

# **TEXT FILTERING KATA PORNO DENGAN METODE BOYER MOORE PADA APLIKASI ELEARNING BERBASIS CMS DI SDN 02 TUREN**

Edward Satya Irawan  
Moh. Sulhan

<sup>1</sup> Sistem Informasi, Universitas Kanjuruhan Malang, edward\_satya@yahoo.com

<sup>2</sup> Sistem Informasi, Universitas Kanjuruhan Malang, hans2net@yahoo.com

## **ABSTRAK**

SDN 02 Turen terdapat media pembelajaran *elearning* yang berbasis cms yang sangat memungkinkan tersisip kata-kata yang dianggap porno dalam proses belajar mengajar yang disajikan oleh guru di SDN 02 Turen. Jika tersisip kata yang buruk mengakibatkan pengaruh buruk terhadap mental, kepribadian dan tata bahasa yang kurang sopan yang akan ditirukan oleh anak sekolah dasar. Untuk mengantisipasi tersisipnya kata-kata yang dianggap porno maka dirancang sebuah *filtering* kata yang dianggap porno pada aplikasi *elearning* SDN 02 Turen yang belum terdapat *filtering* kata untuk menyaring kata-kata yang dianggap porno agar informasi yang disampaikan lebih berkualitas dan dapat menjadi acuan dalam belajar siswa di SDN 02 Turen.

Kesimpulan dari *filtering* kata yang dianggap porno dengan metode pencarian kata *Boyer Moore* ini terbukti dapat mengurangi terjadinya kesalahan dalam tersisipnya kata yang dianggap porno.

## **ABSTRACT**

*SDN 02 Turen are media-based learning elearning cms is very possible inserted the words that are considered pornographic in the learning process is presented by a teacher at SDN 02 Turen. If inserted a bad word resulting impacts on our mental, personality and irreverent grammar that will imitated by elementary school children. To anticipate tersisipnya words that are considered pornographic then designed a filtering words that are considered pornographic at SDN 02 Turen elearning applications that is not in word filtering to filter out words that are considered pornographic information submitted to higher quality and can be a reference for student learning in SDN 02 Turen.*

*The result of the conclusion of the decision support system is proven to reduce the occurrence of errors in words that are considered pornographic tersisipnya*

**Keyword :** *Text Filtering, Word Porn, Booyer Moore*

### **1. Pendahuluan**

Saat ini kebutuhan akan teknologi meningkat drastis, seiring dengan perkembangan teknologi dunia yang sangat pesat. Hampir semua lapisan masyarakat seperti yang hidup di jaman ini mengandalkan teknologi dalam segala kegiatannya. (Menurut Hery Agus dalam *paper* yang berjudul Perkembangan Teknologi Informasi, 2009), negara-negara besar di dunia Amerika Serikat, Jepang, dan Cina terus berlomba-

lomba dalam menghadirkan inovasi baru dalam perkembangan teknologi dunia. Teknologi yang terpesat dalam perkembangannya adalah teknologi informasi yang menghadirkan beragam pilihan bentuk informasi yang dapat langsung diakses melalui berbagai media informasi.

Dalam undang-undang no 44 tahun 2008 dijelaskan bahwa pornografi dalam bentuk apapun baik gambar, tulisan dalam media apapun yang mampu merangsang gairah

seksual maka dianggap melanggar undang-undang yang telah ditetapkan tersebut. Artinya sebuah media yang sebagai jembatan informasi masyarakat seharusnya menyaring informasi-informasi sehingga tidak memunculkan unsur-unsur pornografi karena telah dilarang apa lagi informasi ini mengarah pada dunia pendidikan (Anik Imawati Nur Rohimah, 2008). Karena sesuai dengan Undang-Undang Nomor 23 Tahun 2002 tentang perlindungan anak yaitu yang menjamin pertumbuhan dan perkembangan anak baik fisik, mental spiritual, maupun sosial anak (UU Tentang Pelindungan Terhadap Anak yang Menjadi Korban atau Pelaku Pornografi, 2008).

