

proses, biasanya diawali dengan diterimanya sebuah message. (Wirfs-Brock, Rebecca, & Wilkerson, 2010)

Untuk objek-objek yang memiliki sifat khusus, standar UML mendefinisikan icon khusus untuk objek *boundary*, *controller* dan *persistent entity*.

## BAB III

### METODE PENELITIAN

#### 3.1 Penelitian Terdahulu

Tabel 3.1 Penelitian terdahulu

Judul Jurnal	Hasil Penelitian	Kesimpulan
<p><b>Judul :</b> <i>Exploiting web scraping in a collaborative filteringbased approach to web advertising</i></p> <p><b>Peneliti</b> Eloisa Vargiu1, Mirko Urru.</p> <p><b>Lokasi</b> Barcelona Digital Technology Centre, Spain</p> <p><b>Tahun</b> 2012</p> <p><b>Nama Jurnal</b> Artificial Intelligence</p>	<p>Dalam makalah ini, mereka menyajikan collaborative filtering berbasis <i>Web Advertising</i> yang menggunakan <i>Web scraping</i> teknik untuk menunjukkan iklan yang cocok untuk dimasukkan pada halaman Web. mereka merancang mengusulkan iklan Web sistem dengan memanfaatkan <i>collaborative filtering</i>. Sistem yang diusulkan, pertama, eksploitasi <i>collaborative filtering</i> dan, kemudian, bergantung pada <i>Web scraping</i> untuk mengekstrak iklan yang</p>	<p>mereka menyajikan sebuah sistem periklanan Web penyaringan kolaboratif yang memanfaatkan Teknik <i>Web Sraping</i> untuk menyarankan iklan yang sesuai ke halaman Web tertentu. Mengingat tugas <i>Web advertisement</i> sebagai penyaringan informasi, Mereka merancang sistem periklanan Web yang mengadopsi fitur penyaringan kolaboratif. Sistem yang diusulkan, pertama, bergantung pada Penyaringan kolaboratif dengan memanfaatkan halaman <i>peer</i> dan,</p>

<p>Research, 2013, Vol. 2, No. 1</p>	<p>disarankan. Ide untuk memanfaatkan collaborative filtering di Web iklan yang telah diusulkan oleh Armano &amp; Vargiu dan diadopsi juga di Armano et al.</p>	<p>selanjutnya, ia melakukan pencarian ke Web Scraping untuk melakukan konten halaman analisis. Untuk menunjukkan bagaimana sistem bekerja dalam praktik, kami menyajikan studi kasus yang sesuai, yaitu, bagaimana menyarankan iklan banner Halaman rumah portal Italia Crastulo.</p>
<p><b>Judul : Implementasi ekstraksi Web (<i>Web Scraping</i>) pada mesin pencari jurnal ilmiah menggunakan metode ekspresi regular.</b></p> <p><b>Peneliti</b> Muhammad Syapi'i Batubara</p> <p><b>Lokasi</b> Sekolah tinggi teknik harapan medan</p> <p><b>Tahun</b> 2016</p> <p><b>Nama Jurnal</b> Pusat Jurnal Sekolah tinggi teknik Harapan Medan</p>	<p>Tujuan penelitian ini adalah melakukan ekstraksi web (<i>web scraping</i>) pada mesin pencari jurnal / artikel ilmiah adalah mengotomatisasi informasi yang didapat dari website yang telah ditentukan (<i>Automate Navigation and Extraction</i>) sehingga hasil pencarian artikel atau jurnal ilmiah lebih khusus atau spesifik. Model perancangan aplikasi implementasi ekstraksi web (<i>web scraping</i>) pada mesin pencari jurnal ilmiah menggunakan metode ekspresi regular digambarkan dalam bentuk flowchart sistem. Model perancangan aplikasi implementasi ekstraksi web (<i>web scraping</i>) pada mesin pencari jurnal ilmiah menggunakan</p>	<p>Kesimpulan dari hasil implementasi ekstraksi web (<i>web scraping</i>) pada mesin pencari jurnal ilmiah menggunakan metode ekspresi regular yaitu:</p> <ol style="list-style-type: none"> <li>1. Salah satu fasilitas pendukung perkembangan internet adalah mesin pencari (<i>search engine</i>).</li> <li>2. Mesin pencari (<i>search engine</i>) merupakan kombinasi perangkat keras dan perangkat lunak komputer yang disediakan oleh perusahaan tertentu melalui website yang telah ditentukan.</li> <li>3. Aplikasi dengan menerapkan implementasi ekstraksi web (<i>web scraping</i>) pada mesin pencari memudahkan pengguna dalam melakukan pencarian jurnal ilmiah atau artikel</li> </ol>

