

**SISTEM MONITORING UPDATE ARTIKEL DAN
JURNAL ILMIAH SECARA OTOMATIS DI UIN
WALISONGO SEMARANG MENGGUNAKAN TEKNIK
WEB SCRAPING**

SKRIPSI

Diajukan untuk Memenuhi Tugas Akhir dan Melengkapi Syarat
Guna Memperoleh Gelar Sarjana Strata Satu (S-1) dalam
Teknologi Informasi



Diajukan oleh :

**HASYIM YAHYA
NIM : 2008096005**

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI WALISONGO SEMARANG
TAHUN 2024**

PERNYATAAN KEASLIAN

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama : Hasyim Yahya

NIM : 2008096005

Jurusan : Teknologi Informasi

Menyatakan bahwa skripsi yang berjudul:

SISTEM MONITORING UPDATE ARTIKEL DAN JURNAL ILMIAH SECARA OTOMATIS DI UIN WALISONGO SEMARANG MENGGUNAKAN TEKNIK WEB SCRAPING

Secara keseluruhan adalah hasil penelitian/karya saya sendiri,
kecuali bagian tertentu yang dirujuk sumbernya.

Semarang, 25 Juni 2024

Dibuat Pernyataan



Hasyim Yahya

NIM. 2008096005

PENGESAHAN



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI WALISONGO
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Prof. Dr. Hamka Ngaliyan Semarang
Telp.024-7601295 Fax.7615387

PENGESAHAN

Naskah skripsi berikut ini:

Judul : SISTEM MONITORING UPDATE ARTIKEL DAN JURNAL
ILMIAH SECARA OTOMATIS DI UIN WALISONGO SEMARANG
MENGUNAKAN TEKNIK WEB SCRAPING

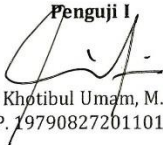
Penulis : **HASYIM YAHYA**
NIM : 2008096005
Jurusan : Teknologi Informasi

Telah diujikan dalam sidang tugas akhir oleh Dewan Penguji
Fakultas Sains dan Teknologi UIN Walisongo dan dapat
diterima sebagai salah satu syarat memperoleh gelar sarjana
dalam Teknologi Informasi.

Semarang, 25 Juni 2024

DEWAN PENGUJI

Penguji I


Dr. Khotibul Umam, M.Kom.
NIP. 197908272011011007

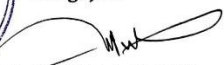
Penguji II


Siti Nur Aini, M.Kom.
NIP. 198401312018012001

Penguji III


Wenty Dwi Yuniarti, S.Pd., M.Kom
NIP. 197706222006042005

Penguji IV


Adzhal Arwani Mahfudh, M.Kom
NIP. 199107032019031006

Pembimbing I


Nur Cahyo Hendro Wibowo, S.T., M.Kom.
NIP. 197312222006041001

Pembimbing II


Siti Nur Aini, M.Kom.
NIP. 198401312018012001

NOTA DINAS

NOTA DINAS

Semarang, 30 Mei 2024

Yth. Ketua Program Studi Teknologi Informasi
Fakultas Sains dan Teknologi
UIN Walisongo Semarang

Assalamu'alaikum. Wr. Wb.

Dengan ini diberitahukan bahwa saya telah melakukan bimbingan,
arahan dan koreksi naskah skripsi dengan :

Judul : SISTEM MONITORING UPDATE ARTIKEL DAN JURNAL
ILMIAH SECARA OTOMATIS DI UIN WALISONGO
SEMARANG MENGGUNAKAN TEKNIK WEB SCRAPING

Nama : **Hasyim Yahya**

NIM : 2008096005

Jurusan : Teknologi Informasi

Saya memandang bahwa naskah skripsi tersebut sudah dapat
diajukan kepada Fakultas Sains dan Teknologi UIN Walisongo
Semarang untuk diujikan dalam Sidang Munaqsyah.

Wassalamu'alaikum. Wr. Wb.

Pembimbing I,



Nur Cahyo Hendro Wibowo, S.T., M.Kom
NIP. 198108122011011007

NOTA DINAS

NOTA DINAS

Semarang, 30 Mei 2024

Yth. Ketua Program Studi Teknologi Informasi
Fakultas Sains dan Teknologi
UIN Walisongo Semarang

Assalamu'alaikum. Wr. Wb.

Dengan ini diberitahukan bahwa saya telah melakukan bimbingan, arahan dan koreksi naskah skripsi dengan :

Judul : SISTEM MONITORING UPDATE ARTIKEL DAN JURNAL
ILMIAH SECARA OTOMATIS DI UIN WALISONGO
SEMARANG MENGGUNAKAN TEKNIK WEB SCRAPING

Nama : **Hasyim Yahya**

NIM : 2008096005

Jurusan : Teknologi Informasi

Saya memandang bahwa naskah skripsi tersebut sudah dapat diajukan kepada Fakultas Sains dan Teknologi UIN Walisongo Semarang untuk diujikan dalam Sidang Munaqsyah.

Wassalamu'alaikum. Wr. Wb.

Pembimbing II,



Siti Nur'aini, M.Kom.
NIP. 198401312018012001

LEMBAR PERSEMBAHAN

Puji dan Syukur penulis ucapkan kepada Allah SWT, laporan tugas akhir ini dapat terselesaikan dengan baik. Karya kecil ini penulis persembahkan untuk:

1. Bapak Subadi dan Ibu Sutimah sebagai orang tua penulis.
2. Kakak Moh Khoiri, Kakak Suciati, Kakak Abdul Majid, Kakak Purisatul Afidah, Kakak Uci Wahyuningsih, dan Kakak Habib Safriyudin sebagai saudara-saudari penulis.
3. Segenap dosen Program Studi Teknologi Informasi
4. Farrel Ibrahim dan Hasan Mukti Wibowo sebagai sahabat sefrekuensi penulis.
5. Sahabat dan teman-teman Teknologi Informasi 2020.
6. Almamater Universitas Islam Negeri Walisongo Semarang.

MOTTO

“Barang siapa menempuh jalan untuk menimba ilmu, niscaya Allah akan mudahkan baginya jalan menuju surga”

—Hadits Riwayat Muslim—

“Might controls everything, and without strength you can not protect anything. Let alone yourself”

—V.S. to D.S—

ABSTRAK

UIN Walisongo Semarang memiliki 46 jurnal ilmiah dengan total ratusan artikel ilmiah di dalamnya. Namun, tingginya jumlah jurnal ilmiah di UIN Walisongo Semarang menimbulkan beberapa kendala. Pertama, tersebarnya artikel ilmiah di berbagai jurnal mempersulit penemuan informasi yang spesifik. Kedua, memeriksa pembaruan terbaru dari setiap jurnal secara manual juga merupakan proses yang memakan waktu dan tidak efisien. Untuk mengatasi permasalahan tersebut, dibuatlah suatu sistem monitoring yang dapat melakukan pengumpulan data yang tersebar dan pengecekan pembaruan secara otomatis menggunakan teknik *web scraping*. Sistem monitoring dibuat menggunakan metode SDLC model *Waterfall*. Pengujian sistem monitoring ini dibagi menjadi dua yaitu pengujian fungsional menggunakan metode *black box testing* dan pengujian *usability* menggunakan metode *user acceptance test*. Berdasarkan pengujian fungsional yang dilakukan, fungsi sistem berjalan sesuai dengan kebutuhan. Pengujian *usability* yang dilakukan memperoleh hasil persentase kelayakan sebesar 76,9% yang menunjukkan bahwa sistem monitoring ini layak untuk diterapkan.

Kata Kunci: Artikel Ilmiah, Jurnal Ilmiah, Sistem Monitoring, Web Scraping

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang senantiasa melimpahkan rahmat dan karunia-Nya kepada penulis sehingga dapat menyelesaikan tugas akhir ini dengan judul **“Sistem Monitoring Update Artikel Dan Jurnal Ilmiah Secara Otomatis di UIN Walisongo Semarang Menggunakan Teknik Web Scraping”** yang digunakan sebagai syarat untuk menyelesaikan pendidikan di Program Studi S1 Teknologi Informasi, Universitas Islam Negeri Walisongo Semarang.

Pada kesempatan ini, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada beberapa pihak yang telah memberikan bantuan, dukungan, dan bimbingan selama penulis menyelesaikan tugas akhir ini, yaitu:

1. Kedua orang tua penulis, Bapak Subadi dan Ibu Sutimah, yang selalu memberikan doa, dukungan, dan semangat kepada penulis sehingga tugas akhir ini dapat terselesaikan.
2. Bapak Dr. Khotibul Umam, ST., M.Kom., selaku Ketua Program Studi Teknologi Informasi UIN Walisongo Semarang, atas bimbingan dan arahnya selama masa perkuliahan.
3. Bapak Nur Cahyo Hendro Wibowo, S.T., M.Kom., selaku Dosen Pembimbing I sekaligus Dosen Wali, yang telah

memberikan arahan, bimbingan, dan masukan yang sangat berharga dalam penyusunan tugas akhir ini.

4. Ibu Siti Nur'aini, M.Kom., selaku Dosen Pembimbing II, yang telah memberikan arahan, bimbingan, dan masukan yang sangat berharga dalam penyusunan tugas akhir ini.
5. Seluruh Dosen Program Studi Teknologi Informasi, serta seluruh dosen dan staf akademik di Universitas Islam Negeri Walisongo Semarang yang telah berkontribusi dan memberikan ilmu pengetahuan selama masa pendidikan.
6. Semua pihak yang tidak dapat disebutkan satu per satu, atas bantuan dan dukungannya dalam menyelesaikan tugas akhir ini.

Penulis berharap tugas akhir ini dapat memberikan manfaat bagi semua pihak, khususnya bagi civitas akademika UIN Walisongo Semarang dan dapat dijadikan sebagai sumber referensi bagi penelitian selanjutnya.

Semarang, 31 Mei 2024

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
PERNYATAAN KEASLIAN	ii
PENGESAHAN	iii
NOTA DINAS	iv
NOTA DINAS	v
LEMBAR PERSEMBAHAN	vi
MOTTO.....	vii
ABSTRAK.....	viii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Identifikasi Masalah.....	5
1.3. Rumusan Masalah	5
1.4. Tujuan Penelitian	6
1.5. Batasan Masalah	6
1.6. Manfaat Penelitian.....	7
BAB II LANDASAN TEORI.....	9
2.1. Kajian Teori	9
2.1.1. Sistem Monitoring	9
2.1.2. Web Scraping.....	10

2.1.3.	Python	12
2.1.4.	Beautiful Soup 4	13
2.1.5.	Cron Job	13
2.1.6.	HTML	14
2.1.7.	System Development Life Cycle	14
2.1.8.	Artikel Ilmiah	15
2.1.9.	Jurnal Ilmiah	16
2.1.10.	Black Box Testing.....	17
2.1.11.	User Acceptance Test	18
2.1.12.	CSS.....	19
2.1.13.	PHP	19
2.1.14.	MySQL.....	20
2.1.15.	XAMPP Server	20
2.1.16.	Bootstrap	21
2.2.	Kajian Penelitian yang Relevan	21
BAB III METODOLOGI PENELITIAN		27
3.1.	Model Pengembangan	27
3.2.	Prosedur Pengembangan	28
3.2.1.	Requirements Definition.....	28
3.2.2.	System and Software Design.....	36
3.2.3.	Implementation and Unit Testing	43
3.2.4.	Integration and System Testing	43
3.2.5.	Operation and Maintenance	50
BAB IV HASIL DAN PEMBAHASAN		51
4.1.	Requirements Definiton	51

4.1.1.	Analisis Proses Scraping Data	51
4.1.2.	Daftar Kebutuhan	69
4.2.	System and Software Design	70
4.3.	Implementation and Unit Testing	75
4.3.1.	Implementasi Web Scraper	75
4.3.2.	Implementasi Website	89
4.4.	Integration and System Testing	95
4.4.1.	Functional Test	95
4.4.2.	User Acceptance Test	100
4.5.	Operation and Maintenance	103
BAB V KESIMPULAN DAN SARAN		104
5.1.	Kesimpulan	104
5.2.	Saran	105
DAFTAR PUSTAKA		106

DAFTAR TABEL

Tabel 2.1 Penelitian yang Relevan	22
Tabel 3.1 Instrumen Pengujian Black Box	44
Tabel 3.2 Instrumen Pengujian User Acceptance Test	47
Tabel 3.3 Konversi Skor.....	48
Tabel 3.4 Hasil Kelayakan	49
Tabel 4.1 Kebutuhan Fungsional Sistem	69
Tabel 4.2 Kebutuhan Non-Fungsional Sistem	70
Tabel 4.3 Desain Tabel Jurnal	71
Tabel 4.4 Desain Tabel Artikel	73
Tabel 4.5 Hasil Pengujian Black Box.....	96
Tabel 4.6 Hasil Pengujian User Acceptance Test.....	101

DAFTAR GAMBAR

Gambar 3.1 Metode Waterfall	27
Gambar 3.2 Blok Diagram Gambaran Umum Sistem	32
Gambar 3.3 Use Case Diagram Sistem Monitoring.....	33
Gambar 3.4 Flowchart Pengambilan Data Update	35
Gambar 3.5 Activity Diagram Berjalannya Sistem	36
Gambar 3.6 ERD Sistem Monitoring.....	38
Gambar 3.7 Desain Mock-up Halaman Home.....	40
Gambar 3.8 Desain Mock-up Halaman Jurnal.....	41
Gambar 3.9 Desain Mock-up Halaman Artikel	42
Gambar 4.1 Contoh Archive Link Jurnal Psikohumaniora.....	54
Gambar 4.2 Struktur HTML Volume Terbaru dari Jurnal Psikohumaniora.....	56
Gambar 4.3 Tampilan Halaman Artikel di Jurnal Ilmiah Psikohumaniora.....	59
Gambar 4.4 Struktur HTML Tanggal Terbit Volume Terbaru dari Jurnal Psikohumaniora.....	60
Gambar 4.5 Struktur HTML Nama dan Link Artikel Ilmiah....	60
Gambar 4.6 Tampilan Pencarian Jurnal Ilmiah di Index SINTA	62
Gambar 4.7 Struktur HTML Akreditasi Jurnal Ilmiah	64
Gambar 4.8 Pola Umum Proses Web Scraping	65
Gambar 4.9 Rancangan Tabel Jurnal	71
Gambar 4.10 Rancangan Tabel Artikel	73
Gambar 4.11 Script Python untuk Koneksi Database	76
Gambar 4.12 Script Python untuk Fungsi Jurnal.....	77
Gambar 4.13 Script Python untuk Fungsi Scrape Issue	80
Gambar 4.14 Script Python untuk Mengakses Database Akreditasi	83
Gambar 4.15 Script Fungsi Iterasi Link Indeks	84

Gambar 4.16 Script Pencarian Jurnal UIN Walisongo Semarang	86
Gambar 4.17 Penggunaan Fitur Cron Job.....	88
Gambar 4.18 Implementasi Halaman Home	89
Gambar 4.19 Implementasi Halaman Jurnal.....	90
Gambar 4.20 Fitur Pagination pada Halaman Jurnal	92
Gambar 4.21 Implementasi Halaman Artikel	93
Gambar 4.22 Fitur Pagination dan Jump Page pada Halaman Artikel	95

DAFTAR LAMPIRAN

Lampiran 1 Tabel Penyaringan Website.....	113
Lampiran 2 Tabel Archive Link	116
Lampiran 3 Tabel Struktur HTML Volume Terbaru	120
Lampiran 4 Tabel Struktur HTML Artikel Ilmiah.....	125
Lampiran 5 Tabel Struktur HTML Akreditasi Jurnal Ilmiah.	134
Lampiran 6 Script Web Scraper Artikel dan Jurnal Ilmiah...	141
Lampiran 7 Script Web Scraper Akreditasi SINTA	153

BAB I

PENDAHULUAN

1.1. Latar Belakang

Teknologi di era modern ini berkembang pesat dan memberikan pengaruh terhadap kehidupan manusia di berbagai aspek kehidupan, baik secara positif maupun negatif. Salah satu faktor yang mendorong perkembangan teknologi adalah pembelajaran serta penelitian yang telah dilakukan oleh para ilmuwan. Belajar adalah proses berubahnya perilaku yang dikarenakan oleh pengalaman dan latihan. Artinya, tujuan kegiatan belajar adalah merubah tingkah laku, baik mengenai aspek pengetahuan, keterampilan dan sikap yang meliputi seluruh organisme atau pribadi (Ningsih, 2016).

Belajar adalah suatu kegiatan yang dapat mempengaruhi dan menentukan masa depan. Hal itu menyebabkan kegiatan belajar harus dilakukan oleh setiap orang, terlebih lagi bagi umat Islam. Di dalam agama Islam perintah untuk belajar sangat ditekankan, yang mana hal itu dibuktikan dengan ayat yang pertama kali diturunkan kepada Nabi Muhammad berupa surat *Al-'Alaq* ayat 1:

اقْرَأْ بِاسْمِ رَبِّكَ الَّذِي خَلَقَ ﴿١﴾

Artinya: "Bacalah dengan (menyebut) nama Tuhanmu yang menciptakan".

Keuntungan manusia dibanding makhluk lainnya terletak pada akal sehat yang dimilikinya. Oleh karena itu, di dalam ayat ini kegiatan membaca, yang merupakan sarana kegiatan belajar, menjadi hal yang pertama kali diperintahkan oleh Allah untuk dilakukan (Ningsih, 2016). Manusia diharapkan dapat mempelajari kekuasaan Allah dengan adanya kegiatan belajar, sehingga mereka dapat meningkatkan dan memperkuat iman mereka kepada Allah.

Kegiatan belajar yang telah dilakukan selama ribuan tahun menghasilkan begitu banyak karya. Contoh dari hasil kegiatan belajar tersebut berupa artikel dan jurnal ilmiah. Artikel ilmiah adalah tulisan yang disusun secara sistematis yang berisi analisis data, temuan ataupun interpretasi hasil penelitian yang dilakukan seseorang (Udil, 2021). Jurnal ilmiah sendiri merupakan publikasi secara berkala kumpulan artikel ilmiah yang telah melalui proses penilaian yang dilakukan oleh para pakar di bidang yang sesuai. Artikel dan jurnal ilmiah hasil kegiatan belajar inilah yang mendorong perkembangan teknologi secara pesat.

Sebagai lembaga pendidikan, UIN Walisongo Semarang turut serta menyumbangkan hasil kegiatan belajar baik berupa artikel maupun jurnal ilmiah untuk mendukung perkembangan teknologi dan ilmu pengetahuan (Mahfudh et al., 2020). UIN Walisongo Semarang memiliki 46 jurnal ilmiah berbeda yang memuat total ratusan artikel ilmiah di dalamnya. Jurnal-jurnal tersebut mempunyai bermacam-macam topik pembahasan, seperti teknologi, politik, ilmu keagamaan dan berbagai topik lainnya. Jurnal-jurnal ilmiah di UIN Walisongo Semarang secara umum mempunyai jadwal terbit dua kali setahun.

Banyaknya jurnal ilmiah di UIN Walisongo Semarang menimbulkan beberapa masalah. Masalah pertama adalah dengan banyaknya jurnal ilmiah di UIN Walisongo Semarang mengakibatkan tersebarnya artikel-artikel ilmiah yang telah diterbitkan. Hal ini dapat menimbulkan masalah karena ketika seseorang ingin mencari artikel ilmiah tetapi tidak mengetahui jurnal yang tepat untuk artikel yang diinginkannya, maka orang itu harus mencari satu per satu jurnal ilmiah yang ada, membuat kegiatan pencarian yang dilakukan membuang-buang waktu (Ulfah & Najiah, 2023).

Banyaknya jurnal ilmiah di UIN Walisongo Semarang juga menyebabkan masalah dimana kebanyakan orang

yang ingin mengetahui pembaruan jurnal-jurnal tersebut harus melakukan pengecekan satu per satu secara manual, dimana hal tersebut kurang efisien untuk dilakukan (Arhandi et al., 2021).

Para peneliti telah menemukan sebuah teknik yang dinamakan *Web Scraping* untuk mengatasi masalah data yang tersebar seperti artikel dan jurnal ilmiah tersebut (Saurkar & Gode, 2018). *Web scraping* merupakan teknik yang digunakan untuk mengekstrak data yang tidak terstruktur dari sebuah situs dimana data tersebut dapat diubah menjadi data terstruktur. Dengan adanya *web scraping*, pengelolaan data yang tersebar menjadi lebih mudah dilakukan.

Penulis mengangkat topik mengenai pembuatan suatu sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang dengan memanfaatkan teknik *web scraping* untuk memecahkan masalah tersebarnya artikel ilmiah dan pemantauan pembaruan jurnal-jurnal ilmiah secara manual. Otomatisasi sistem dilakukan dengan cara mengatur agar sistem melakukan proses *web scraping* minimal sekali dalam jangka waktu satu hari. Dalam pengembangannya, sistem kemudian akan diuji menggunakan metode *Black Box Testing* dan *User Acceptance Test*.

Diharapkan sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis ini dapat memudahkan pencarian artikel ilmiah dan pemantauan *update* jurnal ilmiah di UIN Walisongo Semarang.

1.2. Identifikasi Masalah

Berdasarkan latar belakang di atas, dapat diidentifikasi beberapa masalah yang menjadi fokus dari penelitian ini:

- a. Belum adanya sistem untuk memantau *update* jurnal-jurnal ilmiah secara otomatis di UIN Walisongo Semarang.
- b. Belum adanya sistem yang mengumpulkan artikel-artikel ilmiah yang tersebar di berbagai jurnal ilmiah UIN Walisongo Semarang.

1.3. Rumusan Masalah

Berdasarkan latar belakang di atas, maka peneliti mengambil rumusan masalah yaitu:

- a. Bagaimana cara membangun sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang menggunakan teknik *web scraping* serta menampilkan hasil yang didapat?
- b. Seberapa besar tingkat kelayakan penerapan sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang menggunakan teknik *web scraping*?

1.4. Tujuan Penelitian

Penelitian ini dilakukan dengan maksud dan tujuan sebagai berikut:

- a. Menghasilkan sebuah sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang menggunakan teknik *web scraping*.
- b. Mengukur tingkat kelayakan penggunaan sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang menggunakan teknik *web scraping*.

1.5. Batasan Masalah

Penelitian ini menggunakan beberapa batasan masalah agar proses penelitian dapat dilakukan secara jelas dan memiliki cakupan yang tidak meluas. Batasan masalah pada penelitian ini sebagai berikut:

- a. Sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis ini dibuat hanya terfokus pada artikel dan jurnal-jurnal ilmiah yang ada di UIN Walisongo Semarang.
- b. Jurnal-jurnal ilmiah yang menjadi fokus penelitian ini harus memiliki setidaknya satu volume jurnal yang telah diterbitkan.
- c. Data artikel dan jurnal ilmiah yang diekstrak hanya berupa data umum yang dapat diakses oleh siapa saja.

- d. Teknik *Web Scraping* yang dipilih menggunakan bahasa pemrograman *Python* dan *library* Beautiful Soup 4.
- e. Fitur-fitur sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang ini terdiri atas tiga halaman utama sistem, yaitu Halaman Home, Halaman Jurnal dan Halaman Artikel.
- f. Fitur Halaman Home dibatasi untuk menjalankan proses *web scraping* dan pemantauan *update* artikel dan jurnal ilmiah.
- g. Fitur Halaman Jurnal dibatasi untuk menampilkan informasi mengenai jurnal ilmiah, kapan volume terkini di jurnal ilmiah tersebut diupdate, dan fitur *filter* berdasarkan kata kunci serta pengurutan data jurnal ilmiah.
- h. Fitur Halaman Artikel dibatasi untuk menampilkan informasi artikel ilmiah dan fitur *filter* berdasarkan tahun terbit, asal jurnal, pengurutan data, dan kata kunci yang dimasukkan.

1.6. Manfaat Penelitian

Dengan adanya penelitian ini, diharapkan penulis dapat memberikan manfaat bagi pembaca berupa manfaat teoritis dan manfaat praktis.

a. Manfaat Teoritis

Memberikan ide-ide baru mengenai cara pemanfaatan teknik *web scraping* dan menjadi referensi bagi yang tertarik di bidang *web scraping* dalam melakukan penelitian.

b. Manfaat Praktis

Menghasilkan sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang untuk mengatasi masalah data artikel ilmiah yang tersebar dan pemantauan *update* jurnal ilmiah secara manual.

BAB II

LANDASAN TEORI

2.1. Kajian Teori

2.1.1. Sistem Monitoring

Sistem informasi adalah serangkaian elemen yang saling terhubung yang berkolaborasi dalam proses pengumpulan, pengolahan, penyimpanan, dan presentasi informasi guna mendukung berbagai aspek seperti pengambilan keputusan, koordinasi, pengelolaan, analisis, dan visualisasi di dalam suatu organisasi (Akh Sokhibi, 2020).