SDN 02 Turen terdapat media pembelajaran *elearning* yang berbasis cms yang sangat memungkinkan tersisip kata-kata yang dianggap porno dalam proses belajar mengajar yang disajikan oleh guru di SDN 02 Turen. Menurut Siti Rukiyati, S.Pd selaku Kepala Sekolah SDN 02 Turen menerangkan bahwa proses belajar mengajar yang digunakan oleh SDN 02 Turen adalah aplikasi *elearning* berbasis CMS *Wordpress* yang masih belum terdapat penyaring kata yang dapat menyaring kata yang dianggap porno, jika tersisip kata yang buruk mengakibatkan pengaruh buruk terhadap mental, kepribadian dan tata bahasa yang kurang sopan yang akan ditirukan oleh anak sekolah dasar. Untuk mengantisipasi tersisipnya kata-kata yang dianggap porno maka dirancang sebuah *filtering* kata yang dianggap porno pada aplikasi *elearning* SDN 02 Turen yang belum terdapat *filtering* kata untuk menyaring kata-kata yang dianggap porno agar informasi yang disampaikan lebih berkualitas dan dapat menjadi acuan dalam belajar siswa di SDN 02 Turen.

## 2. Tinjauan Pustaka

*Text mining* memiliki definisi menambang data yang berupa *text* dimana sumber data biasanya didapatkan dari dokumen dan tujuannya adalah mencari kata-kata yang mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antara dokumen (Milkha Harlina Ch, 2011).

*Text mining* menggunakan *natural language processing* untuk memasukkan struktur kedalam kumpulan teks. *Natural*

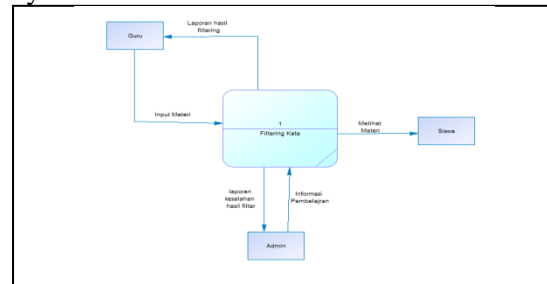
*Language Processing* (NLP) adalah komponen penting dari *text mining* dan merupakan sub-bidang dari *artificial intelligence* dan komputasional linguistik. NLP berupaya memecahkan masalah untuk memahami bahasa alami manusia, dengan segala aturan gramatika dan semantiknya, dan mengubah bahasa tersebut menjadi representasi formal yang dapat diproses oleh komputer (Moh.Sulhan, 2014).

## 3. Pembahasan

*Filtering* kata yang bertujuan untuk menyaring kata-kata yang dianggap porno yang memungkinkan terselip pada aplikasi *elearning* di SDN 02 Turen yang berdampak buruk pada anak-anak sekolah dasar yang dapat mempengaruhi mental dan tata bahasa yang kurang sopan, *filtering* ini juga menggunakan metode *boyer moore* yang memungkinkan pencarian kata yang dianggap porno dengan tepat.

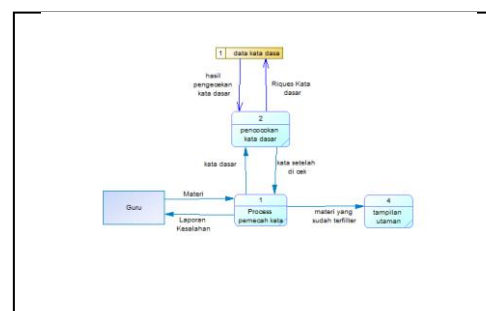
### 3.1 Desain Sistem

*Context Diagram* dalam DFD bisa disebut sebagai DFD level-0. Di dalam *context diagram* terdapat bagian-bagian penting yaitu, beberapa entitas, proses dan arus data. Berikut adalah *context diagram*-nya :



Gambar 1 Context Diagram

Diagram ini menjelaskan alur program secara umum dan *user* yang digunakan

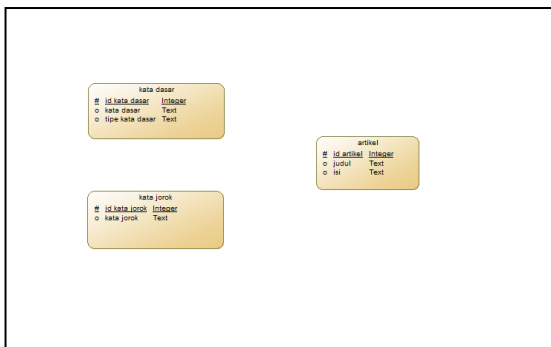


Gambar 2 DFD Level 1

Pada DFD level 1 ini menjelaskan tentang keseluruhan proses dibagi ke proses dalam level-level proses selanjutnya, terdapat 2 proses pada DFD level 1 ini yaitu

Proses yang pertama adalah guru memasukan data materi kedalam sistem filtering kemudian materi masuk kedalam proses pemecahan kata yang akan memecah menjadi kata dasar.

