dapat dilihat pada tabel 1 berikut

Tabel 1 Contoh proses inner product

Kata kunci /Kategori	konser	Hibur				nyanyi	lawak	hasil	...
Politik	0	0	0	0	0	0	0	0	...
Ekonomi	0	0	0	0	0	0	0	0	...
kriminal	0	0	0	0	0	0	0	0	...
Olahraga	0	0	0	0	0	0	0	0	...
Hiburan	1	1	0	0	0	1	1	4	
Input nilai	1	1	1	1	1	1	1		...
Input kata	konser	Hibur	taman	kota	tampil	nyanyi	lawak		...

Prosesnya perhitungan dari tabel 1 diatas adalah sebagai berikut:

- Politik:  
(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)=0
- Ekonomi:  
(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)=0
- Kriminal:  
(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)=0
- Olahraga:  
(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)=0
- Hiburan:  
(1)(1)+(1)(1)+(1)(0)+(1)(0)+(1)(0)+(1)(1)+(1)(1)=4

Dari proses diatas terlihat hasil tertinggi adalah hiburan, sehingga didapatkan output bahwa inputan teks berita termasuk dalam kategori hiburan.

### 3. Perancangan Sistem

#### 3.1 Perancangan Database

Untuk menyimpan kata kunci tiap kategori berita dibutuhkan 5 tabel, yaitu tabel kata\_politik, kata\_ekonomi, kata\_olahraga, kata\_kriminal, dan kata\_hiburan yang masing-masing tabel memiliki struktur sebagai berikut:

Tabel 2 struktur tabel kata kunci

Nama Field	Tipe data
Kata	varchar (2)
Jumlah	Int(10)

Berita yang dikirimkan oleh user akan disimpan dalam sebuah tabel database dengan struktur tabel sebagai berikut:

Tabel 3 struktur tabel berita

Nama field	Tipe data
id (primary key)	int (10)
Judul	varchar (200)
Isiberita	varchar (2000)
Gambar	varchar(15)
Kategori	varchar(15)
Tanggal	varchar(30)
Iduser	varchar(10)

Data-data user terdaftar akan disimpan pada tabel akun yang memiliki struktur sebagai berikut:

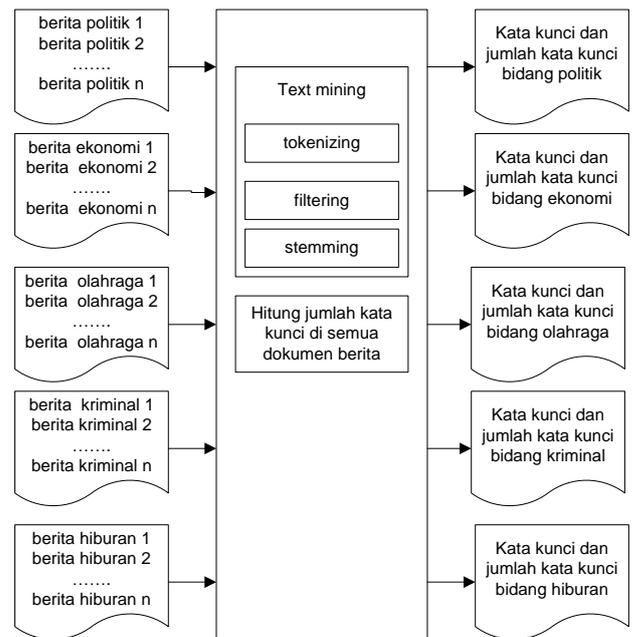
Tabel 4 Struktur tabel akun

Nama field	Tipe data
iduser (primary key)	int (11)
nama	varchar (20)
username	varchar(10)
pass	varchar (10)
email	varchar(20)

#### 3.2 Pengumpulan Data

Data yang dikumpulkan untuk kata kunci tiap kategori berita, yaitu berita kategori politik, ekonomi, olahraga, kriminal dan hiburan didapat dari situs berita (cetak.kompas.com dan www.kabarindonesia.com) kemudian berita-berita ini akan dilakukan proses text mining dan akan didapatkan kata kunci tiap kategori berita sekaligus jumlah kata tersebut di semua berita yang dikumpulkan

Gambaran proses pengumpulan kata kunci dapat diilustrasikan dengan gambar berikut:



Gambar 1 proses mendapatkan kata kunci dari berita

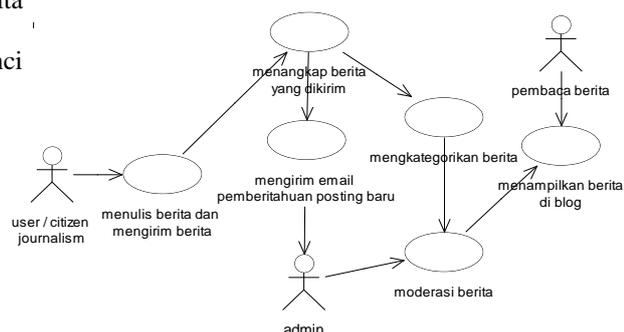
Berikut contoh kata kunci yang dihasilkan dari proses text mining berita politik

Tabel 5 Contoh kata kunci yang dihasilkan dari proses text mining berita politik

Kata	jumlah	Kata	jumlah
Partai	174	Komisi	40
Caleg	117	pemerintah	38
Pemilu	114	Undang	36
Suara	114	Rakyat	36
KPU	102	presiden	34
politik	93	kabupaten	34
anggota	71	kampanye	33
DPR	67	lembaga	33
masyarakat	59	Parpol	32
calon	58	RUU	28
Partai	174	Komisi	40
Caleg	117	pemerintah	38
Pemilu	114	Undang	36
Suara	114	Rakyat	36

#### 3.3 Perancangan Sistem

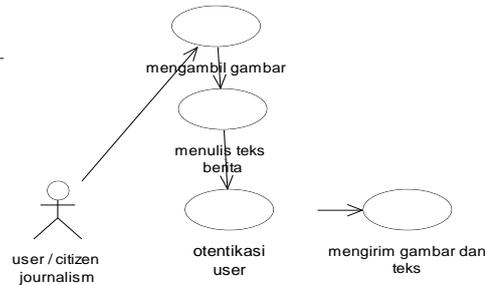
Secara umum sistem yang akan dibuat dapat digambarkan sebagai berikut:



Gambar 2 Use case gambaran system keseluruhan

### 3.4 Aplikasi pada telepon genggam

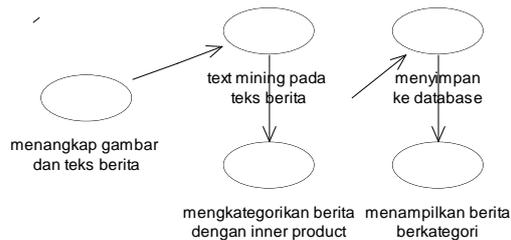
Secara umum sistem aplikasi pada telepon genggam yang akan dibuat dapat digambarkan sebagai berikut



Gambar 3 Use case diagram aplikasi mobile

### 3.5 Aplikasi Blog

Aplikasi Blog pada server seperti di bawah ini:



Gambar 4 activity diagram rancangan sistem aplikasi blog

## 4. Pengujian dan Analisa

Uji coba sistem dilakukan dilakukan dengan menggunakan telepon genggam sony ericsson tipe K630i dan blog sudah diupload dan terkoneksi dengan internet.

### 4.1 Uji coba pengiriman berita menggunakan telepon genggam dan blog yang sudah diupload

Untuk pengujian system, langkah pertama yang dilakukan adalah menginstall file jar dari aplikasi mobile blog pada telepon genggam, pada percobaan kali ini akan digunakan telepon Sony Erricsson dengan tipe K630i

Untuk webnya, dilakukan upload pada web server, dengan alamat URL <http://devie.or.id:8080/mobileblog>. Pertama kali dijalankan akan ditampilkan tampilan awal dari aplikasi mobileblog.



Gambar 5 tampilan awal mobile blog di telepon genggam

Kemudian menuju menu kamera, dan kamera aktif, pilih menu ambil untuk mengambil gambar, dan hasil gambar akan ditampilkan



Gambar 6 menampilkan hasil foto pada telepon genggam

Pilih menu proses dan akan tampil halaman untuk menuliskan judul dan isi berita



Gambar 7 tampilan menulis judul berita di telepon genggam



Gambar 8 tampilan menulis isi berita di telepon genggam

Pilih menu kirim dan menuju form untuk mengisi username dan password untuk otentikasi



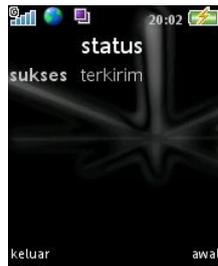
Gambar 9 tampilan otentikasi user pada telepon genggam