Monitoring sendiri merupakan langkah untuk mengkaji kegiatan yang dilaksanakan telah sesuai dengan rencana, mengidentifikasi masalah yang timbul agar langsung dapat diatasi, melakukan penilaian pola kerja dan manajemen yang digunakan sudah tepat untuk mencapai tujuan, mengetahui kaitan antara kegiatan dengan tujuan untuk memperoleh ukuran kemajuan (Kumala et al., 2018)

Monitoring membantu mengingatkan ketika terjadi sesuatu yang salah dan membantu agar pekerjaan tetap pada jalurnya, monitoring bertujuan meningkatkan efektifitas dan efisiensi dari sebuah kegiatan, dan didasarkan pada sasaran dan rencana

kegiatan yang sudah ditentukan, monitoring memungkinkan kita untuk untuk menentukan apakah sumber daya kita telah mencukupi dan telah digunakan dengan baik dan menjadi dasar yang berguna untuk evaluasi selanjutnya (Suaidah & Sidni, 2018). Sistem monitoring akan memberikan dampak yang baik bila dirancang dan dilakukan secara efektif (Setiawansyah, 2020).

2.1.2. Web Scraping

Web scraping merupakan sekumpulan teknik untuk mendapatkan informasi dari sebuah *website* secara otomatis tanpa harus menyalinnya secara manual. Tujuan dari *web scraper* adalah untuk mencari informasi tertentu, mengekstraksi informasi tersebut, kemudian menghimpunnya dalam halaman web yang baru. Beberapa *web scraper* memfokuskan dirinya untuk mengubah data yang tidak terstruktur dan menyimpannya ke dalam data yang terstruktur (Vargiu & Urru, 2012).

Cara kerja *web scraper* adalah dengan membuka halaman *website* tujuan, kemudian mengunduh data yang ada di dalam web, mengekstraknya ke dalam format yang terstruktur dan menyimpannya ke dalam sebuah *file* atau *database* (Joko & Haris, 2020). *Web scraper* dapat mengunduh konten berupa teks, gambar

maupun video yang diformat sebagai HTML dari beberapa halaman web dan mengekstrak data darinya. Beberapa metode yang dapat digunakan dalam teknik *web scraping* di antaranya adalah *Manual Parser*, *HTML Parser*, *XPath Parser*, dan *JSON Parser* (Levi et al., 2020a). Penulis memutuskan untuk menggunakan *HTML Parser* karena *HTML Parser* mempunyai beragam *tools* untuk melakukan proses *web scraping*. Dengan beragam *tools* yang dapat digunakan, pengguna dapat memilih *tools* yang sesuai dengan karakteristik *website* yang akan diekstrak (Levi et al., 2020).

Web scraping secara hukum merupakan hal yang sah untuk dilakukan selama data yang dikumpulkan digunakan untuk kepentingan pribadi dan dalam tidak melanggar undang-undang yang berlaku. Jika data akan dipublikasikan atau jika konten memiliki hak cipta dan melanggar persyaratan layanan maka ada beberapa preseden hukum yang perlu diperhatikan.

Di *Feist Publications, Inc. v. Rural Telephone Service Co.*, Mahkamah Agung Amerika Serikat memutuskan bahwa *scraping* dan penerbitan ulang data, seperti daftar telepon, diizinkan. Hal ini serupa dengan *Australia, Telstra Corporation Limited v. Phone*

Directories Company Pty Ltd, menunjukkan bahwa hanya data dengan penulis yang dapat diidentifikasi dilindungi dengan hak cipta (Jarmul & Lawson, 2017).

Kasus lainnya di Amerika Serikat, mengevaluasi penggunaan kembali berita *Associated Press* untuk produk berita gabungan, dinyatakan sebagai pelanggaran hak cipta di *Associated Press v. Meltwater*. Hal ini menunjukkan bahwa ketika data yang didapatkan merupakan data publik seperti lokasi bisnis dan daftar telepon, data tersebut dapat diterbitkan ulang dengan mengikuti aturan yang berlaku. Namun, jika data tersebut merupakan opini, ulasan atau data pengguna pribadi, kemungkinan besar tidak dapat dipublikasikan ulang karena alasan hak cipta (Jarmul & Lawson, 2017).

Web scraping biasanya dibuat untuk menargetkan situs web atau situs tertentu dengan tujuan mengumpulkan informasi spesifik di situs tersebut. *Scraper* dibuat untuk mengakses halaman spesifik dan perlu dimodifikasi jika situs berubah atau jika informasi yang ada di situs diubah (Baskara, 2022).

2.1.3. Python

Bahasa pemrograman Python pertama kali diciptakan oleh Guidovan Rossum pada awal tahun 1990 di negeri Belanda sebagai pengganti bahasa

pemrograman yang disebut ABC. Walaupun Guido adalah orang yang pertama kali menciptakan bahasa pemrograman ini, tetapi bahasa pemrograman Python yang digunakan sekarang merupakan kontribusi dari berbagai sumber(Purnomo & Ayub, 2021). Bahasa pemrograman Python merupakan bahasa pemrograman yang dapat dikembangkan oleh siapa saja karena bersifat *Open Source* (Yuniarti, 2019).

2.1.4. Beautiful Soup 4

Beautiful Soup 4 (BS4) adalah sebuah pustaka atau *library* dalam bahasa pemrograman Python yang berfungsi untuk menganalisis dan memanipulasi struktur dokumen web. Beautiful Soup 4 (BS4) pertama kali dirilis pada tanggal 13 September 2004 oleh Paul Prescod. Versi awal BS4 ini hanya mampu menangani dokumen HTML. Seiring waktu, BS4 terus berkembang dengan penambahan fitur baru, termasuk dukungan untuk *parsing* dokumen XML dan peningkatan performa (I Gusti Bagus Indra, 2023).

2.1.5. Cron Job

Cron job merupakan aturan penjadwalan suatu perintah yang dijalankan secara periodik. *Cron* diatur oleh crontab (*cron table*) *file*, yakni adalah sebuah *file* konfigurasi yang berisikan perintah yang akan berjalan secara periodik berdasarkan pengaturan jadwal yang

diberikan. Kontrol untuk mengatur waktu akan diatur dengan menggunakan PHP sebagai bahasa pemrograman, di mana PHP adalah bahasa yang kita gunakan untuk membuat *server* menghasilkan *output* yang berbeda setiap kali *browser* melakukan *request* terhadap suatu halaman (Wirawan, 2020).

Penelitian ini memanfaatkan fasilitas *cron job* untuk melakukan otomatisasi *script web scraper* yang digunakan untuk melakukan ekstraksi data artikel dan jurnal ilmiah di UIN Walisongo Semarang.

2.1.6. HTML

Hyper Text Markup Language atau HTML adalah kode yang digunakan untuk mengatur tata letak tampilan halaman web dan isinya. HTML bukanlah bahasa pemrograman, HTML merupakan bahasa *markup* yang melakukan format pada struktur *content* dokumen. HTML terdiri atas beberapa elemen dan *tag* khusus yang dapat digunakan untuk format *content* agar memiliki tampilan tertentu (Gani & Achmad, 2019). *HTML Parser* merupakan metode ekstraksi data berdasarkan elemen dan *tag* khusus tersebut.

2.1.7. System Development Life Cycle

Proses pengembangan suatu sistem harus melewati beberapa tahapan dimulai dari saat sistem itu direncanakan sampai dengan sistem tersebut

diterapkan, dioperasikan dan dipelihara (Prihatiningsih et al., 2023). Bila pengoperasian dari sistem yang dikembangkan itu masih mengalami masalah kritis yang tidak dapat diatasi dengan pemeliharaan sistem maka perlu dikembangkan sistem untuk mengatasinya. Bila hal ini yang terjadi berarti proses harus kembali ke tahap pertama, yaitu tahap perencanaan sistem. Siklus ini disebut dengan siklus hidup suatu sistem (*sistem life cycle*).

Daur atau siklus hidup hasil pengembangan sistem (*System Deveopment Life Cycle*) merupakan suatu bentuk yang digunakan untuk menggambarkan tahapan utama dan langkah-langkah didalam tahapan tersebut dalam proses pengembangannya (Fransisca Khurnia Erkhani, 2021). Kegiatan pengembangan sistem dapat diartikan sebagai kegiatan membangun sistem baru untuk mengganti, memperbaiki atau meningkatkan fungsi sistem yang sudah ada.

2.1.8. Artikel Ilmiah

Karya tulis ilmiah merupakan karya manusia yang dituangkan secara tertulis dalam mengusung suatu perkembangan ilmu (Kusmana, 2020). Artikel ilmiah sebagai bagian dari karya ilmiah adalah karya ilmu pengetahuan yang menyajikan fakta umum dan ditulis menurut metodologi penulisan yang baik dan

benar (Brotowijoyo, 2002). Menurut Brotowijoyo (2002), artikel ilmiah sebagai bagian karya ilmiah selalu ditulis dengan bahasa yang konkret, gaya bahasanya formal, kata-katanya teknis, dan didukung dengan fakta umum yang dapat dibuktikan benar tidaknya.

2.1.9. Jurnal Ilmiah

Jurnal ilmiah merupakan publikasi berkala yang memuat karya tulis ilmiah hasil penelitian yang telah melalui proses penelaahan sejawat (*peer review*). Jurnal ilmiah menjadi media penting bagi para peneliti untuk menyebarkan temuan dan berkontribusi pada perkembangan ilmu pengetahuan (Erfani, 2021).

Saat ini ada banyak jenis jurnal yang dapat dipilih sesuai dengan topik tulisan. Umumnya, terdapat 4 jenis jurnal, diantaranya adalah: 1) Jurnal dengan topik spesifik versus topik umum. Terdapat jurnal yang menerbitkan artikel-artikel dengan topik sangat spesifik, jurnal pendidikan anak usia dini, jurnal ilmu tanah, jurnal gigi dan mulut. Tetapi, ada juga jurnal yang menerbitkan topik secara umum saja, seperti jurnal ilmu pendidikan; 2) Jenis riset (*qualitative vs quantitative*). Ada jurnal yang hanya menerima tulisan riset kualitatif, atau ada jurnal yang hanya menerima riset kuantitatif saja. Ada juga jurnal yang menerima

keduanya. Seorang penulis perlu mengetahui kecenderungan jurnal yang akan dia tuju; 3) *High impact* vs *low impact*. Jurnal juga mempunyai ingkat kepopuleran masing-masing. Jika suatu jurnal banyak dikutip oleh penulis lain dalam menulis artikel ilmiah mereka, maka jurnal tersebut dikategorikan sebagai jurnal '*high impact*'. Begitu juga sebaliknya, jika jurnal tersebut sedikit dikutip oleh khalayak, maka jurnal tersebut termasuk ke dalam jurnal *low impact* (Daud et al., 2020).

2.1.10. Black Box Testing

Black Box Testing merupakan teknik pengujian *software* yang fokus pada spesifikasi fungsi-fungsi yang ada pada perangkat lunak yang dikembangkan (Roohullah Jan et al., 2016). *Black box testing* cenderung dapat menemukan beberapa hal seperti fungsional yang tidak benar atau tidak ada, kesalahan struktur data, kesalahan akses basis data, kesalahan antar muka, kesalahan *performance* serta kesalahan inisialisasi dan terminasi (Pratama & Junianto, 2016).

Keuntungan menggunakan metode *black box testing* adalah: (1) Penguji tidak perlu memiliki pengetahuan tentang Bahasa pemrograman tertentu (2) Pengujian dilakukan dari sudut pandang pengguna, hal ini membantu untuk mengungkapkan ambiguitas

atau inkonsistensi dalam spesifikasi persyaratan; (3) *Programmer* dan *tester* keduanya saling bergantung satu sama lain (Jaya, 2018).

2.1.11. User Acceptance Test

User Acceptance Testing merupakan pengujian yang dilakukan oleh *end user* yang langsung berinteraksi dengan sistem dan dilakukan verifikasi apakah fungsi yang ada telah berjalan sesuai dengan kebutuhan/fungsinya. *Acceptance testing* merupakan pengujian yang dilakukan oleh pengguna yang menggunakan teknik pengujian *black box* untuk menguji sistem terhadap spesifikasinya.

User Acceptance Testing menguji yang dilakukan oleh pengguna sistem. Hasil dari pengujian dapat dijadikan bukti bahwa sistem dapat membantu para pengguna (Wahyudi & Alameka, 2023). *User Acceptance Testing* dilakukan pada pengembangan perangkat lunak bertujuan untuk memastikan sistem memenuhi kebutuhan sebenarnya dari pengguna, bukan hanya spesifikasi sistem.

User Acceptance Testing (UAT) adalah fase terakhir dari proses pengujian perangkat lunak. Selama UAT, perangkat lunak perangkat lunak diuji untuk memastikan tugas-tugas apakah sudah sesuai dengan spesifikasinya. UAT adalah salah satu prosedur

proyek perangkat lunak final dan paling penting yang harus terjadi sebelum perangkat lunak tersebut dikembangkan dan diluncurkan ke pasar. UAT juga dikenal sebagai pengujian beta, pengujian aplikasi atau pengujian pengguna akhir (Cimperman, 2006).

2.1.12. CSS

CSS (*Cascading Style Sheet*) merupakan salah satu bahasa pemrograman web yang bertujuan untuk membuat *website* agar lebih menarik dan terstruktur. Cara kerja CSS dalam memodifikasi HTML dengan memilih elemen HTML yang akan diatur kemudian memberikan properti yang sesuai dengan tampilan yang diinginkan. Dalam memberikan aturan pada elemen HTML, skrip CSS terdiri atas 3 bagian yaitu selector untuk memilih elemen yang akan diberi aturan, *property* yang merupakan aturan yang diberikan dan *value* sebagai nilai dari aturan yang diberikan (Noviantoro et al., 2022).

2.1.13. PHP

PHP adalah salah satu bahasan pemrograman skrip yang dirancang untuk membangun aplikasi web (Solahudin, 2021). Ketika dipanggil dari *web browser*, program yang ditulis dengan PHP akan di-*parsing* di dalam *web server* oleh *interpreter* PHP dan

diterjemahkan ke dalam dokumen HTML, yang selanjutnya akan ditampilkan kembali *web server*.

Karena pemrosesan program PHP dilakukan didalam lingkungan *web browser*, PHP dikatakan sebagai bahasa sisi server (*server-side*). Oleh sebab itu, seperti yang telah dikemukakan sebelumnya, kode PHP tidak akan terlihat pada saat user memilih perintah “*View Source*” pada *web browser* yang mereka gunakan.

2.1.14. MySQL

MySQL merupakan *software database open source* yang sering digunakan untuk mengolah basis data yang menggunakan bahasa SQL (Putra & Nita, 2019). Dapat ditarik kesimpulan bahwa MySQL merupakan aplikasi yang digunakan untuk mengolah basis data yang banyak digunakan untuk membangun aplikasi yang menggunakan database.

2.1.15. XAMPP Server

Menurut Supono & Putratama (2018) mendefinisikan bahwa “XAMPP terdiri dari *apache web server*, MySQL, PHP, Pearl, FTP *server* dan phpMyAdmin”. Sedangkan, menurut Risnandar mengemukakan bahwa “XAMPP adalah suatu program yang digunakan sebagai *server* untuk mengeksekusi fungsi yang ada dalam halaman website”. Berdasarkan pengertian dari para ahli di atas, dapat disimpulkan

bahwa XAMPP merupakan *software* yang digunakan sebagai *server* untuk mengeksekusi fungsi *website* tanpa jaringan internet.

2.1.16. Bootstrap

Bootstrap merupakan merupakan sebuah *framework* dari CSS yang digunakan untuk mempermudah pengembang dalam membangun sebuah *website* yang mempunyai tampilan menarik. Bootstrap adalah sebuah *platform* yang digunakan untuk membuat tampilan antarmuka dari sebuah *website* dan aplikasi yang berbasis web. Bootstrap sendiri berisikan kode dari HTML dan CSS yang telah dilengkapi dari segi desain untuk kebutuhan tipografi, bentuk, tombol, dan lain sebagainya. Tujuan adanya Bootstrap sendiri digunakan untuk meringankan dalam pembuatan dan pengembangan sebuah web (Jantce TJ Sitinjak et al., 2020).

2.2. Kajian Penelitian yang Relevan

Berikut ini adalah tabel berisi kumpulan penelitian terdahulu yang berkaitan dengan penelitian yang ada pada tugas akhir kali ini.

Tabel 2.1 Penelitian yang Relevan

Judul Penelitian	Hasil Penelitian
Sistem Informasi Web Scraping untuk Pengisian Deskripsi Produk Penjual Online pada E-Marketplace (Dana Febri Setiawan, 2019)	Perancangan sistem ini menggunakan Teknik <i>Web Scraping</i> dengan metode <i>HTML Parser</i> dan UML. Hasil penelitian ini adalah kesuksesan proses <i>web scraping</i> pada <i>e-marketplace</i> Tokopedia, Bukalapak, Lazada, Zalora, Blanja.com, Kriya.co.id, Zilingo, Amazon, dan Alibaba. Penelitian dianggap sukses karena sistem yang dibuat dapat mengisi deskripsi produk dengan akurat dan sesuai data yang di- <i>scraping</i> . Kelebihan penelitian ini adalah penelitian ini dapat meningkatkan efisiensi dan efektivitas waktu dalam pengisian deskripsi produk serta memungkinkan pengisian deskripsi produk secara otomatis untuk berbagai produk.

Judul Penelitian	Hasil Penelitian
Implementation of Web Scraping to Build a Web-Based Instagram Account Data Downloader Application (Arif Himawan, Adri Priadana, dan Aris Wahyu Murdiyanto, 2020)	Teknik <i>Web Scraping</i> yang digunakan adalah <i>JSON Parser</i> . Hasil akhir penelitian ini adalah kesuksesan pembuatan aplikasi <i>web-based</i> yang dapat mengunduh data akun Instagram, seperti profil, postingan, dan pengikut. Penelitian ini dianggap sukses karena data akun Instagram yang diunduh lengkap dan akurat. Kelebihan penelitian ini adalah aplikasi yang dibuat dapat membantu pengguna dalam menganalisis data akun Instagram untuk keperluan bisnis atau penelitian, mempermudah pengguna dalam mengumpulkan data akun Instagram secara terstruktur, dan pengunduhan data akun Instagram dapat dilakukan secara otomatis.
Web Scraping with HTML DOM Method for Data Collection of	Teknik <i>Web Scraping</i> yang dipakai adalah <i>HTML DOM</i>

Judul Penelitian	Hasil Penelitian
Scientific Articles from Google Scholar (Alam Rahmatullah dan Rohmat Gunawan, 2020)	<i>Parser</i> . Hasil akhir penelitian ini adalah kesuksesan merekap data peneliti dari Google Scholar berupa data profil, afiliasi, sitasi, dan judul artikel peneliti. Penelitian ini dianggap sukses karena data peneliti yang diekstrak merupakan data yang lengkap dan akurat. Kelebihan sistem ini adalah Sistem ini terintegrasi dengan Google Scholar, platform populer untuk pencarian artikel ilmiah, sehingga memudahkan akses dan pengumpulan data.
Perancangan Aplikasi Perbandingan Harga Produk (Historical Data) Menggunakan Teknik Scraping Web (Ufi Mufidah dan Manasse Siahaan, 2021)	Teknik <i>Web Scraping</i> yang dipakai adalah <i>HTML Parser</i> . Hasil akhir penelitian ini berupa <i>website</i> yang mempunyai perbandingan perbedaan harga produk setiap hari dari situs Zalora. Kelebihan dari penelitian ini adalah <i>website</i> yang dibuat dapat membantu memprediksi

Judul Penelitian	Hasil Penelitian
	kenaikan atau penurunan harga suatu barang di situs Zalora.

Masing-masing penelitian tersebut adalah penelitian yang berorientasi pada penggunaan teknik *web scraping* untuk mengekstrak data tidak terstruktur dari suatu *website* dan menyusun kembali data tersebut agar mempermudah pengolahan kembali data yang didapat. Data yang telah diperoleh dapat digunakan untuk berbagai penelitian yang berbeda.

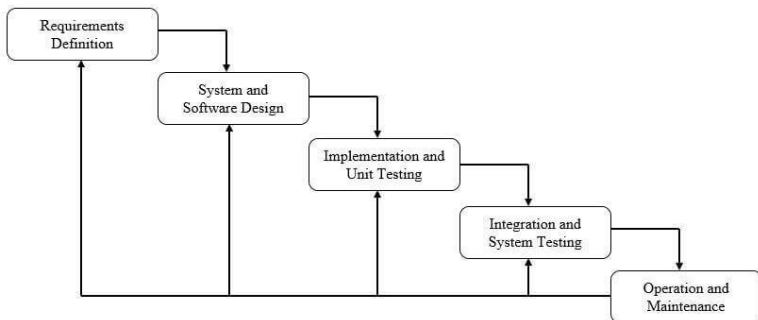
Berdasarkan penelitian-penelitian di atas, peneliti bermaksud merancang sistem monitoring *update* artikel dan jurnal ilmiah di UIN Walisongo Semarang secara otomatis menggunakan teknik *web scraping*. Teknik *web scraping* yang digunakan akan memanfaatkan metode *HTML Parser* dengan Bahasa Pemrograman Python dan *library* Beautiful Soup 4. Data jurnal dan artikel ilmiah yang didapat kemudian akan ditampilkan dalam *website* yang dikembangkan dengan Bahasa Pemrograman PHP dan *Framework* Bootstrap sesuai model perancangan *waterfall*. Terakhir, keberhasilan sistem akan diuji dengan metode *Black Box Testing* serta *User Acceptance Test*.

Kelebihan dari penelitian ini dibandingkan dengan penelitian-penelitian terdahulu adalah proses pengecekan data artikel dan jurnal ilmiah yang otomatis secara berkala dapat menghemat waktu dan tenaga dalam pengecekan data yang awalnya tersebar di jurnal-jurnal berbeda di UIN Walisongo Semarang.

BAB III METODOLOGI PENELITIAN

3.1. Model Pengembangan

Metode pengembangan sistem yang digunakan dalam penelitian ini yaitu metode pengembangan *System Development Life Cycle* (SDLC) dengan model *Waterfall*.



Gambar 3.1 Metode Waterfall

Metode SDLC dalam rekayasa sistem dan rekayasa perangkat lunak merupakan proses pembuatan dan perubahan sistem serta model dan metodologi yang digunakan untuk mengembangkan sistem-sistem tersebut. Konsep ini umumnya mengacu pada sistem komputer atau informasi (Sugiarto et al., 2021).

SDLC terdiri dari beberapa fase yang dimulai dari fase perencanaan, analisis, perancangan, implementasi hingga pemeliharaan sistem. SDLC memiliki beberapa model dalam penerapan tahapan prosesnya salah satunya yang

digunakan pada penelitian kali ini, yaitu model *Waterfall* (Wahyudin & Rahayu, 2020).

3.2. Prosedur Pengembangan

Model *Waterfall* atau sering disebut dengan *linear sequential model* merupakan pendekatan yang sistematis dan berurutan dalam pengembangan perangkat lunak. Berdasarkan Gambar 3.1, dapat dijelaskan bahwa model *waterfall* terdiri dari lima tahapan, yaitu:

3.2.1. Requirements Definition

Dalam tahap ini, pengembang sistem perlu berkomunikasi dengan tujuan memahami kebutuhan perangkat lunak yang diinginkan oleh pengguna serta batasan-batasan yang harus diperhatikan. Informasi ini dapat diperoleh melalui wawancara, diskusi, atau survei langsung dengan pengguna.

Data yang diperoleh kemudian dianalisis untuk mendapatkan informasi yang sesuai dengan kebutuhan pengguna (Wahid, 2020).

a. Teknik Pengumpulan Data

Data yang diperoleh dalam penelitian ini berupa dua macam data, yaitu data primer dan data sekunder.

1) Data Primer

Data Primer didapatkan melalui wawancara kepada pembimbing serta pelajar yang menggunakan jurnal ilmiah UIN Walisongo Semarang untuk mengetahui hal-hal penting yang diperhatikan ketika menggunakan jurnal ilmiah. Hal-hal penting tersebut kemudian dijadikan dasar dari jenis informasi yang perlu diekstrak untuk digunakan di sistem monitoring update artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang menggunakan teknik *web scraping*.

Hal-hal penting yang didapatkan sebagai data primer berupa adanya data jurnal dan data artikel. Data jurnal berupa data nama, indeks, jadwal terbit, dan volume terbaru dari masing-masing jurnal. Sedangkan untuk data artikel berupa data nama, penulis, asal, dan tanggal terbit masing-masing artikel ilmiah.

2) Data Sekunder

Data sekunder didapatkan dengan cara melakukan studi literatur. Studi literatur merupakan penyelesaian masalah melalui sumber-sumber tulisan terdahulu yang pernah dibuat. Melalui metode ini peneliti memperoleh data dari buku, jurnal, maupun referensi lainya yang berkaitan dengan penelitian. Contoh data sekunder yang didapat berupa data yang berhubungan dengan teknik *web scraping*.

b. Analisis Kebutuhan Sistem

Analisa kebutuhan perangkat lunak ini diawali dengan penggambaran secara umum bagaimana sistem monitoring ini bekerja, pengidentifikasian aktor yang terlibat dalam sistem, penjabaran kebutuhan sistem dan memodelkannya ke dalam *use case diagram* dan penjelasan mengenai proses *scraping data* yang dilakukan. Analisa kebutuhan ini ditujukan untuk menggambarkan kebutuhan-kebutuhan yang harus disediakan oleh sistem agar dapat memenuhi kebutuhan pengguna.

