

# CHANNEL BOT TELEGRAM SIMLITABMAS SECARA OTOMATIS DENGAN MEMANFAATKAN WEB CRAWLER

Muhammad Ulil Fahri

Akademi Manajemen Komputer dan Informatika (AMKI) Ketapang, Indonesia

[Muhammad.fahri001@binus.ac.id](mailto:Muhammad.fahri001@binus.ac.id)

## Article Info

### Article history:

Received, 19/11/2020

Revised, 24/11/2020

Accepted, 31/11/2020

### Kata Kunci:

Web Crawler  
Telegram Bot

## ABSTRAK

Saat ini masih terdapat dosen yang sering terlambat dalam mendapatkan informasi terbaru dari simlitabmas. Sehingga banyak dosen di Indonesia yang tidak mendapat dana hibah. Dikarenakan kurang updatenya dari dosen. Penelitian bertujuan memberikan kemudahan untuk dosen di Indonesia dalam mendapatkan informasi dari website simlitabmas. Metode yang dilakukan dalam penelitian ini adalah dengan memanfaatkan web crawler. Pengujian dan implementasi telah dilakukan pada penelitian ini, dan berjalan dengan sangat baik. Dengan memanfaatkan web crawler dalam mendapatkan informasi tentang web simlitabmas secara berkala, ketika ada update dari web simlitabmas secara otomatis memberikan notifikasi telegram ke dosen.

### Keywords:

Web Crawler  
Telegram Bot

## ABSTRACT

Currently there are still lecturers who are often late in getting the latest information from simlitabmas. So that many lecturers in Indonesia do not receive grant funds. Due to the lack of updates from the lecturers. The research aims to make it easy for lecturers in Indonesia to get information from the Simlitabmas website. The method used in this research is to use a web crawler. Testing and implementation have been carried out in this study, and are going very well. By utilizing the web crawler to get information about the simlitabmas web regularly, when there is an update from the simlitabmas web, it automatically provides telegram notifications to the lecturer.

*This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.*



### Penulis Korespondensi:

Jojo,

Program Studi Informatika,

Universitas Komputer Indonesia,

Email: [jojo@gmail.com](mailto:jojo@gmail.com)

## 1 PENDAHULUAN

Masih banyak dosen di perguruan tinggi di Indonesia yang sering terlambat mendapatkan informasi tentang dana hibah penelitian dan pengabdian. Hal ini dikarenakan salah satu faktor malasnya dosen untuk membuka website simlitabmas secara berkala. Kementerian RisetDikti membangun sistem yang disebut Sistem Informasi Penelitian dan Pengabdian Kepada Masyarakat atau disingkat simlitabmas yang dapat digunakan oleh semua dosen di Indonesia untuk mengajukan proposal penelitian dan pengabdian secara online [1]. Pada penelitian ini bertujuan membuat sebuah channel bot telegram dimana datanya secara otomatis di dapat dari website simlitabmas dengan memanfaatkan web crawler. Dengan harapan, setiap ada informasi terbaru dari simlitabmas langsung secara otomatis memberikan notifikasi ke setiap dosen Indonesia via channel bot telegram.

keyword query based focused crawler in which the web pages are crawled fast [2]. menurut pendapat diatas web crawler berfokus pada kueri dengan kata kunci pada website. Web crawler dengan fungsi merayap secara otomatis pada halaman web dengan sistem bot. Degan mengambil atribut html yang dipilih sebagai hasil

