

# **PEMBANGUNAN APLIKASI BERITA BERBASIS WEB DENGAN PENERAPAN TEKNIK WEB SCRAPING**

## **TUGAS AKHIR**

Disusun sebagai salah satu syarat untuk kelulusan  
Program Strata I, Program Studi Teknik Informatika  
Universitas Pasundan Bandung

Oleh :

Mohamad Egi Rahayu  
NPM : 20.304.0045



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS PASUNDAN BANDUNG  
OKTOBER 2024**



**LEMBAR PENGESAHAN  
LAPORAN TUGAS AKHIR**

Telah diujikan dan dipertahankan dalam Sidang Sarjana Program Studi Teknik Informatika Universitas Pasundan Bandung, pada hari dan tanggal sidang sesuai berita acara sidang, tugas akhir dari :

Nama : Mohamad Egi Rahayu

Nrp. : 20.304.0045

Dengan judul :

**“PEMBANGUNAN APLIKASI BERITA BERBASIS WEB  
DENGAN PENERAPAN TEKNIK WEB SCRAPING”**

Bandung, 02 Oktober 2024

Menyetujui,

Pembimbing Utama



(Dr. Ayu Purbasari, ST., MT.)



## ABSTRAK

Seiring dengan perkembangan layanan situs berita, pembaca sering kali harus berpindah-pindah antar berbagai website untuk mendapatkan berita berkualitas. Hal ini dapat menyulitkan, terutama jika informasi yang dicari tersebar di berbagai platform. Oleh karena itu, diperlukan sebuah website khusus yang dapat menggabungkan beberapa situs berita menjadi satu.

Penelitian ini dilakukan dengan mengembangkan aplikasi berita berbasis web bernama HoyNews, yang dibangun menggunakan bahasa pemrograman PHP. Proses pengembangan mencakup perancangan antarmuka pengguna, pengumpulan berita secara otomatis dari berbagai situs, penyimpanan data ke dalam database, serta pengujian alpha testing menggunakan metode black box testing. Aplikasi ini dirancang untuk memudahkan pengguna mengakses berita terkini dari berbagai sumber tanpa harus mengunjungi setiap situs berita satu per satu.

Hasil dari penelitian tugas akhir ini yaitu aplikasi berita berbasis web dengan penerapan teknik web scraping. Dari penelitian ini aplikasi berita berbasis web menyediakan fitur yang memudahkan pengguna dalam mengakses berita terkini dari berbagai sumber, sehingga informasi dapat diperoleh dengan cepat dan efisien.

**Kata kunci:** Web Scraping, Aplikasi Berita, PHP, Alpha Testing, Black Box Testing

## ABSTRACT

Along with the development of news site services, readers often have to move between various websites to get quality news. This can be difficult, especially if the information sought is spread across various platforms. Therefore, a special website is needed that can combine several news sites into one.

This research was carried out by developing a web-based news application called HoyNews, which was built using the PHP programming language. The development process includes designing a user interface, automatically collecting news from various sites, storing data into a database, and alpha testing using the black box testing method. This application is designed to make it easier for users to access the latest news from various sources without having to visit each news site individually.

The results of this final project research are web-based news applications with the application of web scraping techniques. From this research, web-based news applications provide features that make it easier for users to access the latest news from various sources, so that information can be obtained quickly and efficiently.

**Keywords:** Web Scraping, News Application, PHP, Alpha Testing, Black Box Testing

## DAFTAR ISI

LEMBAR PENGESAHAN.....	
LEMBAR PERNYATAAN KEASLIAN TUGAS AKHIR .....	
ABSTRAK.....	i
ABSTRACT.....	ii
KATA PENGANTAR .....	iii
DAFTAR ISI.....	iii
DAFTAR TABEL.....	iv
DAFTAR GAMBAR .....	vi
BAB 1 PENDAHULUAN.....	1-1
1.1 Latar Belakang .....	1-1
1.2 Identifikasi Masalah .....	1-2
1.3 Tujuan Tugas Akhir.....	1-2
1.4 Lingkup Tugas Akhir .....	1-2
1.5 Metodologi Tugas Akhir .....	1-3
1.6 Sistematika Penulisan .....	1-5
BAB 2 LANDASAN TEORI DAN PENELITIAN TERDAHULU.....	2-1
2.1 Landasan Teori.....	2-1
2.1.1 Aplikasi .....	2-1
2.1.2 Berita .....	2-1
2.1.3 Website.....	2-1
2.1.4 Web Scraping .....	2-2
2.1.5 <i>Software Development Life Cycle Waterfall</i> .....	2-2
2.1.6 MySQL.....	2-3
2.1.7 PHP .....	2-3
2.1.8 <i>Creative Commons</i> .....	2-3
2.1.9 <i>Alpha Testing</i> .....	2-4
2.1.10 <i>Blackbox Testing</i> .....	2-4
2.2 Penelitian Terdahulu.....	2-4
BAB 3 SKEMA PENELITIAN.....	3-1
3.1 Alur Penyelesaian Tugas Akhir .....	3-1
3.2 Perumusan Masalah.....	3-4
3.2.1 Analisis Sebab Akibat .....	3-4
3.2.2 Solusi Masalah .....	3-5
3.3 Kerangka Berpikir Teoritis .....	3-5
3.3.1 Kerangka Berpikir Teoritis .....	3-5