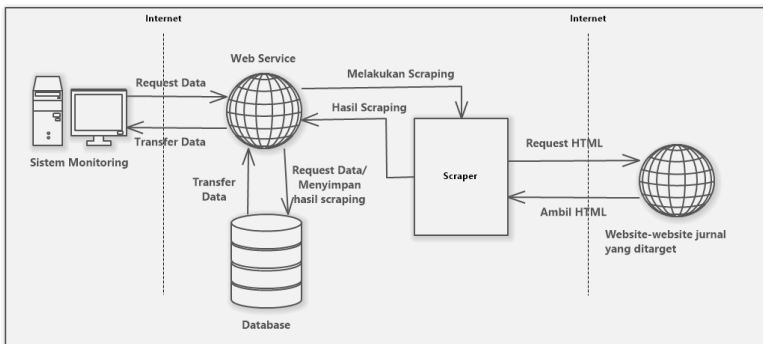
1) Gambaran Umum Sistem

Gambaran umum dari sistem yang berhubungan dengan *web scraping* biasanya dimulai dengan adanya proses pembaruan data dari aplikasi klien, yang kemudian akan disampaikan kepada *scraper* oleh *web service* yang tersedia dan *scraper* akan mengekstrak data dari situs yang dituju jika permintaan akses disetujui oleh situs tersebut. Hasil *scraping* yang dilakukan akan disimpan ke dalam *database*, baik terdapat perubahan data maupun tidak, dan ditampilkan di aplikasi klien yang meminta pembaruan data (Ramadhan, 2014).

Proses kerja sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang menggunakan teknik *web scraping* dapat dilihat pada Gambar 3.2.

Proses berjalannya sistem dimulai dari *user* yang mengakses sistem monitoring dan melakukan *request* data *update* ke *web service* (XAMPP). Request ini kemudian diterima oleh *scraper* yang akan melakukan *request* HTML kepada situs jurnal ilmiah yang telah

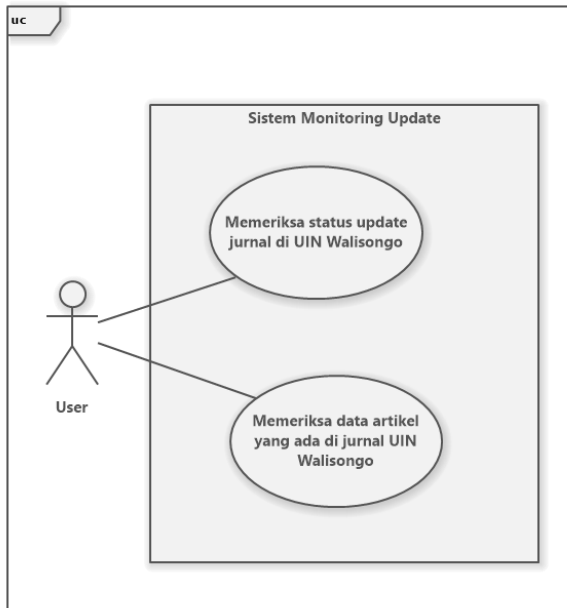
ditentukan. Hasil *scraping* yang dilakukan kemudian dikembalikan kepada *web service* untuk disimpan ke dalam *database*. Setelah data disimpan, baru kemudian data *update* yang di-request oleh *user* ditampilkan di sistem monitoring.



Gambar 3.2 Blok Diagram Gambaran Umum Sistem

2) Use Case Diagram

Use case dan aktor menggambarkan ruang lingkup sistem yang sedang dibangun. *Use case* meliputi semua hal yang ada pada sistem, sedangkan aktor meliputi semua hal yang ada di luar sistem. *Use case diagram* menggambarkan fungsionalitas dari suatu sistem, sehingga pengguna sistem paham dan mengerti mengenai kegunaan sistem yang akan dibangun.



Gambar 3.3 Use Case Diagram Sistem Monitoring

Gambar 3.3 menjelaskan *use case* di dalam sistem monitoring ini. *User* dapat melihat data informasi *update* jurnal ilmiah di UIN Walisongo Semarang. Informasi yang diharapkan untuk ditampilkan oleh sistem berupa ringkasan detail dari masing-masing jurnal ilmiah. Setiap jurnal ilmiah akan mempunyai detail seperti nama jurnal, indeks jurnal, jadwal publish jurnal, dan kapan volume terkini jurnal tersebut diupdate sebagai bentuk status update jurnal.

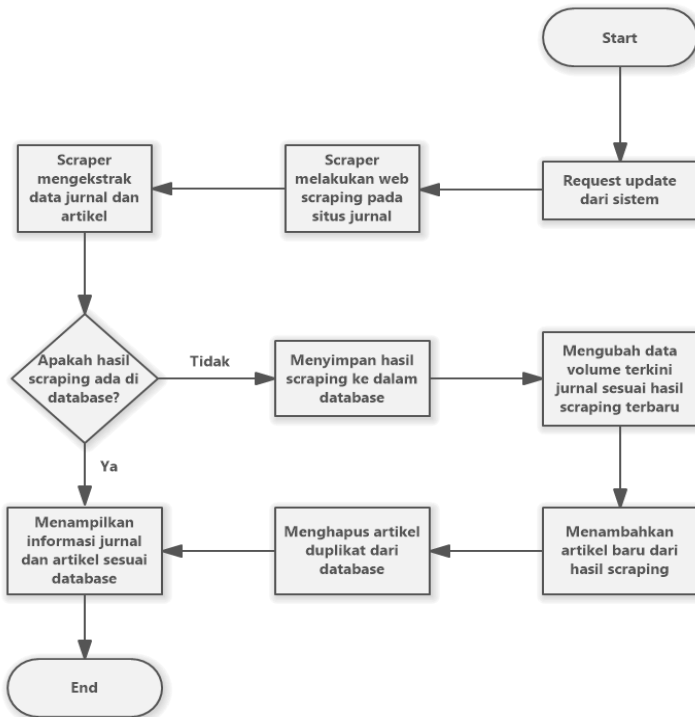
Selain mengakses informasi update jurnal ilmiah, user juga dapat mengakses data artikel ilmiah yang telah diekstrak dari jurnal-jurnal ilmiah yang ada. Setiap artikel ilmiah akan mempunyai informasi nama artikel, tanggal terbit, jurnal ilmiah tempat artikel diekstrak, serta link menuju artikel asli di jurnal ilmiah asalnya.

3) Proses Scraping Data

Proses *scraping* atau ekstraksi data artikel dan jurnal ilmiah UIN Walisongo Semarang mempunyai beberapa tahapan ketika ada *request update* data artikel dan jurnal ilmiah. Tahapan tersebut digambarkan di Gambar 3.4.

Ketika sistem menerima permintaan untuk melakukan *request update* data jurnal terbaru, *scraper* akan melakukan ekstraksi data pada situs jurnal ilmiah yang tersedia. Data yang diekstrak berupa data jurnal dan artikel terbaru dari situs jurnal ilmiah. Setelah *scraper* mengekstrak data jurnal dan artikel, *scraper* akan membandingkannya dengan data yang telah ada di *database* sebelumnya. Apabila hasil *scraping* yang baru sama dengan data yang ada

di *database*, maka sistem akan langsung menampilkan data yang tersedia di *database* tanpa melakukan perubahan apapun terhadap *database*.



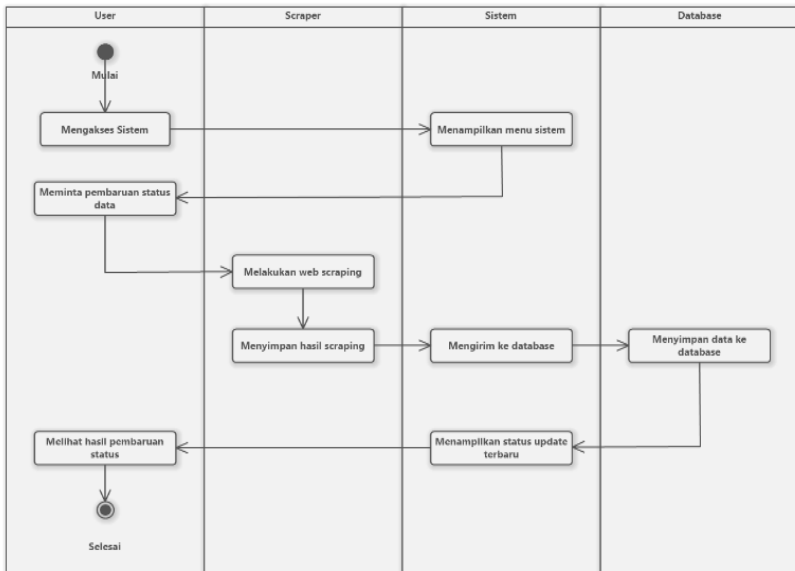
Gambar 3.4 Flowchart Pengambilan Data Update

Ketika hasil *scraping* mempunyai data yang baru dan belum ada di *database*, maka sistem akan menyimpan hasil *scraping* tersebut. Sistem kemudian akan mengganti data volume terkini jurnal dengan hasil *scraping* terbaru.

Selain itu, sistem akan menambahkan data artikel terbaru ke dalam *database* dan memeriksa apakah ada data artikel yang terduplikat. Jika memang ada artikel duplikat, sistem kemudian akan menghapus duplikat yang ada untuk menjaga keakuratan sistem. Terakhir, sistem akan menampilkan informasi artikel dan jurnal sesuai dengan *database* terbaru.

3.2.2. System and Software Design

a. System Design



Gambar 3.5 Activity Diagram Berjalannya Sistem

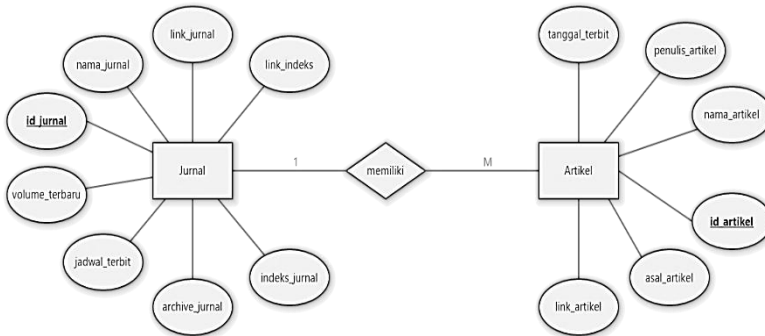
Gambar 3.5 diatas menunjukkan diagram aktivitas sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang menggunakan teknik *web scraping*.

Sistem dimulai ketika ada *user* yang mengakses sistem monitoring *update* artikel dan jurnal ilmiah. Selanjutnya sistem akan menunjukan menu yang tersedia untuk *user*. *User* kemudian melakukan *request* pembaruan data jurnal terbaru yang akan menjalankan fungsi *web scraper*. Sistem kemudian mengalirkan data hasil *scraping* ke *database* untuk disimpan. Setelah data tersimpan di dalam *database*, sistem akan menampilkan data *update* artikel dan jurnal ilmiah terbaru. Proses kerja sistem selesai ketika user dapat melihat hasil update data artikel dan jurnal ilmiah.

b. Entity Relationship Diagram

ERD merupakan suatu model untuk menjelaskan hubungan antar data dalam basis data berdasarkan objek-objek dasar data yang mempunyai hubungan antar relasi. ERD digunakan untuk memodelkan struktur data dan hubungan antar data (Saputra, 2016). Diagram entitas pada perancangan sistem monitoring *update* artikel dan

jurnal ilmiah secara otomatis di UIN Walisongo Semarang menggunakan teknik *web scraping* ini dapat dilihat melalui Gambar 3.6.



Gambar 3.6 ERD Sistem Monitoring

Berdasarkan diagram entitas sistem monitoring ini, basis data akan mempunyai dua entitas yaitu entitas jurnal dan entitas artikel. Hubungan antara entitas jurnal dengan entitas artikel berupa hubungan *one-to-many*, dimana sebuah jurnal dapat mempunyai banyak artikel yang berbeda.

Entitas jurnal sendiri mempunyai beberapa atribut, antara lain atribut *id_jurnal*, *nama_jurnal*, *link_jurnal*, *archive_jurnal*, *jadwal_terbit*, *indeks_jurnal*, *link_indeks* dan *volume_terbaru*. Atribut *id_jurnal* merupakan atribut yang bersifat unik. Atribut *id_jurnal* ini akan menjadi *Primary Key* dari entitas jurnal. Atribut *volume_terbaru*

nantinya akan berisi informasi kapan volume terkini suatu jurnal ilmiah diterbitkan. User dapat menarik kesimpulan apakah jurnal ilmiah tersebut update atau tidak dengan membandingkannya terhadap isi atribut `jadwal_terbit` yang dimiliki oleh suatu jurnal ilmiah.

Sedangkan untuk atribut-atribut yang dimiliki entitas artikel adalah atribut `id_artikel`, `nama_artikel`, `penulis_artikel`, `asal_artikel`, `link_artikel`, dan `tanggal_terbit`. Atribut `id_artikel` akan menjadi *Primary Key* dari entitas artikel.

c. Interface Design

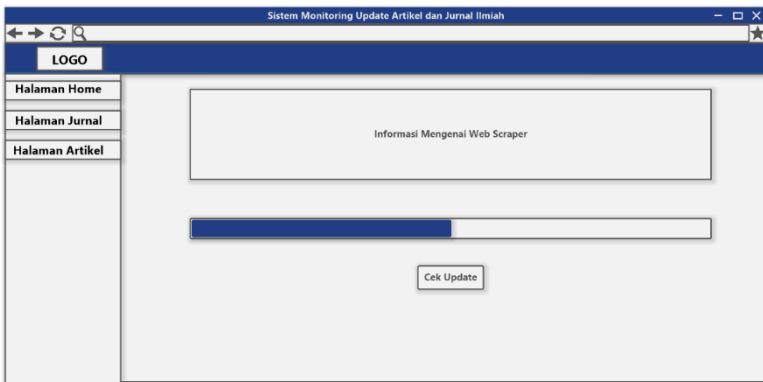
Interface Design atau desain antarmuka merupakan rancangan awal atau gambaran tampilan yang akan di kembangkan pada sistem nantinya. Hasil akhir desain dapat berubah ketika sistem selesai dikembangkan.

1) Halaman Home

Pengecekan *update* oleh sistem monitoring berjalan dengan otomatis dimana sistem akan melakukan pengecekan secara berkala, namun apabila *user* ingin melakukan pengecekan tersendiri untuk hasil yang lebih pasti, *user* dapat melakukan cek *update* secara

manual di Halaman Home. Halaman Home adalah halaman dimana *user* melakukan *request update* data artikel dan jurnal ilmiah.

Halaman ini berisi deskripsi cara melakukan *request update* data dan perintah untuk melakukan *request update* data.



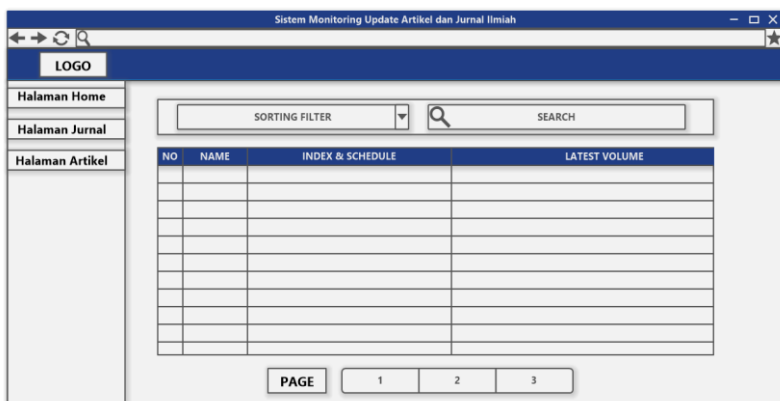
Gambar 3.7 Desain Mock-up Halaman Home

2) Halaman Jurnal

Halaman Jurnal berisi detail informasi jurnal-jurnal ilmiah. Detail informasi yang ada mencakup nama, indeks, jadwal terbit, dan volume terbaru dari suatu jurnal. *Update* atau tidaknya suatu jurnal dapat disimpulkan oleh *user* dengan membandingkan jadwal terbit jurnal tersebut dengan volume terbaru jurnal. Selain itu terdapat fungsi *filter* untuk mempermudah pencarian dan penampilan

data jurnal ilmiah. Fitur filter yang ada di halaman ini berupa filter berdasarkan kata kunci yang akan menyaring data jurnal ilmiah berdasarkan kata kunci yang dimasukkan serta filter pengurutan data berdasarkan nama, indeks, atau volume terbaru jurnal ilmiah.

Semua jurnal yang memenuhi persyaratan untuk dapat diekstrak datanya akan ditampilkan di Halaman Jurnal ini. Terakhir ada fungsi page untuk menavigasi list jurnal ilmiah yang ada.



Gambar 3.8 Desain Mock-up Halaman Jurnal

3) Halaman Artikel

Halaman ini memuat informasi data artikel ilmiah yang telah diekstraksi dari semua jurnal ilmiah yang tersedia di sistem. Di

halaman ini akan ditampilkan fungsi *filter* artikel ilmiah berdasarkan tahun dan asal jurnal. Selain itu, ada fungsi pencarian jika memang user ingin mencari artikel berdasarkan kata kunci tertentu. Selain itu, terdapat filter pengurutan data berdasarkan nama atau tanggal terbit artikel.

The mock-up shows a web application window titled "Sistem Monitoring Update Artikel dan Jurnal Ilmiah". On the left is a sidebar with a "LOGO" button and three navigation links: "Halaman Home", "Halaman Jurnal", and "Halaman Artikel". The main content area contains a filter section with "YEAR START FILTER" and "YEAR END FILTER" dropdowns, a "SORTING FILTER" dropdown, a "JOURNAL FILTER" dropdown, and a "SEARCH" button with a magnifying glass icon. Below the filters is a table with the following structure:

NO	NAME	JOURNAL ORIGIN	PUBLISH DATE

At the bottom of the main area are pagination controls: a "PAGE" button and three numbered buttons (1, 2, 3).

Gambar 3.9 Desain Mock-up Halaman Artikel

Di bawah *filter* akan ditampilkan list semua artikel ilmiah yang ada. Setiap artikel ilmiah mempunyai nama artikel, penulis artikel, link artikel, tanggal terbit artikel, dan asal jurnal artikel tersebut diekstrak. Terakhir fungsi page untuk menavigasi list artikel ilmiah yang ditampilkan.

3.2.3. Implementation and Unit Testing

Pada tahap ini dilakukan pembuatan sistem monitoring *update* artikel dan jurnal ilmiah UIN Walisongo Semarang dengan menggunakan kombinasi PHP, HTML dan CSS. Aplikasi pengedit kode Visual Studio Code digunakan untuk menunjang kegiatan ini. Selain itu, dalam pembuatan sistem ini terdapat penggunaan XAMPP sebagai penyedia layanan web dan MySQL sebagai media pengelolaan *database*. Kerangka kerja yang digunakan dalam sistem ini adalah Bootstrap guna mempermudah pembuatan tampilan website. Pengkodean skrip *web scraping* dibuat menggunakan bahasa pemrograman Python dan *library* Beautiful Soup 4.

3.2.4. Integration and System Testing

Pada tahap ini, dilakukan pengujian terhadap *website* yang dibuat dalam tahap sebelumnya. Pengujian yang dilakukan dalam penelitian ini berupa pengujian fungsional dan pengujian kegunaan. Pengujian fungsional dilakukan untuk menguji bagaimana sistem berperilaku ketika semua modul terintegrasi dan pengujian penerimaan dilakukan

untuk memastikan kebutuhan pelanggan telah terpenuhi (Wahid, 2020).

a. Functional Test

Pengujian fungsional ini menggunakan metode *Black Box Testing*. *Black Box Testing* adalah sebuah metode yang dipakai untuk menguji sebuah *software* tanpa harus memperhatikan detail *software*. Proses pengujian *black box* dengan cara mencoba program yang telah dibuat dengan mencoba memasukkan data pada setiap formnya. Pengujian ini diperlukan untuk mengetahui apakah program tersebut berjalan sesuai dengan yang dibutuhkan (Febriyanti et al., 2021).

Pada penelitian ini digunakan metode *black box testing* berupa metode *Equivalence Partitioning*. *Equivalence Partitioning* merupakan salah satu metode *black box testing* yang akan membagi domain input dari suatu program ke dalam kelas-kelas data, dimana test cases dapat diturunkan.

Tabel 3.1 Instrumen Pengujian Black Box

NO	Fungsi	Hasil yang Diharapkan
1.	Melakukan <i>request update</i> data artikel dan jurnal ilmiah	Fungsi <i>request update</i> dan <i>scraper</i> untuk mengekstrak data artikel dan jurnal ilmiah berjalan dengan benar

NO	Fungsi	Hasil yang Diharapkan
2.	Mencari jurnal ilmiah menurut kata kunci	Fungsi pencarian jurnal ilmiah berdasarkan kata kunci yang dimasukkan berjalan dengan benar
3.	Mengurutkan data jurnal ilmiah	Fungsi <i>filter sorting</i> data jurnal ilmiah berjalan dengan benar
4.	Navigasi jurnal ilmiah	Fungsi <i>page</i> untuk menavigasi list data jurnal ilmiah berjalan dengan benar
5.	Memfilter data artikel ilmiah menurut tahun	Fungsi <i>filter</i> data artikel ilmiah yang telah diekstrak berdasarkan tahun berjalan dengan benar
6.	Memfilter data artikel ilmiah menurut asal jurnal	Fungsi <i>filter</i> data artikel ilmiah berdasarkan asal jurnal ilmiah yang sama berjalan dengan benar
7.	Mencari artikel ilmiah menurut kata kunci	Fungsi pencarian artikel ilmiah berdasarkan kata kunci yang dimasukkan berjalan dengan benar
8.	Mengurutkan data artikel ilmiah	Fungsi <i>filter sorting</i> data artikel ilmiah berjalan dengan benar
9.	Navigasi artikel ilmiah	Fungsi <i>page</i> untuk menavigasi list data artikel ilmiah berjalan dengan benar

Metode *Equivalence Eartitioning* berdasarkan pada premis masukan dan keluaran dari suatu komponen yang dipartisi ke dalam kelas-kelas, menurut spesifikasi dari komponen tersebut, yang akan diperlakukan sama (ekuivalen) oleh komponen tersebut, dan dapat diasumsikan bahwa masukan yang sama akan menghasilkan respon yang sama (Rahadi & Vikasari, 2020).

Test case akan diuji dengan menggunakan metode *Equivalence Partitions*, tahapan ini dilakukan untuk mendapat *dataset*, dan kemudian *dataset* yang berisi data hasil *test case* akan dihitung keefektifannya (MZ, 2016). Instrumen pengujian *black box* pada sistem monitoring ini dapat dilihat di Tabel 3.1.

Hasil pengujian akan dihitung dengan menggunakan rumus untuk menghitung keberhasilan dari sistem sebagai berikut:

$$\text{Persentase Keberhasilan} = \frac{\text{Jumlah berhasil}}{\text{Jumlah pertanyaan}} \times 100\% \quad (3.1)$$

b. *Usability Test*

Pada pengujian ini digunakan metode *User Acceptance Test* (UAT). *User Acceptance Test* (UAT) adalah pengujian terakhir dari aplikasi yang telah dibangun sebelum aplikasi dirilis dan digunakan

langsung oleh pengguna. *User Acceptance Test* (UAT) dapat meyakinkan pengguna aplikasi yang telah dibangun apakah aplikasi dapat diterima pengguna dengan baik (Listiyani & Subhiyakto, 2021). Dalam pengujian ini peneliti memberikan beberapa pernyataan seperti yang dapat dilihat pada Tabel 3.2 (Ramadhan, 2014).

Tabel 3.2 Instrumen Pengujian User Acceptance Test

No	Pertanyaan
1.	Apakah sistem ini mudah digunakan?
2.	Apakah fitur-fitur di dalam sistem ini merespon dengan cepat dan sesuai?
3.	Apakah data artikel dan jurnal ilmiah yang diekstrak sistem ini akurat?
4.	Apakah data artikel dan jurnal ilmiah ditampilkan dengan jelas dan mudah dipahami?
5.	Apakah sistem dapat beroperasi secara stabil dan tidak mengalami crash?

Pertanyaan yang diberikan berkaitan dengan kelayakan sistem kepada beberapa responden menggunakan penilaian berupa skala Likert dengan 5 kategori yaitu SS (Sangat Setuju), S

(Setuju), N (Netral), TS (Tidak Setuju), STS (Sangat Tidak Setuju).

Skala yang digunakan untuk menguji sistem dalam pengujian UAT diukur menggunakan metode skala Likert yang memiliki 5 kategori yang telah dikonversi menjadi skor angka seperti yang dapat dilihat pada Tabel 3.3 berikut:

Tabel 3.3 Konversi Skor

Kode	Jawaban	Skor
SS	Sangat Setuju	5
S	Setuju	4
N	Netral	3
TS	Tidak Setuju	2
STS	Sangat Tidak Setuju	1

Kemudian, untuk menentukan skor total hasil pengujian menggunakan rumus berikut:

$$Skor\ total = (J_{STS} \times 1) + (J_{TS} \times 2) + (J_N \times 3) + (J_S \times 4) + (J_{SS} \times 5) \quad (3.2)$$

Keterangan:

J_{STS} : Jumlah jawaban sangat tidak setuju.