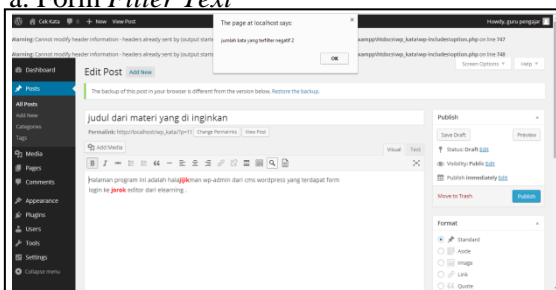
Proses yang kedua adalah proses input *file* ini adalah guru menginputkan sebuah *file* yang akan dipublikasikan di web *elearning*, pertama adalah *file* di inputkan lalu di tampung sementara di master input *file* lalu disimpan, setelah itu *file* akan di cek di *filtering* kata yang berfungsi mengecek kata-kata yang berada dalam *file* tersebut, setelah selesai maka *file* akan ditampilkan dihalaman utama *elearning* dan dapat didownload oleh siswa.



Gambar 3 Conceptual Data Model (CDM)

ERD di atas menjelaskan tentang tabel *database* yang digunakan dalam *filtering* kata, tidak adanya relasi antar tabel yang saling menghubungkan karena setiap tabel berdiri sendiri dan tidak saling berhubungan.

a. Form Filter Text



Gambar 4 Filter Kata

Pada halaman ini pemfilteran kata yang dianggap porno dengan memberi warna merah pada input kata.

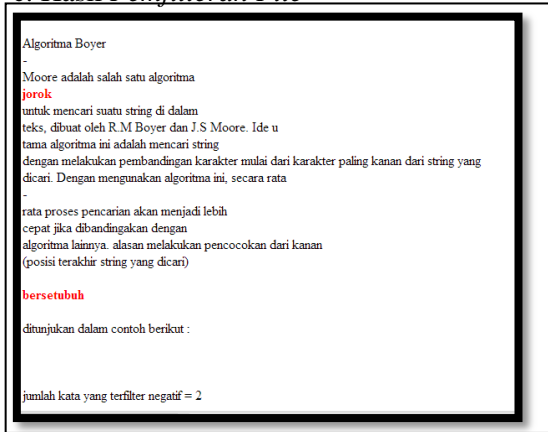
b. Form Upload File



Gambar 5 Form Upload File

Pada halaman ini mengupload data *file* berbentuk pdf.

c. Hasil Pemfilteran File



Gambar 6 Hasil Filtering File

Form ini adalah form *filter file* yang sudah di lengkapi dengan pendeteksi kata porno dalam *file* jika di dalam *file* terdapat kata porno maka akan terdeteksi isi sebagai kata porno.

Berikut *script* untuk pemfilteran kata :

```

<?php
include "/koneksi.php";
// ambil daftar kata jorok
$katajorok = array();
$kata = mysql_query("select *
from filter_kata");
while($jorok =
mysql_fetch_array($kata)){
    $katajorok[] = $jorok;}
$BM = new BM();
$var = $_POST['text'];
$hasil = $var;
//$hasil = $var;
  
```

```

foreach($katajorok as $pattern){
    $hasil = $BM-
>BMReplace($pattern['kata_jorok']
,'<b style="color:red"
>'.$pattern['kata_jorok'].'</b>',
$hasil);
    $judul = $BM-
>BMReplace($pattern['kata_jorok']
,'<b style="color:red"
>'.$pattern['kata_jorok'].'</b>',
$judul);
    $cek = $BM-
>BMCek($pattern['kata_jorok'],$ha
sil);
    if($cek==1){
        $jmlkt++;
    }
}
echo
json_encode(array("text"=>n12br($
hasil), "count"=>$jmlkt));
?>