	metode ekspresi regular dengan flowchart sistem.	yang lebih terfokus dan khusus 4. Penerapan metode ekspresi regular membantu pengguna dalam melakukan pencarian jurnal ilmiah atau artikel yang diinginkan.
<p><b>Judul :</b> <i>Web Scraping of Social Networks.</i></p> <p><b>Peneliti</b> Renita Crystal Pereira, Vanitha T</p> <p><b>Lokasi</b> Dept of M.Sc Software Technology, AIMIT, St. Aloysius College (Autonomous), Mangalore, Karnataka, India</p> <p><b>Tahun</b> 2015</p> <p><b>Nama Jurnal</b> <i>International Journal of Innovative Research in Computer and Communication Engineering</i> Vol. 3, Special Issue 7, October 2015</p>	Dalam paper ini mereka meneliti tools dan teknik apa saja yang dapat di gunakan untuk <i>Web scraping</i> serta dampaknya pada social networks. sebagai ekstraksi data tidak semudah itu. Teknik ini memastikan bahwa informasi yang dikumpulkan akurat, andal dan dapat dipercaya karena seperti yang diketahui data hadir dalam jumlah besar yang sulit dikelola dan dipelihara.	Meskipun teknik ini sangat bagus, ada beberapa tantangan yang dihadapi seperti volume <i>scraping web</i> yang tinggi dapat menyebabkan Kerusakan pada halaman web. Skala ukuran scraper web bisa berbeda dengan satuan ukuran sumbernya Sehingga membuat file sulit untuk interpretasi data. Tingkat kompleksitas sumber jika terjadi Informasi yang diekstraksi memungkinkan <i>scraping web</i> yang sangat rumit juga akan tidak berjalan

### 3.2 Objek Penelitian

Pada penelitian ini yang dijadikan objek penelitian adalah salon yang memiliki *website* di daerah Banten. Salon yang dipilih dapat berupa *Beauty Salon*, *Hair Salon*, *Nail Salon*, *Bridal Salon* maupun *Salon & Day Spa*. Berikut daftar salon yang diteliti:

**Tabel 3.2 Daftar Objek Salon Yang Diteliti**

No	Nama Salon	Lokasi
1	Salon Cantik	Curug Sangereng, Klp. Dua, Tangerang, Banten 15810
2	MayMay Salon	Blok M5 No, One Pm Building, Jl. Gading Serpong Boulevard, RT.18, Curug Sangereng, Klp. Dua, Tangerang, Banten 15810
3	Irwanteam	Bintaro Jaya Xchange Mall, Jl. Lkr. BXC Mall No.307, Pd. Jaya, Pd. Aren, Kota Tangerang Selatan, Banten 15224
4	Tokyo Belle	Supermal Karawaci FF#22, 105 Bulevar Diponegoro, Lippo Karawaci, Bencongan, Klp. Dua, Tangerang, Banten 15811
5	Natasha Skin	Jl. Boulevard Gading Serpong Blok AA-4 no 17-18, Pakulonan Barat, Kelapa Dua, Tangerang, Pakulonan Bar., Klp. Dua, Tangerang, Banten 15810
6	Joanne	Jl. Komp. Bsd, Lengkong Wetan, Serpong, Kota Tangerang Selatan, Banten 15310
7	Royal Garden	Ruko BSD Sektor IV Blok RE No. 50, Lengkong Wetan, Serpong, Lengkong Wetan, Serpong, Kota Tangerang Selatan, Banten 15310
8	Moz5	Perumnas 2, Jl. Borobudur Blok B No. 5, Karawaci, Bencongan Indah, Klp. Dua, Kota Tangerang, Banten 15810
9	Poetre Spa	jalan ciater raya, blok B1 nomer 13