output yang diinginkan. age relevance calculator uses DOM structure of relevant web pages. This method takes the webpage along with the Meta tags into consideration while deciding the relevancy of a webpage [2]. Masih pada penelitian tersebut, dengan mencari website yang relevansi dengan tujuan yang dapat mengambil data DOM dari website. Memanfaatkan tag meta html di dalam website. Teknik ekstraksi informasi merupakan sebagian dari teknik terkini yang memecahkan konten serta mempersembangkannya kembali dalam bentuk yang lebih kecil [3]. setiap dari halaman website memiliki suatu tag html yang dapat di olah dan di ekstraks menjadi bentuk – bentuk yang lebih kecil. DOM adalah proses yang berjalan dalam aplikasi web scraping dengan metode HTML DOM adalah proses scraping, tokenisasi, cleaning, dan lowercased, semua proses tersebut berjalan secara otomatis [4] . Dengan kata lain, semua website dapat dilakukan crawler untuk mengambil data teks website dengan tujuan tertentu. Pada penelitian ini berfokus pada tujuan yang dapat mengambil informasi update secara otomatis. Bot adalah sebuah program komputer yang dirancang untuk menstimulasikan percakapan intelektual dengan satu atau lebih manusia baik secara audio maupun teks [5]. Pada penelitian ini, proses pengambilan data di lakukan secara otomatis lewat bot. Seakan – akan bot dapat membaca isi dari website simlitabmas yang selanjutnya di terjemahkan ke dalam bahasa mesin. Sehingga ketika ada update terbaru dari simlitabmas, maka web crawler langsung berjalan berdasarkan perintah perayapan.

Web crawler adalah program komputer atau perangkat lunak yang menjelajahi halaman halaman internet dengan otomatis sesuai dengan kata kunci yang dimasukkan [6]. Halaman website dapat dijelajahi secara otomatis sehingga dapat melakukan cloning isi dari website. aplikasi web crawler dapat mengambil link yang terdapat pada sitemap dan mengambil isi metadata yang terdapat pada link [7]. Dengan memanfaatkan web crawler penelitian ini dapat membuat sebuah implementasi, dimana dapat mengambil teks dan link yang mengarah langsung ke website yang di crawler. Sitemap adalah salah satu alat bantu untuk mempermudah dalam pengenalan peta situs pada sebuah website [8]. Dalam hal ini target website yang menjadi crawler adalah website simlitabmas. Dari hasil web crawler tersebut kemudian diolah ke dalam sebuah algoritma yang hasil akhirnya dapat mengirim notifikasi secara otomatis ke channel bot telegram. Algoritma adalah urutan langkah-langkah logis penyelesaian masalah yang disusun secara sistematis dan logis [9]. Pembuatan langkah logis berdasarkan aturan sistematis yang dapat memberikan perintah ke server secara langsung.

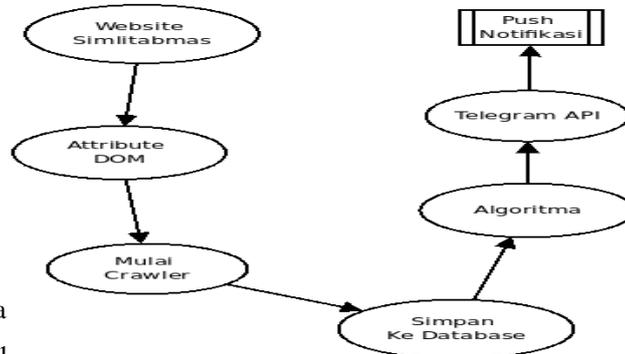
Telegram adalah sebuah aplikasi layanan pengirim pesan instan multiplatform berbasis awan yang bersifat gratis dan nirlaba [10]. kelebihan dari Telegram ini adalah adanya landasan untuk menggunakan Application Programming Interface(API) untuk masyarakat luas. Salah satu API yang disediakan adalah fitur bot. Bot Telegram adalah bot yang saat ini mulai populer dipergunakan [11]. Dengan memanfaatkan API dari Telegram, maka dari hasil crawler website dapat dialihkan menggunakan algoritma yang dimana hasil akhirnya membuat sebuah notifikasi yang cepat. Hasil notifikasi juga berdasarkan md5 sehingga hanya hasil crawler yang terbaru saja yang diberikan akses API. Proses mencari berita secara manual memerlukan waktu dan usaha yang lebih dikarenakan harus melakukan pencarian satu demi satu [12]. Dari pendapat tersebut, dapat di studi kasuskan pada website simlitabmas dimana masih terdapat dosen yang belum mengetahui informasi terbaru. Ketinggalan atas informasi tentang simlitabmas membuat dosen tidak mengetahui perkembangan penelitian dan pengabdian tentang dana hibah yang diberikan oleh pemerintah. Dengan adanya kombinasi yang menarik antara web crawler untuk mengambil data dimana hasil akhirnya adalah berupa informasi terbaru yang dikirim secara API menggunakan channel telegram. Format pertukaran data selain XML yaitu JSON (Javascript Object Notation) yang merupakan bagian dari JavaScript sehingga parsing dilakukan oleh JavaScript tersebut dan memiliki sintaks yang lebih sederhana daripada XML. Secara spesifikasi JSON lebih baik performansinya daripada XML [13] .