J_{TS} : Jumlah jawaban tidak setuju.

J_N : Jumlah jawaban netral.

J_S : Jumlah jawaban setuju.

J_{SS} : Jumlah jawaban sangat setuju.

Berdasarkan perhitungan diatas, maka didapatkan skor total yang akan digunakan dalam penentuan rata-rata skor. Untuk mencari rata-rata skor, digunakan rumus berikut:

$$\text{Rata - rata skor} = \frac{\text{skor total}}{\text{skor maksimal} \times \text{jumlah responden}} \quad (3.3)$$

Setelah rata-rata skor diketahui, persentase hasil kelayakan dapat dihitung. Untuk menghitung persentase tersebut, digunakan rumus berikut:

$$\text{Persentase hasil} = \frac{(\text{rata-rata skor} - \text{skor minimal})}{(\text{skor maksimal} - \text{skor minimal})} \times 100\% \quad (3.4)$$

Hasil persentase yang diperoleh dari perhitungan tersebut kemudian dikonversi sesuai kategori pada Tabel 3.4.

Tabel 3.4 Hasil Kelayakan

Hasil	Kategori
0%-20%	Sangat tidak layak
21%-40%	Tidak layak
41%-60%	Cukup Layak
61%-80%	Layak
81%-100%	Sangat Layak

3.2.5. Operation and Maintenance

Pada tahap ini, perangkat lunak yang telah selesai dikembangkan dijalankan, dan pemeliharaan dilakukan, termasuk perbaikan kesalahan yang mungkin tidak terdeteksi pada tahap sebelumnya (Wahid, 2020).

BAB IV

HASIL DAN PEMBAHASAN

Pada bab ini akan menjelaskan lebih lanjut terkait dengan hasil dan pembahasan dari sistem yang sudah dibangun. Alur penelitian yang dilakukan terdiri dari lima tahap yaitu, *requirements definition, system and software design, implementation and unit testing, integration and system testing, dan operation and maintenance.*

4.1. Requirements Definiton

Tahap ini diawali dengan proses analisis cara kerja *web scraping* yang akan dibuat dan dilanjutkan dengan pembuatan daftar kebutuhan fungsional dan kebutuhan non-fungsional sistem monitoring ini.

4.1.1. Analisis Proses Scraping Data

Pada proses ini, terdapat tiga hal yang perlu dilakukan terlebih dahulu sebelum pembuatan *web scraper* dapat dilakukan. Ketiga hal tersebut adalah penyaringan *website* jurnal ilmiah yang memenuhi persyaratan, analisis struktur *website* jurnal ilmiah yang telah disaring, dan pengembangan pola *scraping* pada *web scraper*.

a. Penyaringan Website

Proses *web scraping* memerlukan target yang jelas agar dapat berjalan dengan lancar. Sistem monitoring *update* artikel dan jurnal

ilmiah secara otomatis di UIN Walisongo Semarang menggunakan teknik *web scraping* ini mempunyai target *scraping* berupa *website* jurnal ilmiah. Oleh karena itu, perlu dilakukan penentuan apakah suatu *website* jurnal ilmiah memenuhi persyaratan untuk dimasukkan ke dalam sistem monitoring.

Agar suatu jurnal ilmiah dapat dimasukkan ke dalam sistem monitoring, jurnal tersebut perlu memenuhi beberapa persyaratan. Persyaratan-persyaratan yang perlu dipenuhi suatu jurnal ilmiah adalah jurnal ilmiah termasuk jurnal ilmiah di UIN Walisongo Semarang, mempunyai setidaknya satu volume yang telah diterbitkan, dan dapat diakses oleh semua orang.

Berdasarkan Lampiran 1, dapat disimpulkan bahwa *website* jurnal ilmiah yang memenuhi ketiga persyaratan yang ditetapkan berjumlah 45 dari 46 *website* jurnal ilmiah.

b. Analisis Struktur Website

Setelah mendapatkan data *website* jurnal ilmiah yang memenuhi persyaratan untuk dimasukkan ke dalam sistem monitoring, perlu

dilakukan analisis terhadap masing-masing *website* jurnal ilmiah sebelum proses pembuatan *web scraper* dilakukan.

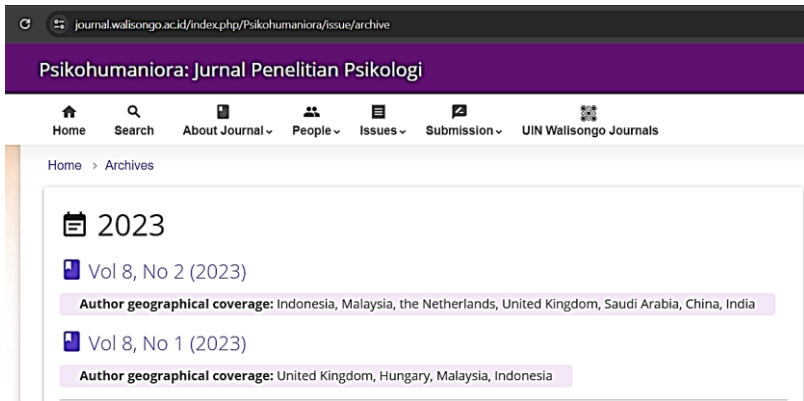
Penelitian ini menggunakan metode *web scraping* HTML Parser, sehingga proses analisis yang dilakukan terhadap *website* jurnal ilmiah adalah analisis struktur HTML *website* tersebut. Analisis digunakan untuk mencari data artikel dan jurnal ilmiah yang ingin diekstrak, yaitu volume terbaru jurnal, nama artikel, link artikel dan tanggal terbit artikel.

Proses analisis struktur HTML *website* untuk mendapatkan data artikel dan jurnal ilmiah tersebut dibagi menjadi beberapa langkah proses.

1) Penentuan Archive Link

Langkah pertama adalah penentuan *archive link* tempat daftar volume terbitan (*issue*) disimpan. Penentuan *archive link* ini diperlukan untuk mengetahui dimana setiap volume terbitan disimpan. Setiap *website* jurnal ilmiah memiliki *archive link* yang berbeda, oleh karena itu perlu dilakukan penentuan *archive link* dari masing-masing

jurnal ilmiah terlebih dahulu. Berikut ini adalah contoh penentuan *archive link* dari salah satu *website* jurnal ilmiah yang memenuhi persyaratan sistem monitoring.



Gambar 4.1 Contoh Archive Link Jurnal Psikohumaniora

Pada *website* jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi, didapatkan hasil *archive link* yaitu “<https://journal.walisongo.ac.id/index.php/Psikohumaniora/issue/archive>”. Daftar volume terbitan (*issue*) dari jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi dapat dilihat di halaman archive ini.

Berdasarkan contoh pada Gambar 4.1, dilakukan penentuan *archive link* dari masing-masing *website* jurnal ilmiah. Hasil penentuan

archive link dari masing-masing jurnal ilmiah dapat dilihat pada Lampiran 2.

2) Penentuan Struktur HTML Volume Terbaru

Langkah kedua adalah penentuan struktur HTML tempat volume terbitan terbaru dari setiap jurnal ilmiah disimpan. Penentuan struktur HTML volume terbaru di setiap jurnal diperlukan karena volume terbitan tersebut merupakan tempat penyimpanan data artikel ilmiah yang akan diekstrak nantinya.

Setiap *website* jurnal ilmiah dapat memiliki struktur HTML yang berbeda untuk bagian volume terbaru, oleh karena itu perlu dilakukan penentuan struktur HTML volume terbaru dari masing-masing jurnal ilmiah terlebih dahulu. Berikut ini adalah contoh penentuan struktur HTML volume terbaru dari jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi.

```

<div id="issue-766" style="clear:left;">
  <div class="list">
    <div class="issueDescription"> == $0
      <h4>
        <a href="https://journal.walisongo.ac.id/index.php/Psikohumaniora/issue/view/766">
          <i class="material-icons h4">book</i>
          " Vol 8, No 2 (2023)"
        </a>
      </h4>
      <div class="issuecoverage-a">
        <p>
          <strong>Author geographical coverage:</strong>
          "&nbsp;Indonesia, Malaysia,&nbsp;the Netherlands, United Kingdom,
          Saudi Arabia, China,&nbsp;India"
        </p>
      </div>
    </div>
  </div>
</div>

```

Gambar 4.2 Struktur HTML Volume Terbaru dari Jurnal Psikohumaniora

Pada Gambar 4.2 di pembahasan sebelumnya, *website* jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi diketahui mempunyai volume terbitan terbaru bernama "Vol 8, No 2 (2023)". Berdasarkan nama volume terbaru tersebut, pada Gambar 4.2 ditemukan bahwa volume terbaru jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi berada diantara struktur HTML dengan tag `<div class="issueDescription"></div>`.

Di antara tag `<div class="issueDescription"></div>`, terdapat

beberapa *tag* yang menyimpan data yang perlu diekstrak oleh *web scraper* yang akan dikembangkan. Data-data tersebut berupa data elemen induk (*parent element*) volume terbaru, nama volume terbaru, dan *link* menuju isi volume terbitan (artikel ilmiah) dari jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi.

Data *parent element* volume terbaru disimpan di dalam *tag* elemen `<div, class = "issueDescription">`. Untuk data nama volume terbaru disimpan di dalam *tag* elemen `<a>`. Sedangkan data *link* menuju isi volume terbaru disimpan di dalam atribut `<href>` dari *tag* elemen `<a>`. Setelah mendapatkan data-data *tag* elemen dan atribut tersebut, proses analisis struktur HTML untuk volume terbaru jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi dianggap selesai.

Berdasarkan contoh tersebut, dilakukan penentuan struktur HTML volume terbaru dari masing-masing *website* jurnal ilmiah. Hasil penentuan struktur HTML volume

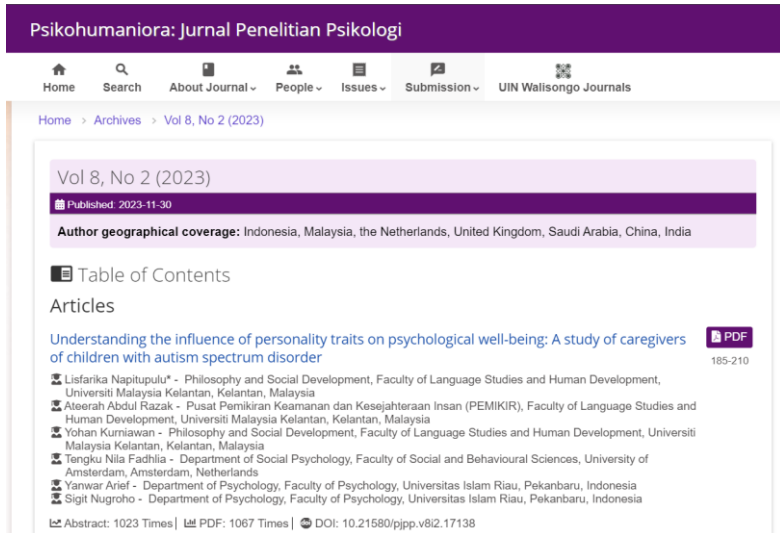
terbaru dari masing-masing jurnal ilmiah dapat dilihat pada Lampiran 3.

3) Penentuan Struktur HTML Artikel Ilmiah

Setelah mendapatkan data volume terbaru, langkah selanjutnya adalah penentuan struktur HTML artikel ilmiah dari setiap volume terbaru jurnal ilmiah. Penentuan struktur HTML artikel ilmiah diperlukan untuk mendapatkan data artikel ilmiah yang akan diekstrak. Data-data artikel ilmiah yang perlu diekstrak pada proses ini adalah data tanggal terbit volume terbaru tempat artikel ilmiah diterbitkan, nama, penulis, dan *link* artikel ilmiah tersebut.

Berikut ini adalah contoh penentuan struktur HTML dari jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi.

Pada Gambar 4.3, dapat dilihat bahwa volume terbaru jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi mempunyai tanggal terbit yaitu 2023-11-30.



Gambar 4.3 Tampilan Halaman Artikel di Jurnal Ilmiah Psikhumaniora

Selain itu, terdapat daftar isi artikel yang diterbitkan pada volume tersebut dimana daftar artikel ilmiah dapat dilihat. Selanjutnya dilakukan analisis untuk mendapatkan struktur HTML tanggal terbit volume terbaru jurnal ilmiah terlebih dahulu.

Pada Gambar 4.4 ditemukan bahwa tanggal terbit volume terbaru jurnal ilmiah Psikhumaniora: Jurnal Penelitian Psikologi berada diantara struktur HTML dengan tag `<div class= "infoissue"></div>`.

Pada Gambar 4.5 ditemukan bahwa nama, penulis dan *link* artikel ilmiah berada di antara struktur HTML dengan *tag* `<td class = "tocArticleTitleAuthors"> </div>`. Oleh karena itu, *parent element* untuk nama, penulis dan *link* artikel ilmiah adalah tag `<td class = "tocArticleTitleAuthors"> </div>`.

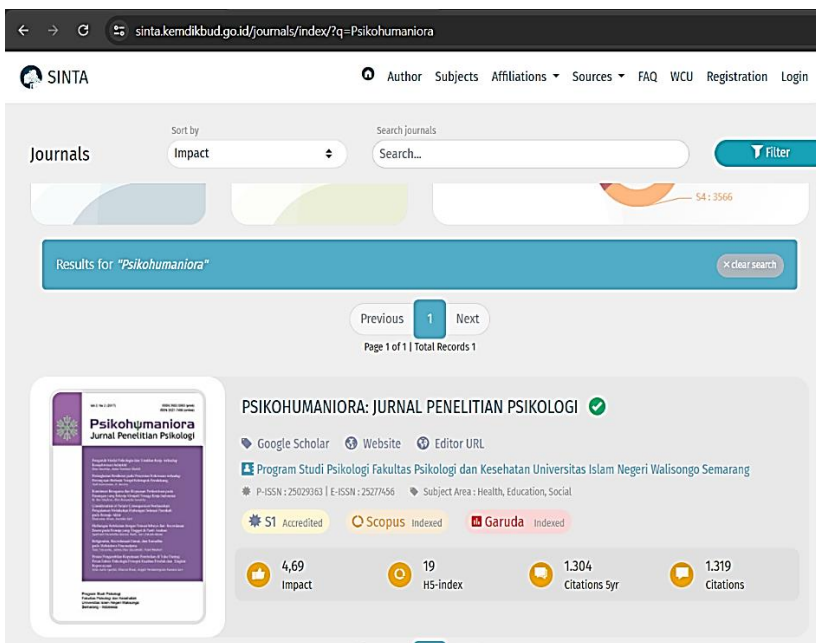
Data nama artikel ilmiah disimpan di dalam *tag* elemen `<a>`. Sedangkan data *link* artikel ilmiah disimpan di dalam atribut `<href>` dari *tag* elemen `<a>`. Data penulis artikel sendiri disimpan pada *tag* yang berbeda, yaitu *tag* elemen `<div id = "authorString">`. Setelah mendapatkan data-data *tag* elemen dan atribut tersebut, proses analisis struktur HTML untuk artikel ilmiah dari volume terbaru jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi dianggap selesai.

Berdasarkan contoh tersebut, dilakukan penentuan struktur HTML artikel ilmiah dari masing-masing *website* jurnal ilmiah. Hasil penentuan struktur HTML

artikel ilmiah dari masing-masing jurnal ilmiah dapat dilihat pada Lampiran 4.

4) Penentuan Struktur HTML Indeks SINTA Jurnal Ilmiah

Langkah keempat yang perlu dilakukan adalah menentukan akreditasi setiap jurnal ilmiah di situs Indeks SINTA. Data-data yang perlu ditentukan strukturnya adalah data link jurnal pada indeks SINTA, *tag parent element* akreditasi, dan *tag elemen* akreditasi jurnal disimpan.



Gambar 4.6 Tampilan Pencarian Jurnal Ilmiah di Index SINTA

Link jurnal pada indeks SINTA berperan penting untuk memastikan apakah suatu jurnal telah terakreditasi SINTA atau belum. Menggunakan jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi sebagai contoh, dapat dilihat tampilan jurnal ilmiah tersebut pada pencarian jurnal pada situs Indeks SINTA. Dari gambar dapat disimpulkan bahwa jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi mempunyai link jurnal yaitu *"<https://sinta.kemdikbud.go.id/journals/index/?q=Psikohumaniora>"*.

Selain itu, dapat dilihat pada gambar tersebut bahwa akreditasi jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi di Indeks SINTA terakreditasi sebagai S1 (Sinta 1). Langkah selanjutnya adalah menentukan struktur HTML yang memuat akreditasi jurnal ilmiah tersebut.

Pada Gambar 4.7, dapat dilihat bahwa nilai akreditasi S1 terdapat pada *tag* elemen `<span class = "num-stat accredited">`. Berdasarkan contoh penentuan struktur

HTML untuk akreditasi jurnal ilmiah Psikohumaniora: Jurnal Penelitian Psikologi tersebut, masing-masing jurnal ilmiah yang memenuhi persyaratan sistem akan dilakukan penentuan struktur akreditasinya pada situs Indeks SINTA.

```
<div class="stat-prev mt-2">
  <span class=" num-stat accredited">
    <a href="#1">
      <i class="el el-certificate">
        ::before
      </i>
      " S1 " == $0
      <span class="text-stat ml-1">Accredited</span>
    </a>
  </span>
  <span class="num-stat scopus-indexed ml-3">⋯</span>
  <a href="https://garuda.kemdikbud.go.id/journal/view/8645">⋯</a>
</div>
```

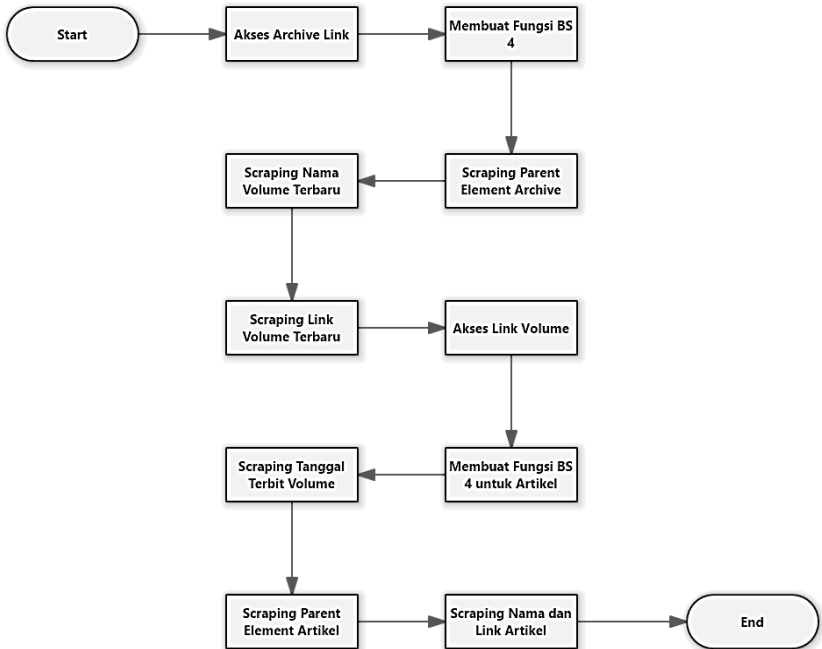
Gambar 4.7 Struktur HTML Akreditasi Jurnal Ilmiah

Hasil penentuan struktur HTML akreditasi pada Indeks SINTA dari masing-masing jurnal ilmiah dapat dilihat pada Lampiran 5.

b. Pengembangan Pola Scraping Data

Setelah mendapatkan data *website* jurnal ilmiah serta struktur HTML data artikel dan jurnal ilmiah, perlu dilakukan pengembangan pola *scraping* data agar ekstraksi data yang dilakukan

dapat berjalan dengan benar. Proses pengembangan pola ini dilakukan dengan cara membuat urutan data mana yang perlu diekstrak terlebih dahulu dan menentukan *method* mana yang perlu digunakan dari *library* Beautiful Soup 4.



Gambar 4.8 Pola Umum Proses Web Scraping

Pada penelitian ini, pola umum proses *web scraping* yang dilakukan pada setiap jurnal ilmiah dapat dilihat pada Gambar 4.8.

Proses pertama dalam pola umum *web scraping* yang dilakukan adalah mengakses *archive*

link. *Archive link* perlu diakses terlebih dahulu untuk mendapatkan respon dari suatu jurnal ilmiah. Jika suatu jurnal ilmiah memberikan respon berupa *error*, maka *web scraper* akan melanjutkan proses *scraping* ke *website* selanjutnya.

Proses kedua yang dilakukan adalah pembuatan variabel Python yang memuat *instance* dari *class* Beautiful Soup 4. Pada proses ini metode *scraping* yaitu *HTML Parser* dideklarasikan.

Proses ketiga yang dilakukan adalah *scraping parent element* volume terbaru. Karena di halaman *archive* jurnal ilmiah mempunyai lebih dari satu volume terbitan, maka proses *scraping* yang dilakukan harus mengekstrak hanya volume terbarunya saja. Proses *scraping* dengan tujuan mendapatkan hanya satu hasil data dapat dilakukan dengan menggunakan *method find* dari *library* Beautiful Soup 4.

Setelah mendapatkan *parent element* volume terbaru, proses *web scraping* kemudian masuk ke bagian selanjutnya, yaitu *scraping* nama volume terbaru. Proses *scraping* nama volume terbaru dari *parent element*-nya adalah dengan cara

menggunakan *method* `get_text` kepada *tag* elemen yang menyimpan data nama volume.

Proses kelima adalah *scraping link* volume terbaru tersebut. Tujuan proses ini adalah agar mendapatkan *link* yang bisa memberikan akses kepada daftar artikel ilmiah dari volume terbitan terbaru. Proses ini dilakukan dengan menggunakan *method* `get('href')` kepada *tag* elemen yang memuat *link* volume terbaru.

Proses selanjutnya adalah mengakses *link* volume terbaru yang telah diekstrak pada proses sebelumnya.

Ketika *link* volume terbaru telah diakses, maka proses ketujuh dapat dilakukan. Pada proses ini perlu adanya pembuatan variabel Python berisi *instance* dari *class* Beautiful Soup 4 yang baru. Pembuatan variabel ini diperlukan agar proses ekstraksi data artikel ilmiah tidak tercampur dengan ekstraksi data volume terbaru di proses sebelumnya.

Proses kedelapan yang dilakukan adalah ekstraksi tanggal terbit dari volume terbitan terbaru. Hal ini dilakukan dengan menggunakan *method* `find` kepada *tag* elemen yang memuat tanggal terbit

karena data yang perlu diekstrak hanya berjumlah satu data saja.

Proses kesembilan adalah ekstraksi *parent element* artikel ilmiah. Suatu volume jurnal ilmiah dapat memuat lebih dari satu artikel ilmiah di dalamnya. Untuk mendapatkan lebih dari satu data artikel ilmiah, *method* `find_all` digunakan pada proses ini. Dengan menggunakan *method* `find_all`, *scraper* akan mengekstrak setiap data yang memiliki *tag* elemen yang sesuai dengan *parent element* artikel ilmiah.

Proses terakhir dalam proses *web scraping* yang dilakukan adalah melakukan ekstraksi nama dan *link* artikel ilmiah. *Method* `get_text` digunakan untuk mendapatkan data nama artikel dan *method* `get('href')` digunakan untuk mendapatkan data *link* artikel ilmiah.

Pola umum proses *web scraping* untuk akreditsai jurnal ilmiah pada indeks SINTA sendiri hampir sama dengan pola pada proses *web scraping* artikel dan jurnal ilmiah. Perbedaan pola *scraping* pada akreditasi SINTA terletak pada jumlah variabel Python berisi *instance* dari *class* Beautiful Soup 4 saja dimana hanya memerlukan satu variabel saja.

4.1.2. Daftar Kebutuhan

Daftar kebutuhan sistem monitoring dibagi menjadi dua bagian yaitu, kebutuhan fungsional dan kebutuhan non fungsional.

a. Kebutuhan Fungsional

Daftar kebutuhan fungsional dari sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang dapat dilihat pada Tabel 4.1.

Tabel 4.1 Kebutuhan Fungsional Sistem

No	Kebutuhan
1	Sistem harus dapat mengidentifikasi dan mengekstrak data artikel dan jurnal ilmiah di UIN Walisongo Semarang.
2	Sistem harus dapat mengidentifikasi dan mengekstrak data akreditasi jurnal ilmiah UIN Walisongo Semarang pada indeks SINTA.
3	Sistem harus dapat menyimpan data artikel dan jurnal ilmiah dalam <i>database</i> .
4	Sistem harus dapat menampilkan informasi artikel dan jurnal ilmiah.
5	Sistem harus dapat menyediakan fitur <i>filter</i> untuk mempermudah dalam menampilkan data artikel dan jurnal ilmiah.
6	Sistem harus dapat melakukan ekstraksi data artikel dan jurnal ilmiah secara otomatis dalam jangka waktu tertentu.

b. Kebutuhan Non-fungsional

Daftar kebutuhan non-fungsional dari sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang menggunakan teknik *web scraping* dapat dilihat pada Tabel 4.2.