3.3.2	Skema Analisis Teori.....	3-6
3.4	Planning Tugas Akhir.....	3-7
<b>BAB 4</b>	<b>ANALISIS DAN PERANCANGAN .....</b>	<b>4-1</b>
4.1	Analisis dan Perancangan .....	4-1
4.1.1	Model Analisis.....	4-1
4.1.2	User Requirement .....	4-1
4.1.3	Software Requirement .....	4-2
4.1.4	Use Case Diagram .....	4-2
4.1.5	Activity Diagram .....	4-3
4.1.6	Skenario Use Case & Prototype.....	4-5
4.1.7	Class Diagram.....	4-28
4.1.8	Perancangan Struktur Database .....	4-28
4.1.9	Mekanisme Kerja Sistem .....	4-30
<b>BAB 5</b>	<b>IMPLEMENTASI DAN PENGUJIAN .....</b>	<b>5-1</b>
5.1	Implementasi .....	5-1
5.1.1	Implementasi Prototype .....	5-1
5.1.2	Penulisan Kode.....	5-10
5.1.2	Struktur Folder.....	5-11
5.2	Pengujian.....	5-17
5.2.1	Test Skenario .....	5-17
5.2.2	Test Case .....	5-18
<b>BAB 6</b>	<b>KESIMPULAN DAN SARAN .....</b>	<b>6-1</b>
6.1	Kesimpulan.....	6-1
6.2	Saran.....	6-1
<b>DAFTAR PUSTAKA.....</b>		



# **BAB 1**

## **PENDAHULUAN**

Bab ini memberikan penjelasan umum mengenai tugas akhir yang penulis lakukan. Penjelasan tersebut meliputi latar belakang, identifikasi masalah, tujuan tugas akhir, lingkup tugas akhir, metodologi tugas akhir, dan sistematika penulisan.

### **1.1 Latar Belakang**

Dalam era digital yang semakin maju, kebutuhan akan informasi yang cepat, dan mudah diakses semakin meningkat. Website berita hadir sebagai solusi untuk memenuhi kebutuhan masyarakat akan informasi terkini, menggantikan media cetak tradisional yang mulai ditinggalkan. Perubahan perilaku pembaca yang cenderung mengakses berita melalui perangkat digital mendorong banyak media untuk beralih ke platform online. Selain itu, persaingan yang ketat di industri media mendorong pengembangan website berita yang tidak hanya menarik dan mudah diakses, tetapi juga kredibel dan berintegritas. Website berita berperan penting dalam mencerdaskan masyarakat dengan menyajikan informasi yang mendidik dan terpercaya.

Seiring dengan perkembangan layanan situs berita, pembaca sering kali harus berpindah-pindah antar berbagai website untuk mendapatkan berita yang berkualitas. Hal ini bisa menyulitkan, terutama jika informasi yang dicari tersebar di berbagai platform. Selain itu, tidak semua situs berita menyajikan informasi dengan standar yang sama, sehingga pembaca perlu selektif dalam memilih sumber yang dapat dipercaya. Situasi ini tidak hanya membuat pembaca harus mengalokasikan waktu lebih banyak, tetapi juga dapat menimbulkan rasa frustrasi karena proses pencarian informasi yang tidak efisien. Hal tersebut membuat tersitanya waktu dan merepotkan untuk membaca berita.

