Development of A Content Grabbing using php curl to Read News Online

Syahtriatna Djusar1*, Zamzami1, Elvira Asril1, Jeffri Supriatna2

1Fakultas Ilmu Komputer, Universitas Lancang Kuning, Pekanbaru Riau, Indonesia
2Swastika Universal, Pekanbaru Riau, Indonesia

eet@unilak.ac.id, zamzami@unilak.ac.id, elvira@unilak.ac.id, jeffrisupriatna@gmail.com

*Corresponding Author
Received: 10 October 2016, Accepted: 4 November 2016
Published online: 14 February 2017

Abstract: News is a form of public demand for information. To access many sources of news, then the reader must be typed and open many sites. To facilitate the reader, we need a function cURL (Client URL) in a web application, which is able to scan the data from or to a server, without interaction from each user. The workings of this function is to open a web page, scan it, and printing again. The difference of this study with previous research, Grabbing Content Techniques is applied in this system, namely by taking content from other web then grabbing results stored in the database, this is done in order to keep it safe from editing content on the source content. In contrast, if the result of grabbing instantly displayed on the web without a database, if the content is deleted or edited by the source of the content, the content on the page we also erased or edited. Before the content displayed on web pages, first in the selection by admin, so that there are no errors in the content. In the end, will produce a news portal, which can access the latest news from multiple sources with a single site only.

Keywords: cURL; Grabbing Content; Application; Read News; PHP.

1. Introduction

Indispensable information needs, wherever and whenever. There is now available a variety of information provider's portal. Many news sites are providing news on the latest, but different news sources, of course depending news content. People who have a curiosity with the latest news, should get the news from various sources. To access many news sources, will make the news reader was complicated because they have typed and open many sites. To that made the news portal that can access the latest news from multiple sources with a single site only. It has been easier for readers to access the latest news, the study.

The Structure

a. How to develop the CURL functions to collect a news content using a database within a site, so the news content will be safe from news source site changes.

b. This study is required to facilitate online news seekers, to get the news in one site collection. So the ease and save time newsreader

c. To gather news into a website, use function CURL. Curl is a program script function that can scan data from or to a server without the interaction of each user. This script will open a web page, scan it, take it, and redisplay the contents of a web page on a site that had been prepared.
2. Related Works

In the Teknologi Informasi journal Vol 1, No. 5, May 2015, CRAWL SOCMED Search Data by Using PHP CURL (Case Study: ISC PT. TELKOM Indonesia), Sudjoko said, Software development aims to generate and complete the customer profile based on a specific reference. Reference data is the email address of the customer who reported at the time of registration Speedy subscription. These email address is used to discover Speedy customer profile by taking a customer profile on the Facebook social networking media. Facebook social networking chosen for reasons of popularity compared to other social media. In software development, PT. Telkom Indonesia determining the php cURL library to grab Facebook profile. With socmed crawl method using php cURL, destination web page can be instantiated directly. At the time of the software (script) is executed, Speedy customer profile data will automatically be updated. In one execution, the software can update the customer profile in accordance with the number of records the email address that is exported. In the end, update customer profiles can help management to select potential customers and management can provide continuous promotions to them as a customer profile can describe the customer needs.

Yusuf, in a study published in the journal Teknik Informatika Vol 8 No. 1 In 2015, it was said that the search application products in electronic trading involving consumers, manufacturers, service providers and intermediary traders using computer networks, namely the Internet. By using the technique of retrieving data on other sites automatically, this application can locate and notify users via e-mail (electronic mail) whenever a product that is sought. CURL is a command function for transferring data to and from the server through the URL with various protocols using a technique that divides a series of parsing the input. With this application, users can search for a variety of products without having to open multiple trading sites and connect to the Internet for 24 hours, simply access this application and this application is going to do. With this application as well, users no longer need to search for products in a wide range of trading sites one by one.

3. Material & Methodology

3.1. Data

In this study, the data obtained from the site content for content in an existing news site, in this study the sites will be taken of data detik.com news is, tribunnews.com, okezone.com, bola.com. Furthermore, the data present in the news this news will be read using the cURL functions. Furthermore, the news content will be stored in a site riau24.com.