Tabel 4.2 Kebutuhan Non-Fungsional Sistem

No	Kebutuhan
1	Sistem harus dapat digunakan dengan mudah oleh pengguna berdasarkan perhitungan <i>usability test</i> .
2	Sistem harus dapat diakses oleh pengguna kapan saja.

4.2. System and Software Design

Tahap *system and software design* pada penelitian ini dilakukan dengan melakukan perancangan *database* sistem monitoring ini. Sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang menggunakan *database* untuk menyimpan hasil data hasil proses *web scraping*. *Database* yang digunakan dibuat dengan menggunakan sistem manajemen *database* (DBMS) berupa MySQL.

Database pada sistem monitoring ini mempunyai dua tabel yaitu tabel jurnal dan tabel artikel. Tabel-tabel tersebut memuat data atribut yang telah dirancang dalam desain ERD pada Gambar 3.6 sebelumnya.

Gambar 4.9 menunjukkan perancangan tabel jurnal pada MySQL.

#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra	Action
☐	1 id_jurnal	int(10)			No	None			Change Drop More
☐	2 nama_jurnal	varchar(100)	utf8mb4_general_ci		No	None			Change Drop More
☐	3 link_jurnal	varchar(100)	utf8mb4_general_ci		No	None			Change Drop More
☐	4 jadwal_terbit	varchar(100)	utf8mb4_general_ci		No	None			Change Drop More
☐	5 archive_jurnal	varchar(100)	utf8mb4_general_ci		No	None			Change Drop More
☐	6 volume_terbaru	varchar(100)	utf8mb4_general_ci		Yes	NULL			Change Drop More
☐	7 link_indeks	varchar(1000)	utf8mb4_general_ci		No	None			Change Drop More
☐	8 indeks_jurnal	varchar(10)	utf8mb4_general_ci		No	None			Change Drop More

Gambar 4.9 Rancangan Tabel Jurnal

Tabel 4.3 di bawah ini berisi detail dan contoh data dari desain tabel jurnal yang telah dibuat.

Tabel 4.3 Desain Tabel Jurnal

Kolom	Type Data	Keterangan	Contoh Data
id_jurnal	Int	Primary Key	1
nama_jurnal	varchar	-	Psikohumaniora: Jurnal Penelitian Psikologi
link_jurnal	varchar	-	https://journal. walisongo.ac.id/i ndex.php/Psikoh umaniora

Kolom	Tipe Data	Keterangan	Contoh Data
jadwal_terbit	varchar	-	<i>May and November</i>
archive_jurnal	varchar	-	https://journal.walisongo.ac.id/index.php/Psikohumaniora/issue/archive
volume_terbaru	varchar	-	Vol 8, No 2 (2023) Diterbitkan Pada: 2023-11-30
link_indeks	varchar	-	https://sinta.kemdikbud.go.id/journals/index/?q=Psikohumaniora
indeks_jurnal	varchar	-	S1

Tabel 4.3 merupakan desain tabel untuk menyimpan data jurnal ilmiah yang akan dilakukan scraping. Atribut `link_jurnal` mempunyai data yang berbeda dari atribut `archive_jurnal`. Data yang dimasukkan ke dalam atribut `link_jurnal` adalah data *link* menuju halaman beranda suatu

jurnal ilmiah. Sedangkan data yang dimasukkan ke dalam atribut `archive_jurnal` adalah data *link* yang memuat halaman kumpulan volume terbitan jurnal ilmiah.

Selanjutnya, Gambar 4.10 menunjukkan perancangan tabel artikel pada MySQL.

#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra	Action
☐	1 <code>id_artikel</code>	<code>int(255)</code>			No	<i>None</i>		<code>AUTO_INCREMENT</code>	Change Drop More
☐	2 <code>nama_artikel</code>	<code>varchar(1000)</code>	<code>utf8mb4_general_ci</code>		No	<i>None</i>			Change Drop More
☐	3 <code>asal_artikel</code>	<code>int(10)</code>			No	<i>None</i>			Change Drop More
☐	4 <code>link_artikel</code>	<code>varchar(1000)</code>	<code>utf8mb4_general_ci</code>		No	<i>None</i>			Change Drop More
☐	5 <code>tanggal_terbit</code>	<code>date</code>			No	<i>None</i>			Change Drop More
☐	6 <code>penulis_artikel</code>	<code>varchar(10000)</code>	<code>utf8mb4_general_ci</code>		No	<i>None</i>			Change Drop More

Gambar 4.10 Rancangan Tabel Artikel

Tabel 4.4 di bawah ini berisi detail dan contoh data dari desain tabel artikel yang telah dibuat.

Tabel 4.4 Desain Tabel Artikel

Kolom	Type Data	Keterangan	Contoh Data
<code>id_artikel</code>	Int	Primary Key	1
<code>nama_artikel</code>	varchar	-	Understanding the influence of personality traits on psychological well-being: A study of caregivers of children with

Kolom	Tipe Data	Keterangan	Contoh Data
			autism spectrum disorder
asal_artikel	int	-	1
link_artikel	varchar	-	https://journal.walisongo.ac.id/index.php/Psikohumana/article/view/17138
tanggal_terbit	date	-	2023-11-30
penulis_artikel	varchar	-	Lisfarika Napitupulu* - Philosophy and Social Development, Faculty of Language Studies and Human Development, Universiti Malaysia Kelantan, Kelantan, Malaysia.

4.3. Implementation and Unit Testing

Tahap ini merupakan tahap penerapan analisis dan perancangan desain yang dilakukan pada tahap-tahap sebelumnya. Tahap ini dibagi menjadi dua proses yaitu, implementasi *web scraper* dan implementasi *website*.

Tahap implementasi ini dilakukan memakai komputer dengan *prosesor* core-i5, RAM (memory) 8GB, dan sistem operasi Windows 10. Text editor Visual Studio Code digunakan untuk mempermudah pengembangan sistem monitoring ini.

4.3.1. Implementasi Web Scraper

Pada bagian ini akan dijelaskan implementasi *web scraper* pada artikel dan jurnal ilmiah di UIN Walisongo Semarang berdasarkan proses perancangan pola umum *scraping* data sebelumnya dan direalisasikan pada file program dengan ekstensi *.py.

Gambar 4.11 merupakan *script* Python untuk mengakses database *db_jurnal* yang berasal dari MySQL. Setelah berhasil mengakses database, langkah selanjutnya adalah mengambil semua data *archive link* yang diletakkan pada tabel *jurnal* di kolom *archive_jurnal*. Data *archive link* tersebut kemudian diletakkan pada variabel *archive_links* untuk memudahkan akses.

```

import requests
from bs4 import BeautifulSoup
import mysql.connector # type: ignore
import datetime
import time

# Connect to database
db = mysql.connector.connect(
    host="localhost",
    username="root",
    password="",
    database="db_jurnal"
)

cursor = db.cursor()

cursor.execute("SELECT archive_jurnal FROM jurnal")
archive_links = [x[0] for x in cursor.fetchall()]

```

Gambar 4.11 Script Python untuk Koneksi Database

Pada Gambar 4.12, langkah selanjutnya yang dilakukan adalah pembuatan fungsi Jurnal dengan argumen `archive_links`. Fungsi Jurnal ini memuat proses *scraping* data volume terbaru setiap jurnal ilmiah. Langkah awal yang dilakukan fungsi ini adalah melakukan iterasi melalui setiap *archive link* yang disimpan di variabel `archive_links`. Hal ini dilakukan untuk memastikan setiap *archive link* melalui proses *scraping* data.

```

def Jurnal(archive_links):
    link_number = 1
    for i, base_url in enumerate(archive_links):
        page = requests.get(base_url)
        soup = BeautifulSoup(page.content, "html.parser")

        issue_element = soup.find("div", class_=("issueDescription", "issueDescriptionImage"))
        if not issue_element:
            print(f"Issue element not found for {base_url}, skipping.")
            continue

        issue_link = issue_element.find("a")["href"]

        latest_volume = issue_element.find("a").text.replace("book ", "")

        update_query = """
            UPDATE jurnal
            SET volume_terbaru = %s
            WHERE archive_jurnal = %s
        """
        cursor.execute(update_query, (latest_volume, base_url))
        db.commit()

        print(f"Latest volume for {base_url}: {latest_volume}")

        sql_get_id = "SELECT id_jurnal FROM jurnal WHERE archive_jurnal = (%s)"
        data = (base_url,)
        cursor.execute(sql_get_id, data)
        archive_jurnal_id = cursor.fetchone()[0]

        scrape_issue(issue_link, archive_jurnal_id)

    link_number += 1

```

Gambar 4.12 Script Python untuk Fungsi Jurnal

Langkah selanjutnya adalah mengakses *archive link* dan mengambil struktur HTML halaman *archive* menggunakan *library* requests dari setiap *archive link*. Kemudian sebuah variabel Python yang memuat *instance class* BeautifulSoup dan metode parser yang

digunakan dibuat. Variabel tersebut pada *script* ini dinamakan variabel soup dengan metode *parser* HTML Parser. Variabel soup ini kemudian digunakan untuk mengekstrak data volume terbaru dari setiap *archive link*.

Langkah *scraping* data berikutnya adalah mencari elemen yang mengandung informasi tentang volume terbaru dan tautan ke isu terbaru dari jurnal tersebut. Langkah ini dilakukan dengan cara menerapkan *method* find pada *parent element* volume terbaru yang telah dianalisis di pembahasan sebelumnya.

Setelah menemukan *parent element* volume terbaru menggunakan metode find dengan argumen *tag* dan *class* volume terbaru, langkah selanjutnya adalah mengekstrak data nama volume terbaru dan *link* volume terbaru. Data nama volume terbaru didapatkan dengan menggunakan *method* find dan *get_text* pada tag <a> dari *parent element* volume terbaru. Selanjutnya nama volume terbaru yang telah diekstrak disimpan ke dalam database menggunakan *update_query*.

Untuk mengekstrak data *link* volume terbaru, digunakan *method* find dan *get('href')* pada tag <a>. Data *link* ini kemudian disimpan dalam variabel

sementara berupa *issue_link* untuk digunakan dalam proses *scraping* artikel ilmiah bersama dengan variabel yang memuat *id_jurnal* dari *archive link* yang sedang diekstrak. Penentuan nilai *id_jurnal* dilakukan dengan membandingkan *archive link* yang sedang diekstrak dengan data *archive link* yang ada di database. Apabila data *archive link* yang diekstrak sama dengan data *archive link* di database, maka data *id_jurnal* untuk *archive link* tersebut diambil.

Langkah selanjutnya yang perlu dilakukan adalah memanggil fungsi *scrape_issue* dengan data *link* volume terbaru dan *id_jurnal* yang sesuai. Gambar 4.13 memuat *script* untuk proses *scraping* data artikel ilmiah yang ada di dalam volume terbaru suatu jurnal ilmiah. Fungsi *scrape_issue* pertama-tama akan mengakses *link* volume terbaru yang telah diekstrak sebelumnya menggunakan *library requests*. Hasil akses *link* volume terbaru disimpan pada variabel yang berbeda dari variabel pada fungsi *Jurnal*. Sama seperti variabel hasil akses *link* volume terbaru yang berbeda dari fungsi sebelumnya, variabel yang memuat *instance* dari *class BeautifulSoup* juga berbeda. Hal ini bertujuan untuk membedakan proses ekstraksi data volume terbaru dan data artikel ilmiah.

```

def scrape_issue(issue_link, archive_jurnal_id):
    try:
        page_article = requests.get(issue_link)
        soup2 = BeautifulSoup(page_article.content, "html.parser")
        tanggal_terbit = soup2.find("div", class_="infoissue").text.replace(
            " Published: ", ""
        )
        date = datetime.datetime.strptime(tanggal_terbit, '%Y-%m-%d').date()
        artikel_parent = soup2.find_all("td", class_="tocArticleTitleAuthors")
        num_articles = len(artikel_parent)
        for detail in artikel_parent:
            title_element = detail.find("div", class_="tocTitle")
            title_link = title_element.find("a") if title_element else None
            nama_artikel = title_link.text.strip() if title_link else "Front Matter"
            link_artikel = title_link["href"] if title_link else "#"
            authors = []
            author_elements = detail.find_all("div", id="authorString")
            for author_element in author_elements:
                author_text = author_element.text.strip()
                author_name = re.sub(r"\s+", " ", author_text)
                formatted_author = f"{author_name}."
                authors.append(formatted_author)

            penulis_artikel = "\n".join(authors)

            insert_query = """
                INSERT INTO artikel (nama_artikel, asal_artikel, link_artikel, tanggal_terbit, penulis_artikel)
                VALUES (%s, %s, %s, %s, %s)
            """
            cursor.execute(insert_query, (nama_artikel, archive_jurnal_id, link_artikel, date, penulis_artikel))
            db.commit()
            print(f"{nama_artikel} {archive_jurnal_id} {link_artikel} {date} {penulis_artikel}")
        print("")
        print(f"Number of articles in this issue: {len(artikel_parent)}")
        print("-" * 50)
    except requests.exceptions.RequestException as e:
        print(f"Error fetching issue link: {e}")
    pass

```

Gambar 4.13 Script Python untuk Fungsi Scrape Issue

Langkah yang dilakukan selanjutnya adalah proses *scraping* tanggal terbit volume terbaru. Proses ini dilakukan menggunakan *method* `find` dan `get_text` pada *tag* `<div, class = "infoissue">`. Karena hasil *scraping* tanggal terbit masih berupa tipe data *string*, maka perlu dilakukan konversi dari tipe data *string* ke dalam tipe data *date*. Proses konversi ini memanfaatkan *library* `datetime`. Hasil konversi tanggal terbit ke tipe data *date* disimpan di dalam variabel *date*.

Setelah itu, dilakukan proses *scraping parent element* untuk semua artikel ilmiah yang ada. Proses *scraping* ini memanfaatkan *method* `find_all` pada *tag* `<td, class = "tocArticleTitleAuthors">`. *Method* `find_all` digunakan untuk mengekstrak data yang berjumlah lebih dari satu data. Setelah semua *parent element* artikel ilmiah diekstrak, dilakukanlah iterasi pada setiap *parent element* tersebut untuk mendapatkan data nama, penulis, dan link artikel ilmiah.

Data nama artikel ilmiah diekstrak menggunakan *method* `find` dan `get_text` pada *tag* `<a href>`. Sedangkan data link artikel didapatkan menggunakan *method* `find` dan `get('href')`. Disisi lain, karena suatu artikel ilmiah dapat mempunyai lebih dari satu penulis, maka data

penulis artikel ilmiah diekstrak dengan menggunakan *method* `find_all` pada *tag* `<div id = "authorString">`. Setelah menemukan semua data penulis artikel ilmiah, data tersebut digabungkan menjadi satu menggunakan *method* `join` ke dalam variabel `penulis_artikel`.

Proses terakhir adalah menyimpan data artikel ilmiah yang telah diekstrak ke dalam *database* sesuai dengan hasil *scraping* yang telah dilakukan. Proses *scraping* berakhir ketika semua *archive link* telah menjalani ekstraksi data. Setelah *script web scraper* berhasil dibuat, *script* tersebut disimpan dengan nama *file* `JurnalScraper.py`.

Setelah implementasi *web scraper* pada artikel dan jurnal ilmiah di UIN Walisongo selesai, selanjutnya akan dilakukan implementasi *web scraper* pada akreditasi jurnal ilmiah pada indeks SINTA. Proses implementasi *web scraper* akreditasi jurnal ilmiah pada indeks SINTA sendiri mempunyai beberapa tahapan. Proses pertama adalah mengakses ke *database*. *Database* yang digunakan untuk menjelaskan implementasi *web scraper* akreditasi jurnal ilmiah menggunakan *database* bernama `monito_29` untuk membedakan dengan penjelasan sebelumnya. Gambar 4.14 menampilkan proses akses *database*,

```

import requests
from bs4 import BeautifulSoup
import mysql.connector
import datetime
import time

db = mysql.connector.connect(
    host="localhost",
    username="monito29_admin",
    password="monito29_admin",
    database="monito29_db",
    charset= "utf8"
)

cursor = db.cursor()

cursor.execute("SELECT link_indeks FROM jurnal")
search_urls = [x[0] for x in cursor.fetchall()]

```

Gambar 4.14 Script Python untuk Mengakses Database Akreditasi

Apabila pada implementasi *web scraper* sebelumnya menggunakan data *archive link*, pada implementasi *web scraper* akreditasi jurnal akan menggunakan data *link indeks* yang memuat akses menuju situs indeks SINTA. Data *link indeks* tersebut dimasukkan ke dalam variabel `search_urls` untuk memudahkan akses. Proses selanjutnya adalah membuat dua fungsi yang berbeda. Fungsi pertama merupakan fungsi `scrape_index` dengan argumen

search_urls. Fungsi ini bertujuan untuk melakukan iterasi melalui semua *link indeks* yang disimpan pada variabel search_urls untuk memastikan semua *link indeks* melalui proses scraping data.

```
def scrape_index(search_urls):
    for i, url in enumerate(search_urls):
        print(f"Link no.{i+1}: ", end='')

        sql_get_id = "SELECT id_jurnal FROM jurnal WHERE link_indeks = (%s)"
        data = (url,)
        cursor.execute(sql_get_id, data)
        indeks_jurnal_id = cursor.fetchone()[0]

        journals = scrape_journals_uin_walisongo(url, indeks_jurnal_id)
        if journals:
            for journal in journals:
                if journal['accreditation']:
                    sql = "UPDATE jurnal SET indeks_jurnal = %s WHERE id_jurnal = %s"
                    data = (journal['accreditation'], indeks_jurnal_id)
                    cursor.execute(sql, data)
                    db.commit()

                    print(f"Accreditation: {journal['accreditation']}", "id: ", indeks_jurnal_id)
                else:
                    sql = "UPDATE jurnal SET indeks_jurnal = %s WHERE id_jurnal = %s"
                    not_found = "Not Found"
                    data = (not_found, indeks_jurnal_id)
                    cursor.execute(sql, data)
                    db.commit()

                    print("Index not found.", "id: ", indeks_jurnal_id)
```

Gambar 4.15 Script Fungsi Iterasi Link Indeks

Gambar 4.15 merupakan script fungsi scrape_index. Langkah selanjutnya adalah

mendapatkan id_jurnal dari setiap *link* yang diekstrak agar tidak terjadi kesalahan dalam proses penyimpanan hasil ekstraksi yang didapat. Setelah mendapatkan id_jurnal dari masing-masing *link indeks*, maka fungsi *scrape_index* akan memanggil fungsi kedua yang akan melakukan proses *web scraping* akreditasi jurnal ilmiah. Hasil *web scraping* nantinya akan disimpan ke dalam *database* sesuai kondisi proses *web scraping* pada fungsi kedua.

Apabila proses *web scraping* akreditasi jurnal pada fungsi kedua terpenuhi, maka fungsi *scrape_index* akan menyimpan data akreditasi yang didapatkan. Sedangkan apabila proses *web scraping* akreditasi jurnal pada fungsi kedua tidak terpenuhi, maka fungsi *scrape_index* akan menyimpan data akreditasi jurnal sebagai “Not Found”, yang berarti jurnal tersebut belum memiliki akreditasi pada indeks SINTA.

Fungsi kedua yang dibuat pada *web scraper* akreditasi jurnal pada indeks SINTA ini bernama fungsi *scrape_journals_uin_walisongo*. Fungsi ini bertujuan untuk mencari setiap jurnal UIN Walisongo sesuai *link indeks* yang sedang diiterasi. Hal ini dilakukan karena terdapat kemungkinan bahwa beberapa jurnal ilmiah mempunyai nama yang sama satu sama lain. Oleh

karena itu perlu dilakukan pembatasan ekstraksi data sehingga jurnal ilmiah yang dicari benar-benar berasal termasuk Jurnal UIN Walisongo Semarang.

```
def scrape_journals_uin_walisongo(search_url, indeks_jurnal_id):
    response = requests.get(search_url)
    soup = BeautifulSoup(response.content, 'html.parser')

    journal_listings = soup.find_all('div', class_='list-item row mt-3')

    journals = []
    for listing in journal_listings:
        affiliation_element = listing.find('div', class_='affil-loc mt-2')
        if affiliation_element and "Walisongo" in affiliation_element.text:
            journal_name_element = listing.find('div', class_='affil-name mb-3')
            if journal_name_element:
                journal_name = journal_name_element.find('a').text.strip()

                accreditation_element = listing.find('span', class_='num-stat accredited')
                accreditation = None
                if accreditation_element:
                    accreditation = accreditation_element.text.strip().split()[0]

                journal_info = {
                    "accreditation": accreditation
                }
                journals.append(journal_info)
    return journals
```

Gambar 4.16 Script Pencarian Jurnal UIN Walisongo Semarang

Gambar 4.16 merupakan *script* fungsi pencarian jurnal UIN Walisongo Semarang pada situs indeks SINTA. Cara kerja fungsi ini adalah pertama, fungsi akan mengekstrak seluruh data jurnal yang ditampilkan pada hasil akses masing-masing *link indeks*. Hal ini dilakukan dengan menggunakan *method*

find_all pada tag elemen <div class = "list-item row mt-3". Setelah semua data jurnal ilmiah pada *link indeks* didapatkan, selanjutnya dilakukan sebuah penyaringan afiliasi jurnal ilmiah. Penyaringan tersebut dilakukan dengan menggunakan *method* get_text() dan *method* find pada *tag* elemen <div class = "affil-loc mt-2"> secara bersamaan pada semua data jurnal yang telah diekstrak sebelumnya, dimana *method* get_text() diharuskan mempunyai kata kunci "Walisongo" didalamnya. Penggunaan dua *method* tadi merupakan langkah kunci untuk menemukan jurnal ilmiah yang terafiliasi dengan UIN Walisongo Semarang.

Setelah jurnal ilmiah UIN Walisongo yang dicari teridentifikasi, selanjutnya dilakukan ekstraksi data akreditasi menggunakan *method* find pada *tag* elemen .

Menggunakan kombinasi fungsi scrape_index dan scrape_journals_uin_walisongo, data akreditasi dari setiap jurnal ilmiah UIN Walisongo Semarang pada situs indeks SINTA dapat diekstrak dan disimpan ke dalam *database*. Proses *scraping* berakhir ketika semua *link indeks* telah menjalani ekstraksi data. Setelah *script web scraper* akreditasi jurnal ilmiah berhasil dibuat,

script tersebut disimpan dengan nama *file* IndexScrapper.py.

Untuk melakukan proses *update* data artikel dan jurnal ilmiah secara otomatis oleh sistem monitoring ini, digunakanlah fitur *cron job*. Pada sistem monitoring ini, proses *update* data artikel dan jurnal ilmiah dijadwalkan untuk dilakukan setiap jamnya menggunakan fitur *cron job*. Sedangkan untuk proses *update* data akreditasi jurnal ilmiah dijadwalkan untuk dilakukan setiap 12 jam sekali. Penggunaan fitur *cron job* untuk otomatisasi pengecekan *update* data artikel dan jurnal ilmiah dapat dilihat pada Gambar 4.17.

Minute	Hour	Day	Month	Weekday	Command
0	*	*	*	*	/home/monito29/virtualenv/public_html/journalscraper/3.10/bin/python3 /home/monito29/public_html/JurnalScrapper.py
0	0	*	*	*	/home/monito29/virtualenv/public_html/journalscraper/3.10/bin/python3 /home/monito29/public_html/IndexScrapper.py

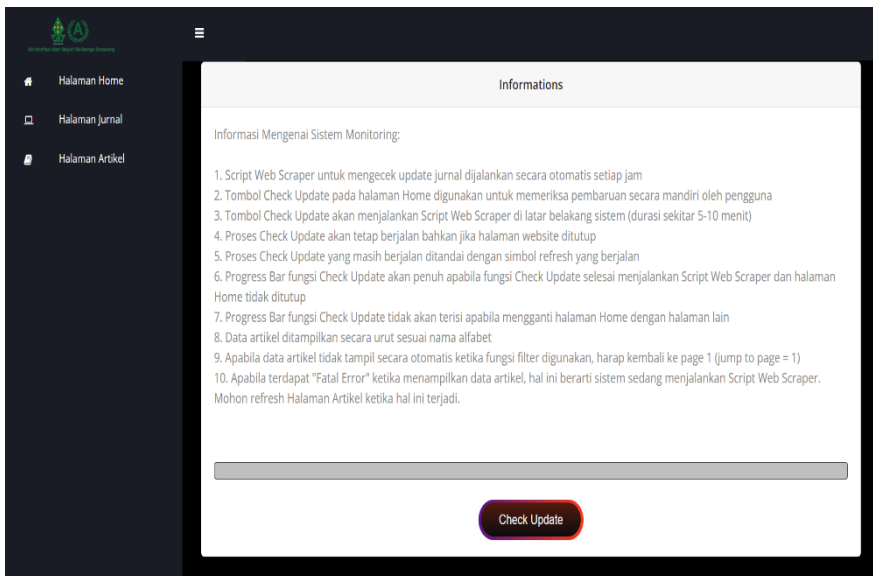
Gambar 4.17 Penggunaan Fitur Cron Job

Berdasarkan gambar diatas, fitur *cron job* bekerja dengan cara mengakses terlebih dahulu *virtual environtment* Python yang akan digunakan untuk menjalankan *script web scraper* yang telah dibuat. Setelah fitur *cron job* berhasil mengakses *virtual*

enviromtmet Python, fitur *cron job* kemudian akan mengakses direktori tempat *script* berada dan menjalankan *file* JurnalScrapper.py dan IndexScrapper.py.