Selain itu, pengumpulan berita di banyak organisasi media masih dilakukan secara manual. Proses ini biasanya melibatkan wawancara langsung, observasi di lapangan, dan penelitian mendalam dari berbagai sumber informasi. Namun, dengan pesatnya perkembangan informasi di era digital, metode manual ini menghadapi berbagai tantangan. Proses manual memerlukan banyak waktu dan tenaga, terutama saat harus menyaring dan memverifikasi informasi dari berbagai sumber online. Volume data yang sangat besar dan terus berubah menyulitkan wartawan dan tim redaksi untuk mengikuti arus informasi yang meningkat. Keterlambatan dalam pengumpulan data juga dapat mengurangi daya saing organisasi media, karena pembaca mengharapkan berita terbaru dan informasi yang selalu diperbarui. Oleh karena itu, saat ini diperlukan sebuah situs web khusus untuk menggabungkan beberapa situs tersebut menjadi satu. Teknik web scraping merupakan salah satu solusi yang cocok digunakan untuk mengambil berita dari berbagai situs tersebut.

Teknik scraping adalah metode untuk mengambil, menganalisis, dan memproses data dari berbagai sistem atau dokumen. Dalam konteks web, teknik ini dikenal sebagai web scraping, yang melibatkan pengambilan dokumen semi-terstruktur dari Internet, biasanya berupa halaman web dalam

HTML atau XHTML. Proses ini bertujuan untuk menganalisis dokumen dan mengekstrak data tertentu untuk digunakan lebih lanjut. Web scraping berbeda dari pengindeksan web karena lebih fokus pada transformasi data tidak terstruktur dalam format HTML menjadi format terstruktur yang dapat disimpan dan dianalisis dalam database atau spreadsheet. Manfaat web scraping adalah meningkatkan fokus informasi yang diambil, sehingga memudahkan pencarian[ADL18].

Dengan adanya teknik web scraping, pengumpulan data dari berbagai sumber berita dapat dilakukan secara otomatis dan efisien, memungkinkan pengelola website untuk selalu menyajikan informasi terkini kepada pengguna. Teknik ini memungkinkan agregasi berita dari berbagai portal secara real-time, memberikan konten yang lebih kaya dan bervariasi. Penggunaan web scraping tidak hanya membantu dalam mempercepat proses pembaruan konten, tetapi juga memungkinkan analisis tren dan preferensi pembaca secara lebih mendalam. Ini menjadi dasar penting dalam pengembangan strategi konten yang relevan dan menarik bagi *audiens*. Dengan demikian, web scraping berperan vital dalam menghadirkan website berita yang responsif, informatif, dan berdaya saing tinggi.

Berdasarkan uraian diatas, penulis ingin mengangkat masalah tersebut sebagai penelitian yang berjudul “PEMBANGUNAN APLIKASI BERITA BERBASIS WEB DENGAN PENERAPAN TEKNIK WEB SCRAPING”. Penulis berharap dengan penelitian ini dapat mengumpulkan konten dari berbagai sumber secara otomatis menggunakan teknik web scraping.

## **1.2 Identifikasi Masalah**

Berdasarkan latar belakang yang telah dipaparkan sebelumnya, permasalahan yang diangkat dalam Tugas Akhir ini adalah bagaimana membangun aplikasi berita dengan konten yang diperoleh melalui teknik web scraping.

## **1.3 Tujuan Tugas Akhir**

Tujuan Tugas Akhir ini adalah membangun aplikasi berita yang dapat mengumpulkan konten dari berbagai sumber secara otomatis menggunakan teknik web scraping.

## **1.4 Lingkup Tugas Akhir**

Penyelesaian tugas akhir ini memiliki lingkup sebagai berikut :

1. Menggunakan sistem manajemen basis data MySQL
2. Menggunakan bahasa pemrograman PHP.
3. Menggunakan *SDLC* dengan model proses *Waterfall*.
4. Menggunakan library scraping simple html dom.
5. Menerapkan prinsip *Creative Commons* untuk melakukan proses scraping pada website yang akan diambil datanya.
6. Menggunakan alpha testing untuk tahap pengujian aplikasi.

### 1.5 Metodologi Tugas Akhir

Bagian ini akan menjelaskan langkah-langkah metodologi penyelesaian yang dilakukan penulis dalam menyelesaikan tugas akhir. Langkah-langkah tersebut dijelaskan dalam bentuk diagram yang dapat dilihat pada gambar 1.1.



Gambar 1.1 Metodologi Tugas Akhir

Berikut merupakan rincian dari metodologi tugas akhir ini, diantaranya:

### 1. Identifikasi Masalah

Tahap identifikasi masalah merupakan langkah awal yang dilakukan dalam penelitian ini, pada tahap ini bermaksud agar memahami permasalahan yang akan diteliti, sehingga pada tahap analisis dan perancangan mengetahui batasan dari permasalahan yang diteliti.