Aplikasi bot telegram dapat digunakan sebagai media informasi penelitian dan HaKI di LP2M [14] . Penelitian tersebut menunjukkan pemanfaat bot telegram untuk memberikan informasi Haki di LP2M berhasil dilakukan dengan baik. Informasi terbaru tentang Lp2m dapat dirasakan secara langsung oleh pengguna sehingga tidak adanya keterlambatan dalam penyebaran informasi. Pengguna hanya membutuhkan waktu kurang dari 1 menit untuk mendapatkan data penelitian atau HaKI yang diinginkan [14] . waktu yang diperlukan kurang dari 1 menit data dapat terkirim dalam bentuk notifikasi ke pengguna. Pada kasus penelitian ini setiap ada update dari simlitabmas maka secara otomatis data terkirim.

Pada penelitian ini, lebih ke fokus dalam bidang bot. Bot adalah program komputer yang melakukan pekerjaan tertentu secara otomatis [15]. Dimana dengan sistem ini akan secara otomatis mengambil isi data dari simlitabmas. Isi data informasi terbaru simlitabmas memanfaatkan web crawler sehingga bot secara otomatis melakukan perayapan terhadap website simlitabmas. Hasil dari perayapan diteruskan ke API telegram yang dimana hasil akhirnya adalah sebuah notifikasi secara cepat dan akurat dalam memberikan informasi tentang simlitabmas. Pada penelitian lainnya hanya berfokus pada satu bidang saja yaitu web crawler, sedangkan pada penelitian langsung membuat implementasi dari hasil crawler. Hasil dari crawler website yang dapat secara langsung dirasakan oleh dosen di Indonesia.

## 2 METODE PENELITIAN

Agar lebih memudahkan dalam membuat penelitian ini, penulis menyajikan kerangka pikir penelitian seperti gambar berikut



Gambar 1. Kerangka

Pada gambar 1 merupakan kerangka pikir yang peneliti gunakan. Agar lebih mudah dipahami, dapat dijabarkan seperti langkah berikut :

1. Memilih website menjadi target
2. Proses seleksi attribute Html yang nantinya akan digunakan sebagai kata kunci web crawler
3. Mulai melakukan web crawler, dengan bantuan cronjob dari server.
4. Proses data crawler yang berupa teks dan link mulai dimasukkan ke dalam sistem database. Pada sistem database mengambil kode unik yang dimana hasilnya di convert menjadi hash md5
5. pembuatan algoritma yang saling terhubung sebagai jembatan web crawler dan API Telegram
6. memanfaatkan API dari telegram.
7. Mengirim push notifikasi ke channel telegram.