		sektor 1.6, Griya Loka, BSD City, Rw. Buntu, Tangerang Selatan, Kota Tangerang Selatan, Banten 15310
10	Dian Mustika	Gading Serpong Ruko PDA no.22 Klp. Dua, Klp. Dua, Tangerang, Banten 15810
11	Naomi Beauty	Citra Raya Cikupa Blok L No. 22, Jalan Boulevard Raya, Cikupa, Tangerang, Banten 15710
12	Esther House	Jl. Jalur Sutera, Pakualam, Serpong Utara, Kota Tangerang Selatan, Banten 15320
13	CBC Beauty Care	Ruko Sektor 7 Extension Blok RM No 40 Serpong, Lengkong Wetan, Serpong, Kota Tangerang Selatan, Banten 15310
14	Dido Salon	Jalan Masjid Nurul Fajri Nomor 17B, RT 02, RW 02, Kelurahan Pondok Aren, Kecamatan Pondok Aren, Pondok Aren, South Tangerang City, Banten 15224
15	Anita salon	Ruko Prominence Blok 38F No. 6, Jl. Jalur Sutera Boulevard, Panunggangan Tim., Pinang, Kota Tangerang Selatan, Banten 15321
16	Umandaru Spa	Jalan Mandar Utama Blok DC1 no 19 Bintaro Sektor 3, Pondok Karya, Pondok Aren, Pd. Karya, Pd. Aren, South Tangerang City, Banten 15225
17	Magdalena Young	Ruko Jalur Sutera 29A No.19 Alam Sutera Serpong, Pakualam, Tangerang, Kota Tangerang Selatan, Banten 15320
18	Fourlen	Pondok Cabe Indah Blok F no. 8, Pamulang, Cipayung, Ciputat, Kota Tangerang Selatan, Banten 15418
19	Zanita	Vila Tangerang Elok Blok E1 No 7, Gelam Jaya, PasarKemis, Tangerang, Banten 15560

Tabel 3.2 merupakan daftar lokasi salon di daerah Banten yang ada dalam penelitian.

**Tabel 3.3 Daftar kategori salon yang diteliti**

<i>No</i>	<i>Hair</i>	<i>Eyelash</i>	<i>Bridal and Spa</i>	<i>Nail</i>	<i>Brow</i>
1	Salon Cantik		Salon Cantik	Salon Cantik	
2	MayMay	MayMay		MayMay	MayMay
3		Tokyo Belle			
4			Natasha Skin		
5	Irwanteam			Irwanteam	
6		Joanne			
7			Royal Garden		
8	Moz5		Moz5	Moz5	
9			Poetre Spa	Poetre Spa	
10			Dian Mustika		
11	Naomi Beauty		Naomi Beauty	Naomi Beauty	
12			Esther House		Esther House
13			CBC Beauty Care		
14	Dido Salon				
15	Anita salon	Anita Salon	Anita Salon	Anita Salon	Anita Salon
16			Umandaru Spa	Umandaru Spa	Umandaru
17			Magdalena Young		
18	Fourlen	Fourlen	Fourlen	Fourlen	
19			Zanita		

Tabel 3.3 merupakan daftar kategori salon yang ada dalam penelitian. Kategori salon dibagi menjadi *Hair, Eyelash, Bridal & Spa, Nail* dan *Brow*.

### 3.3 Metode Penelitian

#### 3.3.1 Metode Penyelesaian Masalah

Dalam penelitian ini metode yang digunakan untuk penyelesaian masalah adalah *Web scraping*. Sedangkan tools yang digunakan adalah Node.js. Mengapa Node.js cocok untuk digunakan karena Node.js dapat berjalan di sistem operasi Windows, Mac OS X dan Linux tanpa perlu ada perubahan kode program. Node.js memiliki pustaka *server HTTP* sendiri sehingga memungkinkan untuk menjalankan *server web* tanpa menggunakan program *server web* seperti Apache atau Nginx. Node.js adalah *project open source*, sehingga siapapun dapat melihat struktur kode dan juga dapat berkontribusi untuk pengembangannya.

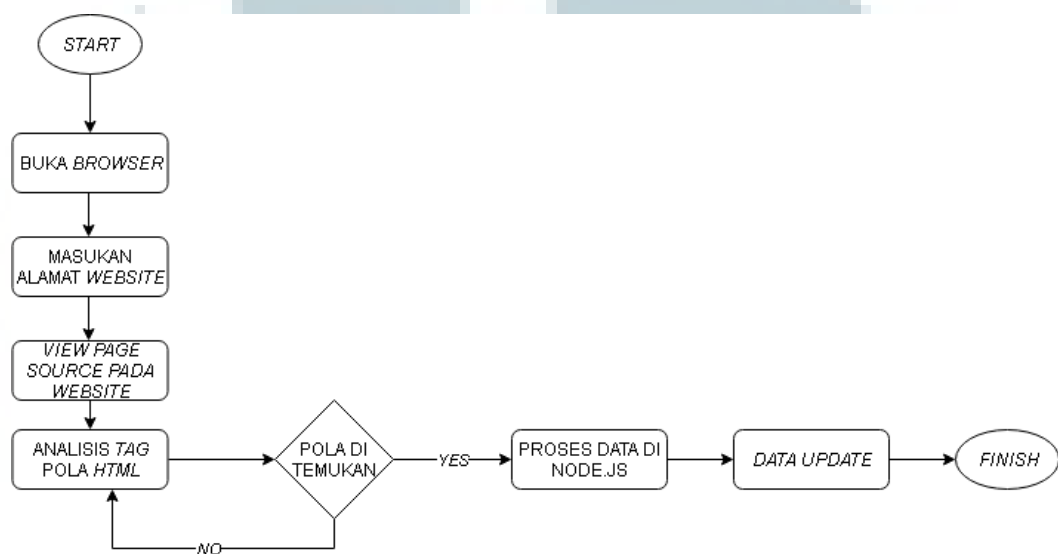
**Tabel 3.4 Perbandingan Pentaho dengan Node.js**