### 2. Pengumpulan Data

Tahap pengumpulan data dilakukan dengan mencari data-data kebutuhan yang relevan dan valid secara teoritis dari sumber yang jelas. Langkah ini dilakukan untuk mendapatkan data-data penunjang yang diperlukan untuk mendukung pada tahap analisis dan perancangan. Adapun cara-cara pengumpulan data yang dilakukan adalah sebagai berikut :

#### 1) Studi Literatur

Studi literatur merupakan tahap awal yang digunakan untuk mengumpulkan data dan informasi yang difokuskan melalui sumber literatur berbagai website. Tujuan dari tahap ini adalah untuk mendapatkan data yang relevan dan sesuai dengan masalah yang telah diidentifikasi dalam tugas akhir ini.

#### 2) Web Scraping

Tahap web scraping dilakukan untuk mengumpulkan data dari beberapa website berita dan menyimpannya ke dalam database untuk dijadikan sebagai analisis data.

### 3. Metode Waterfall

Tahap metode waterfall merupakan tahapan model proses pengembangan perangkat lunak yang mengadopsi pendekatan linier dan berurutan, di mana setiap tahap harus diselesaikan sepenuhnya sebelum melanjutkan ke tahap berikutnya. Berikut adalah tahap-tahap dalam metode Waterfall:

#### 1) Planning

Tahap planning merupakan tahapan yang mencakup estimasi biaya, penjadwalan proyek, dan sumber daya yang dibutuhkan.

#### 2) Modeling

Tahap modeling dilakukan untuk menerjemahkan kebutuhan yang telah dikumpulkan menjadi desain yang dapat diimplementasikan selama fase konstruksi.

#### 3) Construction

Tahap construction terdiri dari pengkodean (coding) dan pengujian (testing). Kode program ditulis sesuai dengan desain yang telah dibuat dan diuji untuk memastikan bahwa setiap unit berfungsi dengan benar.

#### 4) Deployment

Tahap deployment merupakan tahapan yang mencakup penginstalan perangkat lunak di lingkungan produksi, memberikan pelatihan kepada pengguna, dan mendokumentasikan sistem. Setelah penerapan, perangkat lunak mulai digunakan oleh pengguna akhir.

#### 4. Kesimpulan dan Saran

Tahap kesimpulan dan saran dilakukan untuk mendapatkan kesimpulan dari hasil penelitian yang dilakukan terkait dengan masalah yang telah diidentifikasi, serta dilakukan penyampaian saran dan usulan pengembangan penelitian ini kedepannya.

### 1.6 Sistematika Penulisan

Isi dari subbab ini adalah gambaran rencana sistematika penulisan laporan, yang ditunjukkan dengan daftar bab yang akan dibuat.

Laporan tugas akhir dibuat untuk mendokumentasikan pengerjaan tugas akhir. Maka dari itu, diusulkan sistematika penulisan yang menjelaskan mengenai bab-bab pada laporan tugas akhir beserta isinya secara rinci, serta keterkaitan antara bab sebelum dan sesudahnya. Adapun sistematika penulisannya adalah sebagai berikut:

#### **BAB 1. PENDAHULUAN**

Bab ini memberikan penjelasan umum mengenai tugas akhir yang penulis lakukan. Penjelasan tersebut meliputi latar belakang, identifikasi masalah, tujuan tugas akhir, lingkup tugas akhir, metodologi tugas akhir, dan sistematika penulisan

#### **BAB 2. LANDASAN TEORI DAN PENELITIAN TERDAHULU**

Bab ini berisi definisi, teori-teori serta konsep yang diperlukan dalam pengerjaan tugas akhir. Bab ini juga membahas mengenai jurnal-jurnal ilmiah terdahulu yang memiliki kemiripan dengan tugas akhir yang dikerjakan.

#### **BAB 3. SKEMA PENELITIAN**

Bab ini berisi penjelasan alur penyelesaian tugas akhir, analisis persoalan dan manfaat tugas akhir, kerangka pemikiran teoritis, dan profile tempat penelitian.

#### **BAB 4. ANALISIS DAN PERANCANGAN**

Bab ini membahas tahapan awal dalam proses pembangunan yang terdiri dari analisis dan perancangan. Proses analisis melibatkan pendefinisian perangkat lunak dan analisis kebutuhan dengan menggunakan pemodelan spesifikasi kebutuhan

**BAB 5. IMPLEMENTASI DAN PENGUJIAN**

Bab ini berisi tentang implementasi hasil analisis dan perancangan pada bab sebelumnya, dengan proses penerapan konsep metode yang telah ditentukan untuk mendukung manajemen pengembangan aplikasi yang akan dibuat. Implementasi ini mencakup penerapan algoritma, pembuatan kode program, serta pengujian sistem untuk memastikan bahwa aplikasi berfungsi sesuai dengan kebutuhan dan spesifikasi yang telah dirumuskan. Setiap langkah implementasi didokumentasikan secara rinci, termasuk uji coba dan evaluasi terhadap performa sistem, untuk memastikan aplikasi dapat dioperasikan dengan baik.