Jika otentikasi benar, dilanjutkan pada proses pengiriman berita



Gambar 10 tampilan proses mengirimkan berita di telepon genggam

pengiriman berita selesai maka muncul status pengiriman berita



Gambar 11 tampilan status pengiriman berita di telepon genggam

Berita akan ditampilkan di halaman blog



Gambar 12 Hasil pengiriman berita menggunakan telepon genggam

## 4.2 Uji coba pengkategorian berita

Kemudian dilakukan analisa terhadap pengiriman 10 berita pada tiap-tiap kategori, berita diambil dari situs berita, [www.kompas.com](http://www.kompas.com) sebagai berikut

Tabel 6 persentase percobaan pengkategorian

kategori	benar	salah	persentasi
politik	9	1	90%
ekonomi	9	1	90%
olahraga	10	0	100%
kriminal	9	1	90%
hiburan	9	1	90%
Rata-rata			92%

Dari percobaan diatas dapat dilihat rata-rata kemampuan sistem untuk mengkategorian berita cukup baik, sekitar 92%. Rata-rata kesalahan pengkategorian disebabkan karena berita asli terkategori dalam kategori tertentu tapi dalam berita tersebut terkandung banyak kata kunci dari berita lain, misal berita asli dikategorikan sebagai berita kriminal, tetapi di dalam berita tersebut banyak mengandung kata-kata yang termasuk dalam kata kunci kategori ekonomi, sehingga proses perhitungan akan menentukan berita tersebut termasuk dalam berita ekonomi.

Semakin banyak berita yang dimasukkan dalam training maka kata kunci untuk tiap bidang yang didapat juga semakin banyak, hal ini akan meningkatkan tingkat keakuratan dari proses pengklasifikasian berita.

## 5. Kesimpulan.

Setelah dilakukan pengujian dan analisa program, maka dapat diperoleh kesimpulan sebagai berikut :

- dari hasil percobaan, menunjukkan aplikasi ini mampu digunakan untuk mengirimkan berita berupa teks dan gambar melalui aplikasi telepon genggam ke blog
- sistem mampu mengkategorikan secara otomatis berita-berita yang dikirimkan dengan menggunakan metode inner product dengan tingkat keakuratan sekitar 92%.
- Kevalidan output program sangat ditentukan oleh pengklasifikasian kata ke dalam masing-masing bidang. Semakin banyak berita yang dimasukkan dalam data training maka kata kunci yang didapat semakin banyak.

## DAFTAR PUSTAKA

- [1] [id.wikipedia.org/wiki/Blog](http://id.wikipedia.org/wiki/Blog)
- [2] [en.wikipedia.org/wiki/Mobile\\_blogging](http://en.wikipedia.org/wiki/Mobile_blogging)

- [3] Tessy Badriyah, *materi kuliah Text Mining*
- [4] Andri Permana. *klasifikasi topik TA sesuai dengan kemampuan menggunakan metoda inner product.* Surabaya: PENS-ITS. 2008
- [5] Arifin Zainal Agus dan Setiono Novan Ari, *Klasifikasi Dokumen Berita Kejadian Berbahasa Indonesia dengan Algoritma Single Pass Clusterin.* Surabaya: Teknik Informatika-ITS
- [6] [www.gealgeol.com/2008/02/12/mysql.html](http://www.gealgeol.com/2008/02/12/mysql.html)
- [7] [http://www.wirelessdevnet.com/channels/java/features/j2me\\_http.phtml](http://www.wirelessdevnet.com/channels/java/features/j2me_http.phtml)
- [8] Budi Raharjo, *Tuntunan Pemrograman Java untuk Handphone* Bandung. Informatika : 2007
- [9] M Sholahuddin, *Pemrograman J2ME belajar cepat pemrograman perangkat telekomunikasi mobile.* Bandung. Informatika : 2006
- [10] modul kuliah pemrograman lanjut, PENS ITS
- [11] Leo Willyanto Santoso, Sukanto Tedjokusumo, Marcel Renaldy Soetanto. *Aplikasi Pelaporan Berita Emergensi Secara Visual dan Tekstual Lewat Telepon Selular.* Surabaya: Universitas Kristen Petra. 2005
- [12] Jonathan Knudsen. *Introduction to mobile blogging.* <http://developers.sun.com>. 2003