## 3 HASIL DAN ANALISIS

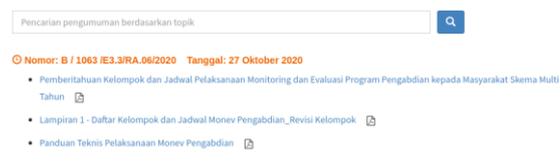
ada penelitian ini melakukan melakukan web crawler, berdasarkan tabel berikut :

Tabel 1. Target Web Crawler

No	Nama Website Target	Link	Target
1	Simlitabmas	<a href="http://simlitabmas.ristekdikti.go.id/">http://simlitabmas.ristekdikti.go.id/</a>	Utama
2	Ristek Brin	<a href="https://www.ristekbrin.go.id/category/pengumuman/">https://www.ristekbrin.go.id/category/pengumuman/</a>	Tambahan

Pada Tabel 1 merupakan website yang menjadi target sasaran web crawler. Ada dua target website yang utama dan tambahan. Dalam hal ini utama adalah website yang menjadi sasaran utama bot dalam mengambil data. Sedangkan website tambahan mengambil informasi pendukung yang dapat membantu web crawler dalam memberikan informasi lainnya.

Pemilihan Attribute HTML berdasarkan Atribut yang diambil. Pada alamat simlitabmas dapat dilihat pada gambar berikut :



Gambar 1. Tampilan

Web Simlitabmas

Pada gambar 1 merupakan pengambilan data attribute Html Dom dari website simlitabmas. Adapun Atribut yang diambil seperti berikut :

`“gvPengumuman_gvInternal_0_lbJudulFile_0”`

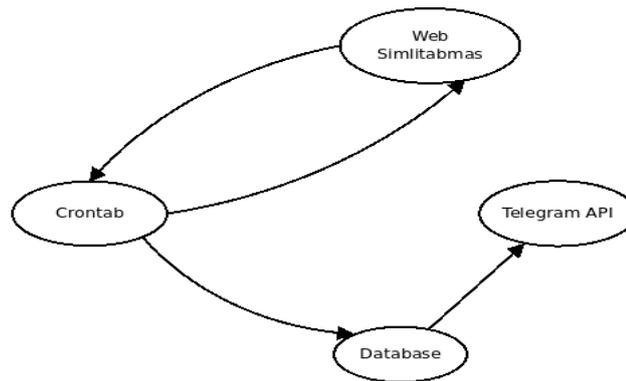
Dengan memanfaatkan attribute html yang telah di dapat selanjutnya di convert menjadi bentuk teks dan link. Agar proses dapat berjalan dengan lancar diperlukan server yang kuat. Dalam hal ini memanfaatkan VPS dengan spesifikasi seperti berikut :

Tabel 2. Spesifikasi Server

No	Spesifikasi	Keterangan
1	Jenis Server	VPS ( Virtual Private Server )
2	Sistem Operasi	Centos 7
3	RAM	4 GB

Server yang diperlukan harus memiliki spesifikasi yang kuat. Hal ini dikarenakan server terus melakukan perayapan pada website simlitabmas. Secara terus menerus dengan melakukan penjadwalan yang otomatis maka server dapat mengambil data dari web simlitabmas dalam waktu setiap jam. Setiap satu jam server mengunjungi website simlitabmas untuk melakukan cek jika terdapat update terbaru.

Berikut adalah gambaran secara umum proses pengambilan data dengan memanfaatkan crontab di server centos 7



Gambar 2. Proses

Fungsi dari crontab adalah melakukan pengecekan secara berkala ke website simlitabmas. Berikut jadwal dari crontab dengan dibuat setiap 3 jam melakukan

Tabel 2. Jadwal Crawler

No	Jam	Waktu Crawler
1	06.00	10 Detik
2	09.00	11 Detik
3	12.00	9 Detik
4	15.00	11 Detik
5	18.00	12 Detik
6	21.00	10 Detik
7	00.00	10 Detik
8	03.00	11 Detik

Pada tabel 1 merupakan jadwal crawler dalam mengambil data website dari simlitabmas. Sebanyak 8 kali proses dalam satu hari web crawler mengunjungi secara otomatis web simlitabmas untuk melakukan cek perubahan update dari informasi. Ketika ada update dari website simlitabmas crontab mengirim dari ke dalam sistem database yang akhirnya dikirim kembali ke API Telegram. Proses web crawler yang sudah berjalan dapat mengambil teks. Berikut data yang diambil.