**BAB 6. KESIMPULAN DAN SARAN**

Bab ini berisi mengenai hasil penelitian serta pernyataan yang didapat berdasarkan identifikasi masalah yang dikemukakan, serta keterkaitan dari semua tahap yang dilakukan dalam penelitian. Di dalamnya terdapat pula saran yang diusulkan untuk penelitian selanjutnya terkait dengan prospek penelitian selanjutnya, serta rekomendasi penerapan di perusahaan terkait.



## DAFTAR PUSTAKA

- [ADL18] Adli, M. A., & Firgia, L. (2018). Rancang Bangun Web Scraping Pada Media Online Berita. *Jurnal ENTER*.
- [AGU20] Agus Prastyo, E. H., Prisma, I. P., & Wiratsongko, R. (2020). Implementasi Teknik Web Scraping Pada Situs Berita Menggunakan Metode Supervised learning. *Jurnal INOVATE*.
- [FLO20] Flores, V. A., Permatasari, P. A., & Jasa, L. (2020). Penerapan Web Scraping Sebagai Media Pencarian dan Menyimpan Artikel Ilmiah Secara Otomatis Berdasarkan Keyword. *Majalah Ilmiah Teknologi Elektro*.
- [HAF23] Hafiz, Y. A., & Sudarmilah, E. (2023). Implementasi Web Scraping Pada Portal Berita Online. *Jurnal INIASI*.
- [HAK21] Hakim, D. A. (2021). Lisensi Kreativitas Bersama (Creative Commons License) Dalam Perspektif Hukum Indonesia. *ISTINBATH : Jurnal Hukum*.
- [HER21] Hermiati, R., Asnawati, & Kanedi, I. (2021). Pembuatan E-Commerce Pada Raja Komputer Menggunakan Bahasa Pemrograman PHP dan Database MySQL. *Jurnal Media Infotama*.
- [IMA21] Imanullah, M., Onsardi, Toyib, R., & Hidayat, M. T. (2021). Pengenalan Website Sekolah Dasar Muhammadiyah 1 Unggulan Dalam Rangka Meningkatkan Mutu Pembelajaran Dan Promosi Sekolah. *Jurnal Pengabdian Masyarakat Bumi Rafflesia*.
- [JAM14] Jamanti, R. (2014). Pengaruh Berita Banjir Di Koran Kaltim Terhadap Kesadaran Lingkungan Masyarakat Kelurahan Termindung Permai Samarinda. *eJurnal Ilmu Komunikasi*.
- [KIN19] Kinaswara, T. A., Hidayati, N. R., & Nugrahanti, F. (2019). Rancang Bangun Aplikasi Inventaris Berbasis Website pada Kelurahan Bantengan. *Jurnal SENATIK*.
- [PRA03] Pramana, H. W. (2003). *Kunci Sukses Aplikasi Penjualan Berbasis Access*. Jakarta: PT. Elex Media Komputindo.
- [PRA21] Prayitno, E., Suprawoto, T., & Riyanto, B. F. (2021). Optimasi Hasil Pencarian Pada Web Scraping Menggunakan Pembobotan Kata TF-IDF. *Journal of Innovation Research and Knowledge*.
- [PUT20] Putra, A. P., Adriyanto, F., Karisman, Muji Harti, T. D., & Puspitasari, W. (2020). Pengujian Aplikasi Point of Sale Menggunakan Blackbox Testing. *Jurnal Bina Komputer*.
- [SAR20] Sarah Suryaningsih, Riandika, Y. A., Hasanah, A. N., & Anggraito, S. (2020). Aplikasi Wakaf Indonesia Berbasis Blockchain. *Jurnal EDUMATIC*.
- [SAT20] Satriaajati, S., Panuntun, S. B., & Pramana, S. (2020). Implementasi Web Scraping Dalam Pengumpulan Berita Kriminal Pada Masa Pandemi COVID-19. *Seminar Nasional Official Statistics*.
- [YAN18] Yanuardi, & Permana, A. A. (2018). Rancang Bangun Sistem Informasi Keuangan Pada PT. Secret Discoveries Travel And Leisure Berbasis Web. *Jurnal Teknik Informatika (JIKA)*.