```

#### 4. Kesimpulan

Setelah dilakukan Uji Coba dengan beberapa sampel data *input* dengan menggunakan metode pengujian pada beberapa mata pelajaran yang sudah ditentukan sebelumnya didapatkan hasil bahwasanya penggunaan metode kombinasi yang memperhatikan spesifikasi data *input* dalam hal butuh tidaknya metode *stemming* sebagai *preprocessing* disisipkan pada proses *filtering* dapat dikatakan lebih baik dibandingkan dengan menyisipkan metode *stemming* untuk semua *filtering* atau bahkan tidak menyisipkan metode *stemming* tersebut pada semua proses *filtering*. Dan untuk metode pencarian yang digunakan adalah *boyer moore* yang memiliki kemampuan pencarian terbaik dan tercepat. Hal ini ditunjukkan dari hasil beberapa pengujian bahwasanya penggunaan metode kombinasi pada *False Negative* = 0, artinya semua kata porno dapat terdeteksi dengan baik, berbeda dengan metode yang lain yang masih ada beberapa kata porno yang tidak terdeteksi. Hal ini terjadi karena kata yang dianggap porno bisa berupa kata berimbuhan, kata dasar, dan kata frase, dimana pengecekan kata pada kata porno yang berupa imbuhan tidak memerlukan proses *stemming*, sedangkan pengecekan kata porno yang berupa kata dasar dan kata frase

membutuhkan proses *stemming*. *Filtering* kata dengan metode *boyer moore* terbukti dapat melakukan pencarian secara akurat dan cepat dalam melakukan pemfilteran kata yang dianggap porno dan dapat memenuhi kriteria pencarian dan dapat melakukan pendeteksian kata yang dianggap porno.

#### 5. Saran

Pada sistem ini telah membantu dalam pemfilteran kata yang dianggap porno, namun dalam pembuatan program masih jauh dari sempurna. Dalam sistem ini terdapat kekurangan diantaranya yaitu :

1. Untuk mengembangkan sistem selanjutnya di harapkan dapat lebih akurat dan efisien dalam penggunaan metode.
2. Untuk pemfilteran agar dapat di kembangkan lebih lanjut dengan metode pemfilteran gambar yang dianggap porno.

#### 3. Daftar Pustaka

- \_\_\_\_\_. Metode *Boyer Moore*.  
[http://id.wikipedia.org/wiki/Algoritma\\_Boyer-Moore](http://id.wikipedia.org/wiki/Algoritma_Boyer-Moore)  
 Diunduh : 10 Maret 2014  
<http://informatika.stei.itb.ac.id/~rinaldi.munir/Stmik/.../MakalahStmik03.pdf>  
 Diunduh : 10 Maret 2014  
<http://prezi.com/.../copy-of-penerapan-algoritma-boyer-moore-untuk-aplikasi>.  
 Diunduh : 10 Maret 2014  
<http://pelita-informatika.com/berkas/jurnal/4316.pdf>  
 Diunduh : 12 Maret 2014  
<http://www.lbh-apik.or.id/uu-pornografi.htm>  
 .Diunduh : 14 Maret 2014.  
<http://www.kemenag.go.id/file/dokumen/442008.pdf>  
 Diunduh :14 Maret 2014  
 Kusri, 2006. *Optimasi Query Untuk Pencarian Data dengan Subset Query*, Bandung.  
 Diunduh : 13 April 2014  
 Chan, Immanuel, 2008. *Oracle Database Performance Tuning Guide*, 10g Release 2 (10.2),Redwood City, CA, pp. 379-403.

Sumathi, S. dkk, 2007. *Fundamentals of Relational Database Management Systems*, SpringerVerlag, Berlin Heidelberg. *MySQL 5.1 Reference Manual*.

<http://dev.mysql.com/doc/refman/5.1/en/partitioning.html>

Diakses pada [20 Oktober 2013, 09:00]

Setiawan, M.A, 2004. *Optimasi SQL Query untuk Informasi Retrieval pada Aplikasi Berbasis Web*, *Proceedings Seminar Nasional Aplikasi Teknologi Informasi UII*, Yogyakarta.