4.3.2. Implementasi Website

Implementasi interface *website* pada sistem monitoring *update* artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang ini terdiri dari tiga halaman yaitu halaman Home, halaman Jurnal, dan halaman Artikel.



Gambar 4.18 Implementasi Halaman Home

Gambar 4.18 di atas merupakan halaman Home sistem. Pada bagian kiri halaman, terdapat daftar

sidebar yang dapat digunakan untuk mengakses halaman-halaman yang ada di sistem ini. Pada halaman Home ini, terdapat *information box* yang berisi informasi mengenai *web scraper*. Selain itu, di bawah *information box*, terdapat *progress bar* serta tombol *Check Update* yang memungkinkan *user* untuk melakukan cek *update* data artikel dan jurnal ilmiah secara pribadi. Ketika ditekan, tombol *Check Update* nantinya akan menjalankan *script web scraper JurnalScraper.py* di latar belakang sistem, tanpa menunggu cek *update* otomatis oleh fitur *cron job* yang berjalan setiap jamnya.

NO	NAMA JURNAL	INDEKS SINTA & JADWAL TERBIT	VOLUME TERBARU
1	Al-Ahkam	52 April And October	Vol 34, No 1 (2024): April Diterbitkan Pada: 2024-04-30
2	AL-ARBAH: Journal Of Islamic Finance And Banking	54 April And October	Vol 5, No 2 (2023) Diterbitkan Pada: 2023-11-01
3	Al-Hayat: Journal Of Biology And Applied Biology	53 May And November	Vol 7, No 1 (2024) Diterbitkan Pada: 2024-05-16
4	Al-Hilal: Journal Of Islamic Astronomy	Not Found April And October	Vol 5, No 2, 2023 Diterbitkan Pada: 2023-11-03

Gambar 4.19 Implementasi Halaman Jurnal

Gambar 4.19 di atas merupakan implementasi halaman Jurnal. Halaman Jurnal ini berfungsi untuk menampilkan data jurnal ilmiah yang ada di dalam *database*. Pada halaman ini, terdapat *Sorting Filter* dan *Journal Search* untuk memudahkan pencarian data jurnal ilmiah. *Sorting Filter* mempunyai empat pilihan pengurutan data jurnal ilmiah yang dapat dipilih. Empat macam pilihan tersebut adalah None, Nama, Indeks, dan Volume Terbaru.

Secara *default*, *Sorting Filter* ditetapkan sebagai pilihan None, yaitu data jurnal ilmiah akan ditampilkan menurut alfabet dari A-Z. Pilihan Nama mempunyai pengurutan yang sama dengan pilihan None, yaitu pengurutan sesuai alfabet dari A-Z. Untuk pilihan Indeks, *Sorting Filter* akan menampilkan data jurnal ilmiah dari akreditasi SINTA tertinggi hingga terendah (S1, S2, S3, S4, S5, hingga Not Found). Terakhir, pilihan Volume Terbaru akan menyebabkan data jurnal ilmiah ditampilkan berdasarkan urutan publikasi terbaru hingga terlama. *Journal Search* merupakan *filter* yang bekerja dengan cara mencocokkan kata kunci yang dimasukkan oleh *user* terhadap nama jurnal ilmiah di dalam *database*. Apabila kata kunci yang dimasukkan ditemukan di dalam *database*, maka jurnal ilmiah

dengan nama yang cocok dengan kata kunci akan ditampilkan.

Halaman Jurnal akan menampilkan maksimal 20 data jurnal ilmiah setiap *page*-nya. *User* dapat menavigasi antara *page* lain dari data jurnal ilmiah menggunakan fitur *pagination* di bagian bawah halaman Jurnal.

18	Journal Of Educational Chemistry (JEC)	S3 June And December	Vol 6, No 1 (2024) Diterbitkan Pada: 2024-06-08
19	Journal Of Integrated Elementary Education	S2 March And September	Vol 4, No 1 (2024): March 2024 Diterbitkan Pada: 2024-06-01
20	Journal Of Islamic Accounting And Finance Research	S2 April And October	Vol 5, No 2 (2023) Diterbitkan Pada: 2023-11-16

1

2

3

Next

Gambar 4.20 Fitur *Pagination* pada Halaman Jurnal

Gambar 4.20 diatas merupakan fitur *pagination* yang berada pada halaman Jurnal. Selanjutnya, Gambar 4.21 merupakan implementasi halaman Artikel. Halaman Artikel mempunyai cara menampilkan data seperti halaman Jurnal, yaitu menampilkan maksimal 20 data artikel ilmiah setiap *page*-nya. Selain itu, sama seperti halaman Jurnal, halaman Artikel juga mempunyai beberapa fungsi *filter* yang dapat digunakan untuk mempermudah penampilan dan pencarian data artikel ilmiah yang ada di dalam database.

Artikel Ilmiah			
NO	NAMA ARTIKEL	TANGGAL TERBIT	ASAL ARTIKEL
1	Urf Of Cyberspace: Solutions To The Problems Of Islamic Law In The Digital Age Penulis: Ita Musarrofa* - Universitas Islam Negeri Sunan Ampel Surabaya, Surabaya, Indonesia. Hoilur Rohman - Universitas Islam Negeri Sunan Ampel Surabaya, Surabaya, Indonesia.	2023-04-30	Al-Ahkam

Gambar 4.21 Implementasi Halaman Artikel

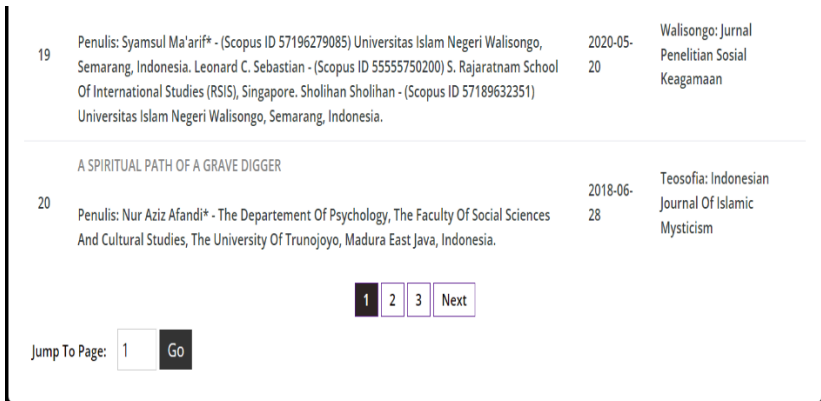
Halaman Artikel juga mempunyai empat *filter* yang dapat membantu pencarian artikel ilmiah yang diinginkan, yaitu *Year Filter*, *Sorting Filter*, *Journal Filter*, dan *Keyword Search*. *Year Filter* bekerja dengan cara menyaring data artikel ilmiah sesuai dengan *Start Year* dan *End Year* yang dipilih oleh *user*. Apabila *user* tidak memilih tahun tertentu, maka *Year Filter* tidak akan menyaring data artikel ilmiah. *Year Filter* memuat daftar tahun dimulai dari tahun terbit terlama artikel ilmiah hingga tahun terbit terbaru artikel ilmiah yang ada di *database*.

Sorting Filter bekerja hampir sama dengan yang ada pada halaman Jurnal. Perbedaan diantara keduanya adalah *Sorting Filter* pada halaman Artikel hanya mempunyai tiga macam pilihan, yaitu None, Nama, dan tanggal terbit.

Journal Filter bekerja dengan cara menyaring data artikel ilmiah sesuai dengan asal jurnal setiap artikel ilmiah. Jika *Journal Filter* digunakan, maka data artikel ilmiah yang ditampilkan hanya akan memuat artikel ilmiah dari jurnal yang dipilih. Terakhir, *Keyword Search* merupakan *filter* yang bekerja dengan cara mencocokkan kata kunci yang dimasukkan oleh *user* dengan nama dan penulis artikel ilmiah di *database*. Apabila kata kunci yang dimasukkan ditemukan di dalam *database*, maka artikel ilmiah dengan nama/penulis yang cocok dengan kata kunci akan ditampilkan. Apabila kata kunci yang dimasukkan oleh *user* tidak ditemukan di dalam *database*, maka tidak ada artikel ilmiah yang akan ditampilkan.

Data artikel ilmiah yang ditampilkan dipengaruhi oleh *filter* yang digunakan oleh *user*. Keempat *filter* di halaman Artikel dapat digunakan secara bersamaan untuk membantu pencarian artikel ilmiah. Pada akhir halaman Artikel, terdapat fungsi *pagination* dan *page*

jump untuk menavigasi data artikel ilmiah yang ditampilkan.



Gambar 4.22 Fitur Pagination dan Jump Page pada Halaman Artikel

Gambar 4.22 merupakan implementasi dari fitur *pagination* dan *page jump* di halaman Artikel. Fitur *page jump* akan berisi nilai bawaan berupa 1.

4.4. Integration and System Testing

Tahap ini dilakukan untuk menguji sistem yang telah dikembangkan. Pengujian pada tahap ini dibagi menjadi dua macam proses pengujian yaitu *functional test* dan *user acceptance test*. *Functional test* yang dilakukan memanfaatkan metode *black box testing* dengan menggunakan Teknik *Equivalence Partitioning*.

4.4.1. Functional Test

Metode *Black Box Testing* merupakan pengujian untuk menunjukkan kesalahan pada system aplikasi seperti kesalahan pada fungsi *system* aplikasi, serta

menu aplikasi yang hilang. Jadi *Black Box testing* merupakan metode uji fungsionalitas system aplikasi.

Dalam melakukan pengujian menggunakan masukan data acak dengan tujuan untuk mendapatkan hasil yang pasti. Dikatakan pasti artinya bila salah, maka di tolak oleh system informasi atau data input tersebut tidak dapat disimpan dalam *database*, sedangkan bila data input benar maka dapat diterima di *database* system informasi (Rahadi & Vikasari, 2020).

Pada penelitian ini teknik Pengujian yang dilakukan pada aplikasi Sistem Informasi perpustakaan adalah Teknik *Equivalence Partitioning*. Hasil pengujian yang dilakukan dapat dilihat pada Tabel 4.5.

Tabel 4.5 Hasil Pengujian Black Box

NO	Fungsi	Hasil yang Diharapkan	Hasil Pengujian	Kesimpulan
1.	Melakukan <i>request update</i> data artikel dan jurnal ilmiah	Fungsi <i>request update</i> dan <i>scraper</i> untuk mengekstrak data artikel dan jurnal	Fungsi <i>scraper</i> berjalan dengan lancar	Berhasil

NO	Fungsi	Hasil yang Diharapkan	Hasil Pengujian	Kesimpulan
		ilmiah berjalan dengan benar		
2.	Mencari jurnal ilmiah menurut kata kunci	Fungsi pencarian jurnal ilmiah berdasarkan kata kunci yang dimasukkan berjalan dengan benar	Fungsi pencarian dapat menyaring data jurnal ilmiah yang ditampilkan berdasarkan kata kunci	Berhasil
3.	Mengurutkan data jurnal ilmiah	Fungsi <i>filter sorting</i> data jurnal ilmiah berjalan dengan benar	Fungsi <i>filter sorting</i> dapat mengurutkan data jurnal ilmiah sesuai pilihan <i>user</i>	Berhasil
4.	Navigasi jurnal ilmiah	Fungsi <i>page</i> untuk menavigasi list data	Fungsi <i>page</i> dapat menavigasi data jurnal	Berhasil

NO	Fungsi	Hasil yang Diharapkan	Hasil Pengujian	Kesimpulan
		jurnal ilmiah berjalan dengan benar	ilmiah yang ditampilkan	
5.	Memfilter data artikel ilmiah menurut tahun	Fungsi <i>filter</i> data artikel ilmiah yang telah diekstrak berdasarkan tahun berjalan dengan benar	Fungsi <i>filter</i> dapat menampilkan data artikel ilmiah berdasarkan <i>start year</i> dan <i>end year</i>	Berhasil
6.	Memfilter data artikel ilmiah menurut asal jurnal	Fungsi <i>filter</i> data artikel ilmiah berdasarkan asal jurnal ilmiah yang sama berjalan dengan benar	Fungsi filter berdasarkan asal jurnal dapat menampilkan data artikel ilmiah berdasarkan jurnal yang dipilih <i>user</i>	Berhasil

NO	Fungsi	Hasil yang Diharapkan	Hasil Pengujian	Kesimpulan
7.	Mencari artikel ilmiah menurut kata kunci	Fungsi pencarian artikel ilmiah berdasarkan kata kunci yang dimasukkan berjalan dengan benar	Fungsi pencarian dapat menyaring data artikel ilmiah yang ditampilkan berdasarkan kata kunci	Berhasil
8.	Mengurutkan data artikel ilmiah	Fungsi <i>filter sorting</i> data artikel ilmiah berjalan dengan benar	Fungsi <i>filter sorting</i> dapat mengurutkan data artikel ilmiah sesuai pilihan <i>user</i>	Berhasil
9.	Navigasi artikel ilmiah	Fungsi <i>page</i> untuk menavigasi list data artikel ilmiah	Fungsi <i>page</i> dapat menavigasi data artikel ilmiah yang ditampilkan	Berhasil

NO	Fungsi	Hasil yang Diharapkan	Hasil Pengujian	Kesimpulan
		berjalan dengan benar		

Setelah didapatkan hasil uji sistem, langkah selanjutnya adalah menghitung persentase keberhasilan dari pengujian *black box* yang dilakukan. Perhitungan dilakukan menggunakan persamaan 3.1 dengan hasil sebagai berikut:

$$\text{Persentase Keberhasilan} = \frac{9}{9} \times 100\%$$

$$\text{Persentase Keberhasilan} = 100\%$$

Berdasarkan *fuctional test* menggunakan *black box* testing yang dilakukan, didapatkan bahwa sistem telah berjalan dengan baik karena memiliki persentase keberhasilan 100%.

4.4.2. User Acceptance Test

Pengujian *usability* dalam penelitian ini dilakukan dengan menggunakan *user acceptance test*. Responden pada pengujian ini berjumlah 40 mahasiswa UIN Walisongo Semarang dari beberapa jurusan yang berbeda.

Hasil kuesioner dan hasil pengujian yang dilakukan dapat dilihat pada Tabel 4.6 berikut:

Tabel 4.6 Hasil Pengujian User Acceptance Test

No	Pertanyaan	STS	TS	N	S	SS
1.	Apakah sistem ini mudah digunakan?	0	0	6	27	7
2.	Apakah fitur-fitur di dalam sistem ini merespon dengan cepat dan sesuai?	0	0	5	29	6
3.	Apakah data artikel dan jurnal ilmiah yang diekstrak sistem ini akurat?	0	0	2	34	4
4.	Apakah data artikel dan jurnal ilmiah ditampilkan dengan jelas dan mudah dipahami?	0	0	6	18	16
5.	Apakah sistem dapat beroperasi secara stabil dan tidak mengalami crash?	0	0	8	23	9

Berdasarkan hasil kuesioner diatas, dilakukan perhitungan skor total menggunakan persamaan 3.2 dengan hasil sebagai berikut:

$$Skor\ total = ((0 + 0 + 0 + 0 + 0) \times 1) + ((0 + 0 + 0 + 0 + 0) \times 2) + ((6 + 5 + 2 + 6 + 8) \times 3) + ((27 + 29 + 34 + 18 + 23) \times 4) + ((7 + 6 + 4 + 16 + 9) \times 5)$$

$$Skor\ total = 815$$

Setelah skor total didapatkan, maka dilakukan perhitungan rata-rata skor. Perhitungan rata-rata skor dilakukan menggunakan persamaan 3.3 dengan hasil di bawah ini:

$$\text{Rata - rata skor} = \frac{815}{5 \times 40}$$

$$\text{Rata - rata skor} = 4.075$$

Setelah rata-rata skor didapatkan, kini perlu dilakukan perhitungan persentase hasil kelayakan sistem monitoring. Perhitungan persentase hasil kelayakan dilakukan sesuai persamaan 3.4 dengan hasil sebagai berikut:

$$\text{Persentase hasil} = \frac{(4.075-1)}{(5-1)} \times 100\%$$

$$\text{Persentase hasil} = \frac{3.075}{4} \times 100\%$$

$$\text{Persentase hasil} = 76,9\%$$

Berdasarkan hasil pengujian *user acceptance test*, didapatkan hasil pengujian kelayakan sistem monitoring update artikel dan jurnal ilmiah secara otomatis di UIN Walisongo Semarang menggunakan teknik *web scraping* sebesar 76,9%. Berdasarkan kategori hasil kelayakan pada tabel 3.4, hasil tersebut menunjukkan bahwa sistem monitoring ini layak untuk diterapkan.

4.5. Operation and Maintenance

Tahap *operation and maintenance* merupakan tahap terakhir yang dilakukan pada penelitian ini. Tahap ini dilakukan dengan dengan cara memperbaiki *error* dan *crash* yang terjadi pada sistem monitoring *update* artikel dan jurnal ilmiah ini.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil pembahasan yang dilakukan dalam penelitian ini, dapat disimpulkan beberapa hal sebagai berikut:

1. Berdasarkan analisis dan perancangan sitem, sistem monitoring ini terdiri dari dua bagian utama yaitu *web scraper*, yang dibuat dengan bahasa pemrograman Python dan *library* Beautiful Soup 4, dan *website* sistem monitoring itu sendiri. *Web scraper* berfungsi sebagai pengekstrak data artikel dan jurnal ilmiah sedangkan *website* sistem monitoring berfungsi untuk menampilkan hasil ekstraksi yang dilakukan oleh *web scraper*. Selain itu, terdapat bagian pendukung yaitu otomatisasi proses *web scraping* menggunakan *cron job*.
2. Berdasarkan hasil pengujian *usability* menggunakan *user acceptance test*, didapatkan hasil kelayakan sistem monitoring sebesar 76,9%. Dengan kata lain, sistem monitoring *update* artikel dan jurnal ilmiah ini layak untuk diterapkan.

5.2. Saran

Berikut ini adalah beberapa saran yang dapat dipertimbangkan untuk penelitian mendatang:

1. Penelitian selanjutnya dapat mencoba menggunakan *library* maupun teknik *parser* yang berbeda selain *library* Beautiful Soup 4 dan metode HTML *Parser*. Library lain yang dapat digunakan untuk *web scraping* seperti Scrapy, Spider, Selenium, dan sebagainya. Sedangkan untuk metode parser selain HTML *Parser* dapat menggunakan LXML *Parser* ataupun JSON *Parser*.
2. Penelitian selanjutnya dapat mempertimbangkan untuk menambahkan fitur pengurutan data artikel selain sesuai alfabet.
3. Penelitian selanjutnya dapat mempertimbangkan untuk meningkatkan fitur *filter* data.

DAFTAR PUSTAKA

- Akh Sokhibi, M. A. A. (2020). Journal Of Industrial Engineering And Technology (Jointech) UNIVERSITAS MURIA KUDUS. *Jointech Umk*, 1(1).
- Arhandi, P. P., Mashudi, I. A., & Nugroho, F. A. (2021). Automated Website Monitoring System Using Web Scraping and Raspberry Pi. *Telematika*, 18(2).
<https://doi.org/10.31315/telematika.v18i2.5506>
- Baskara, R. (2022). *Implementasi Web Scraping Pada Media Sosial Instagram*.
- Brotowijoyo, M. D. (2002). Penulisan karangan ilmiah. (No Title).
- Cimperman, R. (2006). *Uat defined: A guide to practical user acceptance testing (digital short cut)*. Pearson Education.
- Daud, A., Aulia, A. F., Rimayanti, N., & Hardian, M. (2020). Pelatihan terstruktur: usaha peningkatan kompetensi guru dalam penulisan artikel jurnal ilmiah. *Unri Conference Series: Community Engagement*, 2.
<https://doi.org/10.31258/unricsce.2.90-97>
- Erfani, Moch. (2021). Upaya Meningkatkan Kompetensi Guru Untuk Menghadapi Kegiatan Belajar Dari Rumah Dengan Aplikasi Discord Pada SMP Negeri Satu Atap 2 Bungatan Tahun Pelajaran 2020/2021. *Jurnal IKA PGSD (Ikatan Alumni PGSD) UNARS*, 9(1).
<https://doi.org/10.36841/pgsdunars.v9i1.1020>
- Febriyanti, N. M. D., Sudana, A. A. K. O., & Piarsa, I. N. (2021). Implementasi Black Box Testing pada Sistem Informasi

- Manajemen Dosen. *Jurnal Ilmiah Teknologi Dan Komputer*, 2(3), 1–10.
- Fransisca Khurnia Erkhani. (2021). Sistem Informasi Pencatatan Hutang Dalam Pengadaan Obat Berbasis Multiuser (Studi Kasus Di RSUD Ambarawa). *Jurnal Ilmiah Sistem Informasi Dan Ilmu Komputer*, 1(3).
<https://doi.org/10.55606/juisik.v1i3.56>
- Gani, L., & Achmad, A. (2019). Website dan HTML. *Website Dan HTML*, 58.
- I Gusti Bagus Indra, P. (2023). *Analisis Data Produk Marketplace Menggunakan Web Scraping Menggunakan Metode Beautiful Soup Dan Lxml*. Universitas Mataram.
- Jantce TJ Sitinjak, D. D., Maman, ., & Suwita, J. (2020). Analisa Dan Perancangan Sistem Informasi Administrasi Kursus Bahasa Inggris Pada Intensive English Course Di Ciledug Tangerang. *Insan Pembangunan Sistem Informasi Dan Komputer (IPSIKOM)*, 8(1).
<https://doi.org/10.58217/ipsikom.v8i1.164>
- Jarmul, K., & Lawson, R. (2017). *Python web scraping*. Packt Publishing Ltd.
- Jaya, T. S. (2018). Pengujian Aplikasi dengan Metode Blackbox Testing Boundary Value Analysis (Studi Kasus: Kantor Digital Politeknik Negeri Lampung). *Jurnal Informatika: Jurnal Pengembangan IT*, 3(1).
<https://doi.org/10.30591/jpit.v3i1.647>
- Joko, M., & Haris, U. (2020). Scraping Web Marketplace Menggunakan Metode DOM Parsing Untuk Pengumpulan Data Produk. *Jurnal Ilmiah Informatika (JIF)*, 8(1).

- Kumala, A. E., Borman, R. I., & Prasetyawan, P. (2018). Sistem Informasi Monitoring Perkembangan Sapi Di Lokasi Uji Performance (Studi Kasus : Dinas Peternakan Dan Kesehatan Hewan Provinsi Lampung). *Jurnal Tekno Kompak*, 12(1). <https://doi.org/10.33365/jtk.v12i1.52>
- Kusmana, H. S. (2020). *Merancang karya tulis ilmiah*.
- Levi, M., Palit, H. N., Rostianingsih, S., & Siwalankerto, J. (2020a). Perbandingan Performa Tools Web Scraping pada Website dengan Data Statis dan Dinamis. *Jurnal Infra*, 8.
- Levi, M., Palit, H. N., Rostianingsih, S., & Siwalankerto, J. (2020b). Perbandingan Performa Tools Web Scraping pada Website dengan Data Statis dan Dinamis. *Jurnal Infra*, 8.
- Listiyan, E., & Subhiyakto, E. R. (2021). Rancang Bangun Sistem Inventory Gudang Menggunakan Metode Waterfall Studi Kasus Di Cv. Aqualux Duspha Abadi Kudus Jawa Tengah. *KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi*, 1(1). <https://doi.org/10.24002/konstelasi.v1i1.4272>
- Mahfudh, A. A., Rizki, F. A., & Alfaza, A. S. (2020). Tingkat Ketergantungan (Usability) E-learning di Fakultas Saitek UIN Walisongo Semarang. *Walisongo Journal of Information Technology*, 2(2). <https://doi.org/10.21580/wjit.2020.2.2.7172>
- MZ, M. M. K. (2016). Pengujian perangkat lunak metode black-box berbasis equivalence partitions pada aplikasi sistem informasi sekolah. *MIKROTIK: Jurnal Manajemen Informatika*, 6(1).
- Ningsih, W. (2016). Hakikat Belajar Menurut Perspektif Al-Qur'an. *Hakikat Belajarmenurut Perspektif Al-Qur'an*.