Tabel 3. Struktur Database

No	Name	Type	Extra
1	Id	Int	Auto Increment
2	Uuid	Varchar (md5)	-
3	Judul	Text	-
4	Nomor	Varchar	-
5	Tanggal	Date	-

Pada tabel 3 merupakan data penyimpanan yang digunakan untuk menampung data yang didapat dari website simlitabmas. Uji coba dilakukan dari tanggal 1-5 November 2020. berikut rangkuman uji coba yang dilakukan

Tabel 4. Pengujian Crawler

No	Tanggal	Update Simlitabmas	Notifikasi
1	1 November 2020	Ada	Ada
2	2 November 2020	Ada	Ada
3	3 November 2020	Tidak	Tidak
4	4 November 2020	Ada	Ada
5	5 November 2020	Tidak	Tidak

Pada tabel 4 merupakan pengujian sistem yang dilakukan. Dapat disimpulkan program crawler dapat berjalan dengan baik. Ketika ada update informasi dari simlitabmas, secara otomatis sistem mengirim ke notifikasi telegram. Berikut adalah tampilan web simlitabmas pada tanggal 4 November

Nomor: B/1480/E5.1.2/KI.01.02/2020 Tanggal: 04 Nopember 2020

- Pengumuman Nominasi Calon Pemenang Insentif HKI Produktif
- Technical Meeting Presentasi HKI Produktif

Gambar 3. Update

Web Simlitabmas

Pada gambar 3 merupakan gambaran update dari website simlitabmas. Setelah ada update dari website simlitabmas secara otomatis crontab mengirim ke Api Telegram yang mengirim notifikasi, seperti pada gambar berikut



Gambar 4. Notifikasi Telegram

Gambar 4 menunjukkan, pada saat jam 21.00 sesuai jadwal crawler. Ada perubahan informasi pada website simlitabmas. Langsung secara otomatis mengirim notifikasi ke telegram dengan mengirim variabel nomor, tanggal, dan judul. Pengujian sistem kedua dilakukan secara 10 hari. Pengujian ini dilakukan dengan harapan tidak terdapat bug ketika web crawler. Berikut adalah jadwal pengujian yang peneliti lakukan.

Tabel 5. Pengujian Ke-2

No	Tanggal (November 2020)	Update Simlitabmas	Notifikasi
1	5	Tidak	Tidak
2	6	Ada	Ada
3	7	Tidak	Tidak
4	8	Tidak	Tidak
5	9	Ada	Ada
6	10	Tidak	Tidak
7	11	Tidak	Tidak
8	12	Ada	Ada
9	13	Tidak	Tidak
10	14	Tidak	Tidak

Dari tabel 5 dapat dilihat sistem berjalan dengan sangat baik. Tidak terdapat delay keterlambatan informasi update dari simlitabmas. Ketika ada update dari simlitabmas secara otomatis juga notifikasi terkirim.

#### 4 KESIMPULAN

Dari hasil penelitian dapat disimpulkan channel bot telegram simlitabmas secara otomatis dengan memanfaatkan web crawler dapat dimanfaatkan sebagai sarana dosen dalam mendapatkan informasi secara cepat dan akurat. Peneliti mempersilahkan untuk dosen di seluruh indonesia untuk bergabung pada channel bot simlitabmas bisa melalui tautan berikut <https://t.me/simlitabmas>

## UCAPAN TERIMA KASIH

Terima kasih kepada pengelola website simlitabmas yang selalu memberikan informasi tentang penelitian dan pengabdian kepada masyarakat khususnya untuk dosen. Dengan adanya website simlitabmas merupakan suatu kesempatan dosen untuk meningkatkan karir di bidang pendidikan.