3.2. Design System

3.2.1 Use case diagram

![Use Case Diagram](image-url)
3.2.2 Activity Diagram

Activity diagram shows the activity of some parts of the structure of the organizations involved in the system.

a. Activity Diagram Admin

Here is an activity diagram modeling Admin on the application system:

![Activity Diagram Admin](image1)

Figure 2. Activity Diagram Admin

b. Activity Diagram User

![Activity Diagram User](image2)

Figure 3. Activity Diagram User
4. Results and Discussion

4.1. Result

Data News Table

<table>
<thead>
<tr>
<th>No</th>
<th>File Name</th>
<th>Type</th>
<th>Number</th>
<th>Keterangan</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>news_id</td>
<td>Varchar</td>
<td>10</td>
<td>Primary Key</td>
</tr>
<tr>
<td>2</td>
<td>news_link</td>
<td>Varchar</td>
<td>500</td>
<td>Link Berita</td>
</tr>
<tr>
<td>3</td>
<td>news_title</td>
<td>Varchar</td>
<td>500</td>
<td>Judul Berita</td>
</tr>
<tr>
<td>4</td>
<td>news_pict</td>
<td>Varchar</td>
<td>500</td>
<td>Gambar berita</td>
</tr>
<tr>
<td>5</td>
<td>news_titlepict</td>
<td>Varchar</td>
<td>500</td>
<td>Ket. Gambar berita</td>
</tr>
<tr>
<td>6</td>
<td>news_topic</td>
<td>Varchar</td>
<td>1000</td>
<td>Isi / konten berita</td>
</tr>
<tr>
<td>7</td>
<td>news_cat</td>
<td>Varchar</td>
<td>500</td>
<td>Kategori berita</td>
</tr>
<tr>
<td>8</td>
<td>news_source</td>
<td>Varchar</td>
<td>100</td>
<td>Sumber berita</td>
</tr>
<tr>
<td>9</td>
<td>news_date</td>
<td>Datetime</td>
<td></td>
<td>Tanggal update berita</td>
</tr>
<tr>
<td>10</td>
<td>views_date</td>
<td>Date</td>
<td></td>
<td>Tanggal berita dilihat</td>
</tr>
<tr>
<td>11</td>
<td>news_update</td>
<td>varchar</td>
<td>10</td>
<td>Status berita</td>
</tr>
<tr>
<td>12</td>
<td>news_headline</td>
<td>Varchar</td>
<td>5</td>
<td>Status headline</td>
</tr>
<tr>
<td>13</td>
<td>news_viewed</td>
<td>Int</td>
<td></td>
<td>Jumlah pembaca</td>
</tr>
<tr>
<td>14</td>
<td>news_comment</td>
<td>Int</td>
<td></td>
<td>Komentar</td>
</tr>
</tbody>
</table>

Table 1. Tabel news

Source Code

Source code for grabbing the content is:

```php
 Include"../config/koneksi.php";

 function bacaHTML($url){
  $data = curl_init();
  curl_setopt($data, CURLOPT_RETURNTRANSFER, 1);
  curl_setopt($data, CURLOPT_URL, $url);
  $hasil = curl_exec($data);
  curl_close($data);
  return $hasil;
 }
```

Explanation of Source Code

- Include"../config/koneksi.php"; : Make a connection to the database.
- function bacaHTML($url) : create CURL functions that can be used in PHP.
- $data = curl_init() ; : Initialize a cURL session or create a resource.
- curl_setopt($data, CURLOPT_RETURNTRANSFER, 1); and curl_setopt($data, CURLOPT_URL, $url); : To set options and option RETURNTRANSF    ER URL.
- $hasil = curl_exec($data); : Get the page URL and give it to the variable $result.
Before doing admin grabbing content first select menu news News to display data and button for grabbing news. Then the admin can perform grabbing news by selecting one button is provided, as shown below

![Figure 4. Menu Button Grabbing News and Content](image)

After selecting the button grabbing content then the data will be refreshed and displayed on tables news. Then the admin can do news updates by clicking the icon update.

![Figure 5. Result Grabbing Content](image)

### 4.2. Discussion

In this study, the function cURL (Client URL) in the form of a PHP script placed in sites riau24.com provided. Riau24.com this application on the site, should be managed by an admin in search for news on any news site. This application menggunakan multiple tables stored in a database. The use of the database on the application, can solve the existing problems in similar applications that
Development of A Content Grabbing using php curl to Read News Online

already exist. With this database, if this is news in the site removed the original site, it will not cause an error on the site riau24.com, because the removed content has been stored in riau24.com site.

5. Conclusion
   a. The use cURL function can be used to gather news into a site of beberpa existing news sites.
   b. Use of the database on these applications can help overcome riau2.com error on the site, if the news content to the site source is removed.
   c. Is expected in subsequent studies, found an intelligent logic that can perform reading of the necessary search and news automatically, without admin management.

References