- Noviantoro, A., Silviana, A. B., Fitriani, R. R., & Permatasari, H. P. (2022). Rancangan Dan Implementasi Aplikasi Sewa Lapangan Badminton Wilayah Depok Berbasis Web. *Jurnal Teknik Dan Science*, 1(2), 88–103.
<https://doi.org/10.56127/jts.v1i2.108>
- Pratama, Y. A., & Junianto, E. (2016). Sistem Pakar Diagnosa Penyakit Ginjal Dan Saluran Kemih Dengan Metode Breadth First Search. *Jurnal Informatika*, 2(1).
<https://doi.org/10.31311/ji.v2i1.69>
- Prihatiningsih, T., Suhandini, Y., & Santikai, D. (2023). Penggunaan System Development Life Cycle (SDLC) Sebagai Sistem Informasi Penjualan Retail. *Energy - Jurnal Ilmiah Ilmu-Ilmu Teknik*, 13(1).
<https://doi.org/10.51747/energy.v13i1.1505>
- Purnomo, L. M., & Ayub, M. (2021). Analisis Data Hasil Web Scraping untuk Menentukan Kualitas Jurnal Ilmiah. *Jurnal Strategi*, 3(1).
- Putra, A. B., & Nita, S. (2019). Perancangan dan Pembangunan Sistem Informasi E-Learning Berbasis Web (Studi Kasus Pada Madrasah Aliyah Kare Madiun). *Seminar Nasional Teknologi Informasi Dan Komunikasi 2019*, 1(1), 81–85.
- Rahadi, N. W., & Vikasari, C. (2020). Pengujian Software Aplikasi Perawatan Barang Milik Negara Menggunakan Metode Black Box Testing Equivalence Partitions. *Infotekmesin*, 11(1). <https://doi.org/10.35970/infotekmesin.v11i1.124>
- Ramadhan, M. R. (2014). *Implementasi Teknik Web Scraping Untuk Pembuatan Web Service Harga Barang Situs E-Commerce Dengan Perangkat Bergerak Sebagai Klien*. Universitas Brawijaya.

- Roohullah Jan, S., Tauhid Ullah Shah, S., Ullah Johar, Z., Shah, Y., & Khan, F. (2016). An Innovative Approach to Investigate Various Software Testing Techniques and Strategies. *International Journal of Scientific Research in Science, Engineering and Technology*, 2(2).
- Saputra, G. (2016). Pengembangan Sistem Penanganan Keluhan Berbasis Web (Studi Kasus: Fakultas Teknik Universitas Muhammadiyah Yogyakarta). In *Ucv: Vol. 1* (Issue 02).
- Saurkar, A. V, & Gode, S. A. (2018). An Overview On Web Scraping Techniques And Tools. *International Journal on Future Revolution in Computer Science & Communication Engineering*, 4(4).
- Setiawansyah, S. (2020). Monitoring Aplikasi Menggunakan Dashboard Untuk Sistem Informasi Akuntansi Pembelian Dan Penjualan (STUDI KASUS : UD APUNG). *Jurnal Tekno Kompak*, 14(1). <https://doi.org/10.33365/jtk.v14i1.503>
- Solahudin, M. (2021). Rancang Bangun Sistem Informasi Akademik Sekolah (SIAS) Berbasis Website. *DoubleClick: Journal of Computer and Information Technology*, 4(2), 107. <https://doi.org/10.25273/doubleclick.v4i2.8315>
- Suaidah, S., & Sidni, I. (2018). Perancangan Monitoring Prestasi Akademik Dan Aktivitas Siswa Menggunakan Pendekatan Key Performance Indicator (Studi Kasus SMA N 1 Kalirejo). *Jurnal Tekno Kompak*, 12(2). <https://doi.org/10.33365/jtk.v12i2.154>
- Sugiarto, A., Susilawati, S., & Hidayat, R. (2021). Sistem Informasi Pendaftaran Pendampingan Legalisasi Produk UMKM Provinsi Banten Berbasis Web Menggunakan Metode SDLC Pada PKPH UNMA Banten. *International Journal Mathla'ul*

Anwar of Halal Issues, 1(2).

<https://doi.org/10.30653/ijma.202112.27>

Udil, P. A. (2021). Pelatihan Penulisan Artikel Ilmiah Penelitian Tindakan Kelas Untuk Publikasi Pada Jurnal ilmiah. *Jurnal Nasional Pengabdian Masyarakat*, 2(1).
<https://doi.org/10.47747/pengabdiankepadamasyarakat.v2i1.257>

Ulfah, A., & Najiah, I. (2023). Implementasi Web Scraping Pada Situs Jurnal Sinta Menggunakan Framework Selenium Webdriver Python. *JIKA (Jurnal Informatika)*, 7(1).
<https://doi.org/10.31000/jika.v7i1.7037>

Vargiu, E., & Urru, M. (2012). Exploiting web scraping in a collaborative filtering- based approach to web advertising. *Artificial Intelligence Research*, 2(1).
<https://doi.org/10.5430/air.v2n1p44>

Wahid, A. A. (2020). Analisis Metode Waterfall Untuk Pengembangan Sistem Informasi. *Jurnal Ilmu-Ilmu Informatika Dan Manajemen STMIK*, November.

Wahyudi, I., & Alameka, F. (2023). Analisis Blackbox Testing Dan User Acceptance Testing Terhadap Sistem Informasi Solusimedsosku. *Jurnal Teknosains Kodepena* /, 04(01).

Wahyudin, Y., & Rahayu, D. N. (2020). Analisis Metode Pengembangan Sistem Informasi Berbasis Website: A Literatur Review. *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 15(3).
<https://doi.org/10.35969/interkom.v15i3.74>

Wirawan, A. (2020). Sistem Scheduling Pelaporan Data Akademik di UIN Sunan Kalijaga ke Pangkalan Data Pendidikan Tinggi (PDDikti) dengan Menggunakan Fitur Cron Job di Linux.

JISKA (Jurnal Informatika Sunan Kalijaga), 5(3).

<https://doi.org/10.14421/jiska.2020.53-05>

Yuniarti, W. D. (2019). *Dasar-Dasar Pemrograman Dengan Python*. Deepublish.

Lampiran 1 Tabel Penyaringan Website

No	Nama Jurnal	Status Persyaratan
1	Psikohumaniora: Jurnal Penelitian Psikologi	Memenuhi
2	Al-Ahkam	Memenuhi
3	Walisongo: Jurnal Penelitian Sosial Keagamaan	Memenuhi
4	Jurnal Theologia	Memenuhi
5	Nadwa: Jurnal Pendidikan Islam	Memenuhi
6	Sawwa: Jurnal Studi Gender	Memenuhi
7	Economica: Jurnal Ekonomi Islam	Memenuhi
8	JSW (Jurnal Sosiologi Walisongo)	Memenuhi
9	Teosofia: Indonesian Journal of Islamic Mysticism	Memenuhi
10	Jurnal Ilmu Dakwah	Memenuhi
11	Journal of Islamic Accounting and Finance Research	Memenuhi
12	Vision: Journal for Language and Foreign Language Learning	Memenuhi
13	Journal of Educational Chemistry (JEC)	Memenuhi
14	At-Taqaddum	Memenuhi
15	Dimas: Jurnal Pemikiran Agama untuk Pemberdayaan	Memenuhi
16	Walisongo Journal of Chemistry	Memenuhi
17	Islamic Communication Journal	Memenuhi
18	International Journal Ihya' 'Ulum al-Din	Memenuhi
19	Nutri-Sains: Jurnal Gizi, Pangan dan Aplikasinya	Memenuhi
20	Physics Education Research Journal	Memenuhi
21	Phenomenon: Jurnal Pendidikan MIPA	Memenuhi
22	Journal of Natural Sciences and Mathematics Research	Memenuhi
23	Al-Hayat: Journal of Biology and Applied Biology	Memenuhi
24	Wahana Akademika: Jurnal Studi Islam dan Sosial	Memenuhi

No	Nama Jurnal	Status Persyaratan
25	Journal of Islamic Studies and Humanities	Memenuhi
26	Bioeduca: Journal of Biology Education	Memenuhi
27	Square: Journal of Mathematics and Mathematics Education	Memenuhi
28	Walisongo Law Review (Walrev)	Memenuhi
29	Journal of Islamic Economics Management and Business (JIEMB)	Memenuhi
30	Joint International Seminar	Memenuhi
31	AL-ARBAH: Journal of Islamic Finance and Banking	Memenuhi
32	Journal of Digital Marketing and Halal Industry	Memenuhi
33	Walisongo Journal of Information Technology	Memenuhi
34	Al-Hilal: Journal of Islamic Astronomy	Memenuhi
35	Journal of Integrated Elementary Education	Memenuhi
36	Journal of Advanced Guidance and Counseling	Memenuhi
37	Jawda: Journal of Islamic Education Management	Memenuhi
38	Journal of Early Childhood and Character Education	Memenuhi
39	Prosperity: Journal of Society and Empowerment	Memenuhi
40	Jurnal Theohumanitaria	Tidak Memenuhi (tidak mempunyai jurnal yang telah diterbitkan)
41	Alsina: Journal of Arabic Studies	Memenuhi
42	Daluang: Journal of Library and Information Science	Memenuhi
43	Journal of Islamic Art and Architecture (JIAA)	Memenuhi
44	JPW (Jurnal Politik Walisongo)	Memenuhi

No	Nama Jurnal	Status Persyaratan
45	Munazzama: Journal of Islamic Management and Pilgrimage	Memenuhi
46	Research Journal on Teacher Professional Development	Memenuhi

Lampiran 2 Tabel Archive Link

No	Nama Jurnal	Archive Link
1	Psikohumaniora: Jurnal Penelitian Psikologi	https://journal.walisongo.ac.id/index.php/002FPsikohumaniora/issue/archive
2	Al-Ahkam	https://journal.walisongo.ac.id/index.php/ahkam/issue/archive
3	Walisongo: Jurnal Penelitian Sosial Keagamaan	https://journal.walisongo.ac.id/index.php/walisongo/issue/archive
4	Jurnal Theologia	https://journal.walisongo.ac.id/index.php/teologia/issue/archive
5	Nadwa: Jurnal Pendidikan Islam	https://journal.walisongo.ac.id/index.php/Nadwa/issue/archive
6	Sawwa: Jurnal Studi Gender	https://journal.walisongo.ac.id/index.php/sawwa/issue/archive
7	Economica: Jurnal Ekonomi Islam	https://journal.walisongo.ac.id/index.php/economica/issue/archive
8	JSW (Jurnal Sosiologi Walisongo)	https://journal.walisongo.ac.id/index.php/JSW/issue/archive
9	Teosofia: Indonesian Journal of Islamic Mysticism	https://journal.walisongo.ac.id/index.php/teosofia/issue/archive
10	Jurnal Ilmu Dakwah	https://journal.walisongo.ac.id/index.php/dakwah/issue/archive
11	Journal of Islamic Accounting and Finance Research	https://journal.walisongo.ac.id/index.php/jiafr/issue/archive
12	Vision: Journal for Language and Foreign Language Learning	https://journal.walisongo.ac.id/index.php/vision/issue/archive

No	Nama Jurnal	Archive Link
13	Journal of Educational Chemistry (JEC)	https://journal.walisongo.ac.id/index.php/jec/issue/archive
14	At-Taqaddum	https://journal.walisongo.ac.id/index.php/attaqaddum/issue/archive
15	Dimas: Jurnal Pemikiran Agama untuk Pemberdayaan	https://journal.walisongo.ac.id/index.php/dimas/issue/archive
16	Walisongo Journal of Chemistry	https://journal.walisongo.ac.id/index.php/wjc/issue/archive
17	Islamic Communication Journal	https://journal.walisongo.ac.id/index.php/icj/issue/archive
18	International Journal Ihya' 'Ulum al-Din	https://journal.walisongo.ac.id/index.php/ihya/issue/archive
19	Nutri-Sains: Jurnal Gizi, Pangan dan Aplikasinya	https://journal.walisongo.ac.id/index.php/Nutri-Sains/issue/archive
20	Physics Education Research Journal	https://journal.walisongo.ac.id/index.php/perj/issue/archive
21	Phenomenon: Jurnal Pendidikan MIPA	https://journal.walisongo.ac.id/index.php/Phenomenon/issue/archive
22	Journal of Natural Sciences and Mathematics Research	https://journal.walisongo.ac.id/index.php/JNSMR/issue/archive
23	Al-Hayat: Journal of Biology and Applied Biology	https://journal.walisongo.ac.id/index.php/hayat/issue/archive
24	Wahana Akademika: Jurnal Studi Islam dan Sosial	https://journal.walisongo.ac.id/index.php/wahana/issue/archive

No	Nama Jurnal	Archive Link
25	Journal of Islamic Studies and Humanities	https://journal.walisongo.ac.id/index.php/JISH/issue/archive
26	Bioeduca: Journal of Biology Education	https://journal.walisongo.ac.id/index.php/BIOEDUCA/issue/archive
27	Square: Journal of Mathematics and Mathematics Education	https://journal.walisongo.ac.id/index.php/square/issue/archive
28	Walisongo Law Review (Walrev)	https://journal.walisongo.ac.id/index.php/walrev/issue/archive
29	Journal of Islamic Economics Management and Business (JIEMB)	https://journal.walisongo.ac.id/index.php/JIEMB/issue/archive
30	Joint International Seminar	https://journal.walisongo.ac.id/index.php/PJIS/issue/archive
31	AL-ARBAH: Journal of Islamic Finance and Banking	https://journal.walisongo.ac.id/index.php/arbah/issue/archive
32	Journal of Digital Marketing and Halal Industry	https://journal.walisongo.ac.id/index.php/JDMHI/issue/archive
33	Walisongo Journal of Information Technology	https://journal.walisongo.ac.id/index.php/jit/issue/archive
34	Al-Hilal: Journal of Islamic Astronomy	https://journal.walisongo.ac.id/index.php/al-hilal/issue/archive
35	Journal of Integrated Elementary Education	https://journal.walisongo.ac.id/index.php/jieed/issue/archive

No	Nama Jurnal	Archive Link
36	Journal of Advanced Guidance and Counseling	https://journal.walisongo.ac.id/index.php/agc/issue/archive
37	Jawda: Journal of Islamic Education Management	https://journal.walisongo.ac.id/index.php/jawda/issue/archive
38	Journal of Early Childhood and Character Education	https://journal.walisongo.ac.id/index.php/joece/issue/archive
39	Prosperity: Journal of Society and Empowerment	https://journal.walisongo.ac.id/index.php/prosperity/issue/archive
40	Alsina: Journal of Arabic Studies	https://journal.walisongo.ac.id/index.php/alsina/issue/archive
41	Daluang: Journal of Library and Information Science	https://journal.walisongo.ac.id/index.php/daluang/issue/archive
42	Journal of Islamic Art and Architecture (JIAA)	https://journal.walisongo.ac.id/index.php/JIAA/issue/archive
43	JPW (Jurnal Politik Walisongo)	https://journal.walisongo.ac.id/index.php/JPW/issue/archive
44	Munazzama: Journal of Islamic Management and Pilgrimage	https://journal.walisongo.ac.id/index.php/munazzama/issue/archive
45	Research Journal on Teacher Professional Development	https://journal.walisongo.ac.id/index.php/rjtpd/issue/archive

Lampiran 3 Tabel Struktur HTML Volume Terbaru

No	Nama Jurnal	Parent Element	Nama Volume	Link Volume
1	Psikohumaniora: Jurnal Penelitian Psikologi	<div, class = "issueDescription">	<a href>	<a href>
2	Al-Ahkam	<div, class = "issueDescription">	<a href>	<a href>
3	Walisongo: Jurnal Penelitian Sosial Keagamaan	<div, class = "issueDescription">	<a href>	<a href>
4	Jurnal Theologia	<div, class = "issueDescription">	<a href>	<a href>
5	Nadwa: Jurnal Pendidikan Islam	<div, class = "issueDescription">	<a href>	<a href>
6	Sawwa: Jurnal Studi Gender	<div, class = "issueDescription">	<a href>	<a href>
7	Economica: Jurnal Ekonomi Islam	<div, class = "issueDescription">	<a href>	<a href>
8	JSW (Jurnal Sosiologi Walisongo)	<div, class = "issueDescription">	<a href>	<a href>
9	Teosofia: Indonesian Journal of Islamic Mysticism	<div, class = "issueDescription">	<a href>	<a href>
10	Jurnal Ilmu Dakwah	<div, class = "issueDescription">	<a href>	<a href>
11	Journal of Islamic Accounting and Finance Research	<div, class = "issueDescription">	<a href>	<a href>

No	Nama Jurnal	Parent Element	Nama Volume	Link Volume
12	Vision: Journal for Language and Foreign Language Learning	<div, class = "issueDescripti on">	<a href>	<a href>
13	Journal of Educational Chemistry (JEC)	<div, class = "issueDescripti on">	<a href>	<a href>
14	At-Taqaddum	<div, class = "issueDescripti on">	<a href>	<a href>
15	Dimas: Jurnal Pemikiran Agama untuk Pemberdayaan	<div, class = "issueDescripti on">	<a href>	<a href>
16	Walisono Journal of Chemistry	<div, class = "issueDescripti on">	<a href>	<a href>
17	Islamic Communication Journal	<div, class = "issueDescripti on">	<a href>	<a href>
18	International Journal Ihya' 'Ulum al-Din	<div, class = "issueDescripti on">	<a href>	<a href>
19	Nutri-Sains: Jurnal Gizi, Pangan dan Aplikasinya	<div, class = "issueDescripti on">	<a href>	<a href>
20	Physics Education Research Journal	<div, class = "issueDescripti on">	<a href>	<a href>
21	Phenomenon: Jurnal Pendidikan MIPA	<div, class = "issueDescripti on">	<a href>	<a href>
22	Journal of Natural Sciences and Mathematics Research	<div, class = "issueDescripti on">	<a href>	<a href>

No	Nama Jurnal	Parent Element	Nama Volume	Link Volume
23	Al-Hayat: Journal of Biology and Applied Biology	<div, class = "issueDescription">	<a href>	<a href>
24	Wahana Akademika: Jurnal Studi Islam dan Sosial	<div, class = "issueDescription">	<a href>	<a href>
25	Journal of Islamic Studies and Humanities	<div, class = "issueDescription">	<a href>	<a href>
26	Bioeduca: Journal of Biology Education	<div, class = "issueDescription">	<a href>	<a href>
27	Square: Journal of Mathematics and Mathematics Education	<div, class = "issueDescription">	<a href>	<a href>
28	Walisongo Law Review (Walrev)	<div, class = "issueDescription">	<a href>	<a href>
29	Journal of Islamic Economics Management and Business (JIEMB)	<div, class = "issueDescriptionImage">	<a href>	<a href>
30	Joint International Seminar	<h4>	<a href>	<a href>
31	AL-ARBAH: Journal of Islamic Finance and Banking	<div, class = "issueDescription">	<a href>	<a href>
32	Journal of Digital Marketing and Halal Industry	<div, class = "issueDescription">	<a href>	<a href>
33	Walisongo Journal of	<div, class = "issueDescription">	<a href>	<a href>

No	Nama Jurnal	Parent Element	Nama Volume	Link Volume
	Information Technology			
34	Al-Hilal: Journal of Islamic Astronomy	<div, class = "issueDescription">	<a href>	<a href>
35	Journal of Integrated Elementary Education	<div, class = "issueDescriptionImage">	<a href>	<a href>
36	Journal of Advanced Guidance and Counseling	<div, class = "issueDescription">	<a href>	<a href>
37	Jawda: Journal of Islamic Education Management	<div, class = "issueDescription">	<a href>	<a href>
38	Journal of Early Childhood and Character Education	<div, class = "issueDescription">	<a href>	<a href>
39	Prosperity: Journal of Society and Empowerment	<div, class = "issueDescription">	<a href>	<a href>
40	Alsina: Journal of Arabic Studies	<div, class = "issueDescription">	<a href>	<a href>
41	Daluang: Journal of Library and Information Science	<div, class = "issueDescription">	<a href>	<a href>
42	Journal of Islamic Art and Architecture (JIAA)	<div, class = "issueDescriptionImage">	<a href>	<a href>

No	Nama Jurnal	Parent Element	Nama Volume	Link Volume
43	JPW (Jurnal Politik Walisongo)	<div, class = "issueDescription">	<a href>	<a href>
44	Munazzama: Journal of Islamic Management and Pilgrimage	<div, class = "issueDescription">	<a href>	<a href>
45	Research Journal on Teacher Professional Development	<div, style = "clear: left;">	<a href>	<a href>

Lampiran 4 Tabel Struktur HTML Artikel Ilmiah

No	Nama Jurnal	Tanggal Terbit	Parent Element	Nama Artikel	Link Artikel	Penulis Artikel
1	Psikohumaniora : Jurnal Penelitian Psikologi	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
2	Al-Ahkam	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
3	Walisongo: Jurnal Penelitian Sosial Keagamaan	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
4	Jurnal Theologia	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
5	Nadwa: Jurnal Pendidikan Islam	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
6	Sawwa: Jurnal Studi Gender	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
7	Economica: Jurnal Ekonomi Islam	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"

No	Nama Jurnal	Tanggal Terbit	Parent Element	Nama Artikel	Link Artikel	Penulis Artikel
8	JSW (Jurnal Sosiologi Walison go)	<div, class = "infoissu e">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
9	Teosofia : Indonesian Journal of Islamic Mysticism	<div, class = "infoissu e">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
10	Jurnal Ilmu Dakwah	<div, class = "infoissu e">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
11	Journal of Islamic Accounting and Finance Research	<div, class = "infoissu e">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
12	Vision: Journal for Language and Foreign Language Learning	<div, class = "infoissu e">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"

No	Nama Jurnal	Tanggal Terbit	Parent Element	Nama Artikel	Link Artikel	Penulis Artikel
13	Journal of Educational Chemistry (JEC)	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
14	At-Taqaddum	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
15	Dimas: Jurnal Pemikiran Agama untuk Pemberdayaan	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
16	Walisongo Journal of Chemistry	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
17	Islamic Communication Journal	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
18	International Journal Ihya' 'Ulum al-Din	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
19	Nutri-Sains: Jurnal Gizi,	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"

No	Nama Jurnal	Tanggal Terbit	Parent Element	Nama Artikel	Link Artikel	Penulis Artikel
	Pangan dan Aplikasinya					
20	Physics Education Research Journal	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
21	Phenomenon: Jurnal Pendidikan MIPA	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
22	Journal of Natural Sciences and Mathematics Research	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
23	Al-Hayat: Journal of Biology and Applied Biology	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
24	Wahana Akademika: Jurnal Studi	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"

No	Nama Jurnal	Tanggal Terbit	Parent Element	Nama Artikel	Link Artikel	Penulis Artikel
	Islam dan Sosial					
25	Journal of Islamic Studies and Humanities	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
26	Bioeduca: Journal of Biology Education	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
27	Square: Journal of Mathematics and Mathematics Education	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
28	Walisongo Law Review (Walrev)	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
29	Journal of Islamic Economics Manage	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"

No	Nama Jurnal	Tanggal Terbit	Parent Element	Nama Artikel	Link Artikel	Penulis Artikel
	ment and Business (JIEMB)					
30	Joint International Seminar	Tidak ada	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, class="tocAuthors"
31	AL-ARBAH: Journal of Islamic Finance and Banking	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
32	Journal of Digital Marketing and Halal Industry	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
33	Walison go Journal of Information Technology	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
34	Al-Hilal: Journal of Islamic Astronomy	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"

No	Nama Jurnal	Tanggal Terbit	Parent Element	Nama Artikel	Link Artikel	Penulis Artikel
35	Journal of Integrated Elementary Education	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
36	Journal of Advanced Guidance and Counseling	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
37	Jawda: Journal of Islamic Education Management	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
38	Journal of Early Childhood and Character Education	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
39	Prosperity: Journal of Society and	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"

No	Nama Jurnal	Tanggal Terbit	Parent Element	Nama Artikel	Link Artikel	Penulis Artikel
	Empowerment					
40	Alsina: Journal of Arabic Studies	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
41	Daluang : Journal of Library and Information Science	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
42	Journal of Islamic Art and Architecture (JIAA)	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
43	JPW (Jurnal Politik Walison go)	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"
44	Munazzama: Journal of Islamic Management and Pilgrimage	<div, class = "infoissue">	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, id="authorString"

No	Nama Jurnal	Tanggal Terbit	Parent Element	Nama Artikel	Link Artikel	Penulis Artikel
45	Research Journal on Teacher Professional Development	Tidak Ada	td, class="tocArticleTitleAuthors"	<a href>	<a href>	div, class="tocAuthors"