## REFERENSI

- [1] G. B. Putra, "RANCANG BANGUN SISTEM PENELITIAN DAN PENGABDIAN MASYARAKAT UNIVERSITAS BANGKA BELITUNG BERBASIS WEB SERVER," *J. Ecotipe (Electronic, Control, Telecommun. Information, Power Eng.*, vol. 4, no. 1, pp. 17–22, 2017.
- [2] M. Kumar, A. Bindal, R. Gautam, and R. Bhatia, "Keyword query based focused Web crawler," *Procedia Comput. Sci.*, vol. 125, pp. 584–590, 2018.
- [3] E. Mahdiyah, "SISTEM EKSTRAKSI KANDUNGAN URL, TITTLE, META TAG, HYPERLINK PADA HALAMAN WEB," *SEMIRATA 2015*, vol. 5, no. 1.
- [4] V. Mitra, H. Sujaini, and A. B. P. Negara, "Rancang bangun aplikasi web scraping untuk korpus paralel indonesia-inggris dengan metode html dom," *JUSTIN (Jurnal Sist. dan Teknol. Informasi)*, vol. 5, no. 1, pp. 36–41, 2017.
- [5] R. B. SANTOSO, "Rancang Bangun Prototype Chat Bot Customer Service System Berbasis Web." Tugas Akhir, Universitas Islam Negeri Sultan Syarif Kasim Riau, 2011.
- [6] I. Ramadhan and H. Sastramihardja, "Pemanfaatan Web Crawler Dalam Mengumpulkan Informasi Melalui Internet," *Konf. Nas. Sist. Inf. 2018*, pp. 8–9, 2018.
- [7] A. Halim, R. D. Nyoto, and N. Safriadi, "Perancangan Aplikasi Web Crawler untuk Menghasilkan Dokumen Teks pada Domain Tertentu," *JUSTIN (Jurnal Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 114–117, 2017.
- [8] H. Susianawati, A. B. Tjandrarini, and S. H. E. Wulandari, "Design of Web-Based Sales Information System at CV Gemilang Indonesia," *J. Sist. Inf. dan Komput. Akunt.*, vol. 6, no. 1, pp. 29–35, 2017.
- [9] A. Budiyanto, "Pengantar Algoritma dan Pemrograman," *Kuliah Berseri Ilmu Komput.*, 2003.
- [10] A. Tanjung, "SISTEM OTOMASI PENYIRAMAN TANAMAN BERBASIS TELEGRAM," *SIGMA Tek.*, vol. 2, no. 1, pp. 81–94, 2019.
- [11] L. N. Gunawan, J. Anjarwirawan, and A. Handojo, "Aplikasi Bot Telegram Untuk Media Informasi Perkuliahan Program Studi Informatika-Sistem Informasi Bisnis Universitas Kristen Petra," *J. Infra*, vol. 6, no. 1, pp. 134–139, 2018.
- [12] R. Parlika and A. Pratama, "Aplikasi Penampil Berita Realtime Berbasis Bot Telegram menggunakan API Web (APBR Versi 1.0)," in *Prosiding Seminar Nasional SANTIKA Ke-1 2019*, 2019, pp. 17–20.
- [13] Y. F. A. Wibowo, D. S. Kusumo, and A. A. Marta, "Optimasi Performansi Ajax dengan Menggunakan Json," in *Seminar Nasional Informatika (SEMNASIF)*, 2015, vol. 1, no. 1.
- [14] A. D. Mulyanto, "Pemanfaatan Bot Telegram Untuk Media Informasi Penelitian," *MATICS*, vol. 12, no. 1, pp. 49–54, 2020.
- [15] D. Utomo, M. Sholeh, and A. Avorizano, "Membangun Sistem Mobile Monitoring Keamanan Web Aplikasi Menggunakan Suricata dan Bot Telegram Channel," in *Prosiding Seminar Nasional Teknoka*, 2017, vol. 2, pp. I81–I87.