Lampiran 5 Tabel Struktur HTML Akreditasi Jurnal Ilmiah

No	Nama Jurnal	Parent Element	Akreditasi	Link SINTA
1	Psikohumaniora: Jurnal Penelitian Psikologi	'div', class_='aff il-name mb-3'	span, class=" num-stat accredited"	<a href="https://sinta.ke
mdikbud.go.id/
journals/index
/?q=Psikohuma
niora">https://sinta.ke mdikbud.go.id/ journals/index /?q=Psikohuma niora
2	Al-Ahkam	'div', class_='aff il-name mb-3'	span, class=" num-stat accredited"	<a href="https://sinta.ke
mdikbud.go.id/
journals/index
/?q=Al-Ahkam">https://sinta.ke mdikbud.go.id/ journals/index /?q=Al-Ahkam
3	Walisongo: Jurnal Penelitian Sosial Keagamaan	'div', class_='aff il-name mb-3'	span, class=" num-stat accredited"	<a href="https://sinta.ke
mdikbud.go.id/
journals/index
/?q=Jurnal+Pen
elitian+Sosial+
Keagamaan">https://sinta.ke mdikbud.go.id/ journals/index /?q=Jurnal+Pen elitian+Sosial+ Keagamaan
4	Jurnal Theologia	div, class_="ca rd bg- card- orange"	div, class_="stat -num"	<a href="https://sinta.ke
mdikbud.go.id/
journals/profil
e/818">https://sinta.ke mdikbud.go.id/ journals/profil e/818
5	Nadwa: Jurnal Pendidikan Islam	'div', class_='aff il-name mb-3'	span, class=" num-stat accredited"	<a href="https://sinta.ke
mdikbud.go.id/
journals/index
/?q=nadwa">https://sinta.ke mdikbud.go.id/ journals/index /?q=nadwa
6	Sawwa: Jurnal Studi Gender	'div', class_='aff il-name mb-3'	span, class=" num-stat accredited"	<a href="https://sinta.ke
mdikbud.go.id/
journals/index
/?q=sawwa">https://sinta.ke mdikbud.go.id/ journals/index /?q=sawwa
7	Economica: Jurnal Ekonomi Islam	'div', class_='aff il-name mb-3'	span, class=" num-stat accredited"	<a href="https://sinta.ke
mdikbud.go.id/
journals/index
/?q=economica">https://sinta.ke mdikbud.go.id/ journals/index /?q=economica
8	JSW (Jurnal Sosiologi Walisongo)	'div', class_='aff il-name mb-3'	span, class=" num-stat accredited"	<a href="https://sinta.ke
mdikbud.go.id/
journals/index
/?q=JSW">https://sinta.ke mdikbud.go.id/ journals/index /?q=JSW

No	Nama Jurnal	Parent Element	Akreditasi	Link SINTA
9	Teosofia: Indonesian Journal of Islamic Mysticism	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=teosofia
10	Jurnal Ilmu Dakwah	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=Jurnal+Ilmu+Dakwah
11	Journal of Islamic Accounting and Finance Research	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=Journal+of+Islamic+Accounting+and+Finance+Research
12	Vision: Journal for Language and Foreign Language Learning	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=vision
13	Journal of Educational Chemistry (JEC)	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=Journal+of+Educational+Chemistry
14	At-Taqqaddum	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=At-Taqqaddum
15	Dimas: Jurnal Pemikiran Agama untuk Pemberdayaan	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=Dimas

No	Nama Jurnal	Parent Element	Akreditasi	Link SINTA
16	Walisongo Journal of Chemistry	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=Walisongo+Journal+of+Chemistry
17	Islamic Communication Journal	div, class_="card bg-card-orange"	div, class_="stat-num"	https://sinta.ke mdikbud.go.id/journals/profile/382
18	International Journal Ihya' 'Ulum al-Din	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=International+Journal+Ihya'+Ulum+al-Din
19	Nutri-Sains: Jurnal Gizi, Pangan dan Aplikasinya	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=nutri-sains
20	Physics Education Research Journal	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=Physics+Education+Research+Journal
21	Phenomenon: Jurnal Pendidikan MIPA	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=Phenomenon
22	Journal of Natural Sciences and Mathematics Research	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.ke mdikbud.go.id/journals/index/?q=Journal+of+Natural+Scien

No	Nama Jurnal	Parent Element	Akreditasi	Link SINTA
				ces+and+Mathematics+Research
23	Al-Hayat: Journal of Biology and Applied Biology	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=al-hayat
24	Wahana Akademika: Jurnal Studi Islam dan Sosial	div, class_"card bg-card-orange"	div, class_"stat-num"	https://sinta.kemdikbud.go.id/journals/profile/1866
25	Journal of Islamic Studies and Humanities	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Journal+of+Islamic+Studies+and+Humanities
26	Bioeduka: Journal of Biology Education	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=bioeduka
27	Square: Journal of Mathematics and Mathematics Education	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=square
28	Walisongo Law Review (Walrev)	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=walrev
29	Journal of Islamic Economics Management and Business (JIEMB)	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Journal+of+Islamic+Economics+Management+and+Business

No	Nama Jurnal	Parent Element	Akreditasi	Link SINTA
				ent+and+Business
30	Joint International Seminar	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Joint+International+Semin ar
31	AL-ARBAH: Journal of Islamic Finance and Banking	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=al-arbah
32	Journal of Digital Marketing and Halal Industry	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Journal+of +Digital+Marke ting+and+Halal +Industry
33	Walisongo Journal of Information Technology	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Walisongo +Journal+of+Inf ormation+Tech nology
34	Al-Hilal: Journal of Islamic Astronomy	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Al-Hilal
35	Journal of Integrated Elementary Education	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Journal+of +Integrated+El ementary+Educ ation

No	Nama Jurnal	Parent Element	Akreditasi	Link SINTA
36	Journal of Advanced Guidance and Counseling	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Journal+of+Advanced+Guidance+and+Counseling
37	Jawda: Journal of Islamic Education Management	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=jawda
38	Journal of Early Childhood and Character Education	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Journal+of+Early+Childhood+and+Character+Education
39	Prosperity: Journal of Society and Empowerment	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Prosperity
40	Alsina: Journal of Arabic Studies	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Alsina
41	Daluang: Journal of Library and Information Science	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=daluang
42	Journal of Islamic Art and Architecture (JIAA)	'div', class_='aff il-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Journal+of+Islamic+Art+and+Architecture

No	Nama Jurnal	Parent Element	Akreditasi	Link SINTA
43	JPW (Jurnal Politik Walisongo)	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Jurnal+Politik+Walisongo
44	Munazzama: Journal of Islamic Management and Pilgrimage	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Munazzama
45	Research Journal on Teacher Professional Development	'div', class_='affil-name mb-3'	span, class="num-stat accredited"	https://sinta.kemdikbud.go.id/journals/index/?q=Research+Journal+on+Teacher+Professional+Development

Lampiran 6 Script Web Scraper Artikel dan Jurnal Ilmiah

```
import requests
from bs4 import BeautifulSoup
import mysql.connector
import datetime
from threading import Thread # for multithreading
import time # for rate limiting (optional)
import re

start_time = time.time()

def joint_seminar():
    base_url =
"https://journal.walisongo.ac.id/index.php/PJIS/issue/archive"
    cursor.execute("INSERT INTO `jurnal` (`id_jurnal`, `nama_jurnal`,
`link_jurnal`, `jadwal_terbit`, `archive_jurnal`, `volume_terbaru`,
`link_indeks`, `indeks_jurnal`) VALUES ('30', 'Joint International
Seminar', 'https://journal.walisongo.ac.id/index.php/PJIS',
'Unknown',
'https://journal.walisongo.ac.id/index.php/PJIS/issue/archive',
NULL,
'https://sinta.kemdikbud.go.id/journals/index/?q=Joint+Internatio
nal+Seminar', 'Not Found');")
    db.commit()
    delete_query = """
DELETE a
FROM artikel AS a
INNER JOIN (
SELECT link_artikel, MIN(id_artikel) AS min_id
FROM artikel
GROUP BY link_artikel
) AS b ON a.link_artikel = b.link_artikel AND a.id_artikel <>
b.min_id
"""

    cursor.execute(delete_query)
    db.commit()

page = requests.get(base_url)
```

```

soup = BeautifulSoup(page.content, "html.parser")

# Find all issue links (assuming they are within anchor tags)
issue_element = soup.find("div", class_=("issueDescription",
"issueDescriptionImage", "issueCoverImage"))
issue_link = issue_element.find("a")["href"]
latest_volume = "2015"

# Update volume_terbaru in jurnal table for the current
archive_jurnal
update_query = """
    UPDATE jurnal
    SET volume_terbaru = %s
    WHERE archive_jurnal = %s
    """
cursor.execute(update_query, (latest_volume, base_url))
db.commit()

print(f'Latest volume for {base_url}: {latest_volume}') #
Confirmation message

archive_jurnal_id = int(30)

try:
    page_article = requests.get(issue_link)
    soup2 = BeautifulSoup(page_article.content, "html.parser")

    # Try scraping from TOC image link first
    toc_image_element = soup2.find("div", id="issueCoverImage")
    if toc_image_element and toc_image_element.find("a",
href=True):
        toc_image_link = toc_image_element.find("a")["href"]
        toc_page = requests.get(toc_image_link)
        soup3 = BeautifulSoup(toc_page.content, "html.parser")

        # Get current date
        current_date = datetime.date.today()

        # Format the date as YYYY-MM-DD
        formatted_date = current_date.strftime('%Y-%m-%d')

```

```

penulis_artikel = "Unknown"
artikel_parent = soup3.find_all("div", class_="tocTitle")

for detail in artikel_parent:
    nama_artikel = (
        detail.find("a", href=True).text
        if detail.find("a", href=True)
        else "Front Matter"
    )
    link_artikel = (
        detail.find("a", href=True)["href"]
        if detail.find("a", href=True)
        else "#"
    )

    insert_query = """
        INSERT INTO artikel (nama_artikel, asal_artikel, link_artikel,
tanggal_terbit, penulis_artikel)
        VALUES (%s, %s, %s, %s, %s)
    """

    cursor.execute(insert_query, (nama_artikel, archive_jurnal_id,
link_artikel, formatted_date, penulis_artikel))
    db.commit()

    print(f"{nama_artikel} {archive_jurnal_id} {link_artikel}
{formatted_date} {penulis_artikel}")

num_articles = len(artikel_parent)

print("")
print(f"Number of articles in this issue: {num_articles}")
print("-" * 50) # Separator between issues

except requests.exceptions.RequestException as e:
    print(f"Error fetching issue link: {e}")
    pass

def RJPTD():

```

```

base_url =
"https://journal.walisongo.ac.id/index.php/rjtpd/issue/archive"
cursor.execute("INSERT INTO `jurnal` (`id_jurnal`, `nama_jurnal`,
`link_jurnal`, `jadwal_terbit`, `archive_jurnal`, `volume_terbaru`,
`link_indeks`, `indeks_jurnal`) VALUES ('45', 'Research Journal on
Teacher Professional Development',
'https://journal.walisongo.ac.id/index.php/rjtpd', 'Twice a year',
'https://journal.walisongo.ac.id/index.php/rjtpd/issue/archive',
NULL,
'https://sinta.kemdikbud.go.id/journals/index/?q=Research+Journ
al+on+Teacher+Professional+Development', 'Not Found');")
db.commit()

```

```

delete_query = """
DELETE a
FROM artikel AS a
INNER JOIN (
    SELECT link_artikel, MIN(id_artikel) AS min_id
    FROM artikel
    GROUP BY link_artikel
) AS b ON a.link_artikel = b.link_artikel AND a.id_artikel <>
b.min_id
"""

```

```

cursor.execute(delete_query)
db.commit()

```

```

# Fetch the current page
page = requests.get(base_url)
soup = BeautifulSoup(page.content, "html.parser")

```

```

# Find all issue links (assuming they are within anchor tags)
issue_element = soup.find("div", style="clear:left;")
issue_link = issue_element.find("a")["href"]
latest_volume = issue_element.find("a").text

```

```

# Update volume_terbaru in jurnal table for the current
archive_jurnal
update_query = """
UPDATE jurnal

```

```

        SET volume_terbaru = %s
        WHERE archive_jurnal = %s
    """
    cursor.execute(update_query, (latest_volume, base_url))
    db.commit()

    print(f"Latest volume for {base_url}: {latest_volume}") #
Confirmation message

    archive_jurnal_id = int(45)

    try:
        page_article = requests.get(issue_link)
        soup2 = BeautifulSoup(page_article.content, "html.parser")
        # Fallback: scrape from current issue link

        # Get current date
        current_date = datetime.date.today()

        # Format the date as YYYY-MM-DD
        formatted_date = current_date.strftime('%Y-%m-%d')

        artikel_parent = soup2.find_all("td",
class_="tocArticleTitleAuthors")

        for detail in artikel_parent:
            # Extract article title
            title_element = detail.find("div", class_="tocTitle")
            title_link = title_element.find("a") if title_element else None
            nama_artikel = title_link.text.strip() if title_link else "Front
Matter"
            link_artikel = title_link["href"] if title_link else "#"

            authors = []
            author_elements = detail.find_all("div", class_="tocAuthors")
            for author_element in author_elements:
                author_text = author_element.text.strip()
                author_name = re.sub(r"\s+", " ", author_text) # Replace
multiple whitespaces with single space
                formatted_author = f"{author_name}."

```

```

        authors.append(formatted_author)

# Combine and format author names
penulis_artikel = "\n".join(authors)

insert_query = """
    INSERT INTO artikel (nama_artikel, asal_artikel, link_artikel,
tanggal_terbit, penulis_artikel)
    VALUES (%s, %s, %s, %s, %s)
    """
    cursor.execute(insert_query, (nama_artikel, archive_jurnal_id,
link_artikel, formatted_date, penulis_artikel))
    db.commit()

    print(f"{nama_artikel} {archive_jurnal_id} {link_artikel}
{formatted_date} {penulis_artikel}")

num_articles = len(artikel_parent)

print("")
print(f"Number of articles in this issue: {num_articles}")
print("-" * 50) # Separator between issues

except requests.exceptions.RequestException as e:
    print(f"Error fetching issue link: {e}")
    pass

def scrape_issue(issue_link, archive_jurnal_id):
    try:
        page_article = requests.get(issue_link)
        soup2 = BeautifulSoup(page_article.content, "html.parser")

        # Try scraping from TOC image link first
        toc_image_element = soup2.find("div", id="issueCoverImage",
class_="col-sm-3")
        if toc_image_element and toc_image_element.find("a",
href=True):
            toc_image_link = toc_image_element.find("a")["href"]

        # Combine requests for fetching articles and TOC details

```

```

toc_page_response = requests.get(toc_image_link)
soup3 = BeautifulSoup(toc_page_response.content,
"html.parser")

tanggal_terbit = soup3.find("div",
class_="infoissue").text.replace(
    " Published: ", ""
)

date = datetime.datetime.strptime(tanggal_terbit, '%Y-%m-
%d').date() # Format date

artikel_parent = soup3.find_all("td",
class_="tocArticleTitleAuthors")
num_articles = len(artikel_parent)

for detail in artikel_parent:
    # Extract article title
    title_element = detail.find("div", class_="tocTitle")
    title_link = title_element.find("a") if title_element else None
    nama_artikel = title_link.text.strip() if title_link else "Front
Matter"
    link_artikel = title_link["href"] if title_link else "#"

    # Extract and format author names (logic remains the
same)
    authors = []
    author_elements = detail.find_all("div", id="authorString")
    for author_element in author_elements:
        author_text = author_element.text.strip()
        author_name = re.sub(r"\s+", " ", author_text) # Replace
multiple whitespaces with single space
        formatted_author = f"{author_name}."
        authors.append(formatted_author)

    # Combine and format author names
    penulis_artikel = "\n".join(authors)

    # Insert data into artikel table
    insert_query = ""

```

```

        INSERT INTO artikel (nama_artikel, asal_artikel,
link_artikel, tanggal_terbit, penulis_artikel)
        VALUES (%s, %s, %s, %s, %s)
        """

        cursor.execute(insert_query, (nama_artikel,
archive_jurnal_id, link_artikel, date, penulis_artikel))
        db.commit()
        print(f"{nama_artikel} {archive_jurnal_id} {link_artikel}
{date} {penulis_artikel}")

    else:
        # Fallback: scrape from current issue link
        tanggal_terbit = soup2.find("div",
class_="infoissue").text.replace(
            " Published: ", ""
        )
        date = datetime.datetime.strptime(tanggal_terbit, '%Y-%m-
%d').date() # Format date

        artikel_parent = soup2.find_all("td",
class_="tocArticleTitleAuthors")
        num_articles = len(artikel_parent)

        for detail in artikel_parent:
            # Extract article title
            title_element = detail.find("div", class_="tocTitle")
            title_link = title_element.find("a") if title_element else None
            nama_artikel = title_link.text.strip() if title_link else "Front
Matter"
            link_artikel = title_link["href"] if title_link else "#"

            authors = []
            author_elements = detail.find_all("div", id="authorString")
            for author_element in author_elements:
                author_text = author_element.text.strip()
                author_name = re.sub(r"\s+", " ", author_text) # Replace
multiple whitespaces with single space
                formatted_author = f"{author_name}."
                authors.append(formatted_author)

```

```

# Combine and format author names
penulis_artikel = "\n".join(authors)

# Insert data into artikel table
insert_query = """
    INSERT INTO artikel (nama_artikel, asal_artikel,
link_artikel, tanggal_terbit, penulis_artikel)
        VALUES (%s, %s, %s, %s, %s)
    """
    cursor.execute(insert_query, (nama_artikel,
archive_jurnal_id, link_artikel, date, penulis_artikel))
    db.commit()
    print(f"{nama_artikel} {archive_jurnal_id} {link_artikel}
{date} {penulis_artikel}")

print("")
print(f"Number of articles in this issue: {len(artikel_parent)}")
print("-" * 50) # Separator between issues

except requests.exceptions.RequestException as e:
    print(f"Error fetching issue link: {e}")
    pass

def Jurnal(archive_links):
    link_number = 1
    for i, base_url in enumerate(archive_links):
        page = requests.get(base_url)
        soup = BeautifulSoup(page.content, "html.parser")

        # Find all issue links (assuming they are within anchor tags)
        issue_element = soup.find("div", class_=("issueDescription",
"issueDescriptionImage"))
        if not issue_element:
            print(f"Issue element not found for {base_url}, skipping.")
            continue

        issue_link = issue_element.find("a")["href"]

```

```

    # Extract latest volume (assuming it's in the text of the anchor
tag)
    latest_volume = issue_element.find("a").text.replace("book ", "")

    # Update volume_terbaru in jurnal table for the current
archive_jurnal
    update_query = """
        UPDATE jurnal
        SET volume_terbaru = %s
        WHERE archive_jurnal = %s
    """

    cursor.execute(update_query, (latest_volume, base_url))
    db.commit()

    print(f'Latest volume for {base_url}: {latest_volume}') #
Confirmation message

    # Fetch archive_jurnal_id using afetchone()
    sql_get_id = "SELECT id_jurnal FROM jurnal WHERE
archive_jurnal = (%s)"
    data = (base_url,) # Wrap base_url in a tuple for the query
parameter
    cursor.execute(sql_get_id, data)
    archive_jurnal_id = cursor.fetchone()[0] # Extract the first
column (id_jurnal) from the result

    scrape_issue(issue_link, archive_jurnal_id) # Pass
archive_jurnal_id for proper linking

    link_number += 1

# Connect to database
db = mysql.connector.connect(
    host="localhost",
    username="monito29_admin",
    password="monito29_admin",
    database="monito29_db"
),
    charset= "utf8"
)

```

```

cursor = db.cursor()
cursor.execute("DELETE FROM jurnal WHERE archive_jurnal =
'https://journal.walisongo.ac.id/index.php/PJIS/issue/archive'")
cursor.execute("DELETE FROM jurnal WHERE archive_jurnal =
'https://journal.walisongo.ac.id/index.php/rjtpd/issue/archive'")
db.commit()

```

```

delete_query = """
DELETE a
FROM artikel AS a
INNER JOIN (
    SELECT link_artikel, MIN(id_artikel) AS min_id
    FROM artikel
    GROUP BY link_artikel
) AS b ON a.link_artikel = b.link_artikel AND a.id_artikel <>
b.min_id
"""

```

```

cursor.execute(delete_query)
db.commit()

```

```

cursor.execute("SELECT archive_jurnal FROM jurnal")
archive_links = [x[0] for x in cursor.fetchall()]

```

```

Jurnal(archive_links)
joint_seminar()
RJPTD()

```

```

delete_query = """
DELETE a
FROM artikel AS a
INNER JOIN (
    SELECT link_artikel, MIN(id_artikel) AS min_id
    FROM artikel
    GROUP BY link_artikel
) AS b ON a.link_artikel = b.link_artikel AND a.id_artikel <>
b.min_id
"""

```

```
cursor.execute(delete_query)
db.commit()

print("Duplicate rows deleted from artikel table.")

# Close database connection
db.close()

end_time = time.time()
execution_time = end_time - start_time
print("Execution time:", execution_time)
```

Lampiran 7 Script Web Scraper Akreditasi SINTA

```
import requests
from bs4 import BeautifulSoup
import mysql.connector
import datetime
import time # for rate limiting (optional)

def scrape_journals_uin_walisongo(search_url, indeks_jurnal_id):
    """
    Scrapes the accreditation of journals listed on the SINTA search
    page
    that are affiliated with UIN Walisongo Semarang.

    Args:
        search_url (str): The URL of the search page on SINTA website.
        indeks_jurnal_id (int): The ID of the journal in the database.

    Returns:
        str: The extracted accreditation (or None if not found).
    """

    response = requests.get(search_url)
    soup = BeautifulSoup(response.content, 'html.parser')

    # Find all journal listings
    journal_listings = soup.find_all('div', class_='list-item row mt-3')

    journals = []
    for listing in journal_listings:
        # Check affiliation location (more generic selector)
        affiliation_element = listing.find('div', class_=lambda
class_name: class_name and 'affil-loc' in class_name and 'mt-2' in
class_name)
        if affiliation_element and "Walisongo" in affiliation_element.text:
            # Extract journal name
            journal_name_element = listing.find('div', class_='affil-name
mb-3')
            if journal_name_element:
```

```

        journal_name = journal_name_element.find('a').text.strip()

        # Find accreditation element (enhanced selector)
        accreditation_element = listing.find('span', class_=lambda
class_name: class_name and 'num-stat' in class_name and
'accredited' in class_name)
        accreditation = None
        if accreditation_element:
            accreditation = accreditation_element.text.strip().split()[0]

        # Add journal details to dictionary
        journal_info = {
            "accreditation": accreditation
        }
        journals.append(journal_info)
    return journals

```

```

def scrape_sinta_value(search_url, indeks_jurnal_id):
    """

```

Scrapes the value of the ``<stat num>`` tag with the text "Sinta 2" from the provided URL.

Args:

`search_url` (str): The URL of the journal profile page on SINTA website.

`indeks_jurnal_id` (int): The ID of the journal in the database.

Returns:

str: The extracted Sinta value with "Sinta " replaced by "S" (or None if not found).

```

    """

```

```

    response = requests.get(search_url)
    response.raise_for_status() # Raise an exception for failed
requests

```

```

    soup = BeautifulSoup(response.content, "html.parser")

```

```

# Find the card containing "Sinta 2" text
sinta_card = soup.find("div", class_="card bg-card-orange")

# Extract the text from the `<stat num>` tag within the card (if it
exists)
sinta_value = None
if sinta_card:
    sinta_value = sinta_card.find("div", class_="stat-num").text

# Modify Sinta value if found
if sinta_value:
    sinta_value = sinta_value.split()[1] # Get the second word
(number)
    sinta_value = "S" + sinta_value # Prepend "S"

return sinta_value

def scrape_index(search_urls):
    """
    Iterates through a list of search URLs and attempts to scrape
    journals using
    scrape_journals_uin_walisongo first, then scrape_sinta_value if the
    first fails.

    Args:
        search_urls (list): A list of URLs to scrape.

    Returns:
        None
    """

    for i, url in enumerate(search_urls):
        print(f"Link no.{i+1}: ", end="")

        # Fetch journal ID using afetchone()
        sql_get_id = "SELECT id_jurnal FROM jurnal WHERE link_indeks
= (%s)"
        data = (url,) # Wrap base_url in a tuple for the query parameter
        cursor.execute(sql_get_id, data)

```

```

indeks_jurnal_id = cursor.fetchone()[0] # Extract the first
column (id_jurnal) from the result

journals = scrape_journals_uin_walisongo(url, indeks_jurnal_id)
if journals:
    for journal in journals:
        if journal['accreditation']:
            # Store accreditation in database
            sql = "UPDATE jurnal SET indeks_jurnal = %s WHERE
id_jurnal = %s"
            data = (journal['accreditation'], indeks_jurnal_id) # Use
extracted value
            cursor.execute(sql, data)
            db.commit()

            print(f"Accreditation: {journal['accreditation']}", "id: ",
indeks_jurnal_id)
        else:
            sinta_value = scrape_sinta_value(url, indeks_jurnal_id)
            if sinta_value:
                # Store Sinta value in database
                sql = "UPDATE jurnal SET indeks_jurnal = %s WHERE
id_jurnal = %s"
                data = (sinta_value, indeks_jurnal_id)
                cursor.execute(sql, data)
                db.commit()

                print(f"Sinta value: {sinta_value}", "id: ", indeks_jurnal_id)
            else:
                # Store accreditation in database
                sql = "UPDATE jurnal SET indeks_jurnal = %s WHERE
id_jurnal = %s"
                not_found = "Not Found"
                data = (not_found, indeks_jurnal_id)
                cursor.execute(sql, data)
                db.commit()

                print("Index not found.", "id: ", indeks_jurnal_id)

# Connect to database

```

```

db = mysql.connector.connect(
    host="localhost",
    username="monito29_admin",
    password="monito29_admin",
    database="monito29_db",
    charset= "utf8"
)

cursor = db.cursor()

# Retrieve search URLs from the database
cursor.execute("SELECT link_indeks FROM jurnal")
search_urls = [x[0] for x in cursor.fetchall()]

# Iterate through search URLs and scrape data
scrape_index(search_urls)

# Close database connection
db.close()

```