<i>Pentaho Data Integrations</i>	<i>Node.js</i>
Melalui Pentaho tidak dapat mengambil data berupa gambar	Menggunakan Node.js dapat mengambil data berupa gambar
Secara manual <i>import database</i> dari excel	Secara otomatis <i>database</i> akan terisi

Pada Table 3.4 ditulis bahwa jika menggunakan node.js maka dapat menscraping data berupa gambar, sehingga tidak perlu lagi secara manual mengunggah logo dari setiap salon seperti jika menggunakan Pentaho, Lalu jika pada pentaho harus *import database* terlebih dahulu di nodejs sudah terhubung

otomatis pada MongoDB, karena MongoDB *driver* merupakan *third party/Library* yang telah di sediakan oleh Node.js.

### 3.3.1.1 Web Crawling

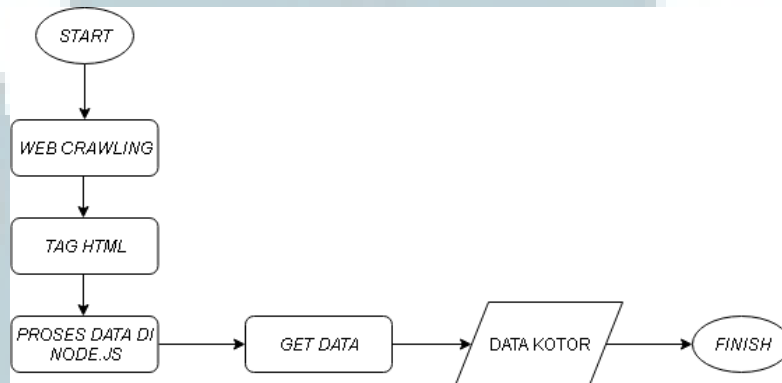


**Gambar 3.1** Flowchart proses *Web Crawling*

*Web crawling* adalah metode pertama yang dilakukan yaitu dengan mengambil data dari dua puluh lima *website* salon di daerah Banten yang telah ditentukan cukup relevan. Langkah pertama yang dilakukan dalam *web crawling* adalah mendapatkan skrip HTML dari *page source website* yang diteliti lalu menganalisa apakah di dalam skrip tersebut terdapat sebuah pola yang sama dalam tiap skrip ketika akan menampilkan data yang diperlukan untuk menampilkan Lokasi dan jasa dari salon. Setelah mendapatkan pola spesifik di dalam *script* HTML, maka akan didapatkan spesifik *tag* yang nantinya akan dipakai untuk menjalankan metode *web scraping*.



### 3.3.1.2 Web Scraping



**Gambar 3.2 Flowchart Web Scraping**

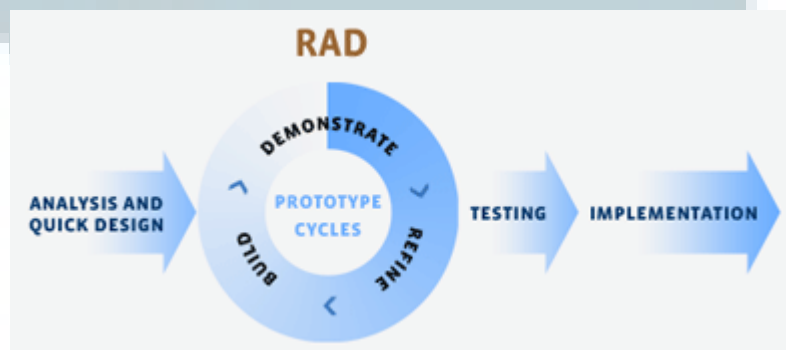
Gambar 3.2 merupakan *flowchart Web Scraping*. Metode pengambilan data yang kedua adalah *web scraping*. Setelah mendapatkan skrip HTML dari *website* yang telah ditentukan, selanjutnya Tag HTML yang sudah didapatkan pada saat melakukan *web crawling* akan diproses dan kemudian akan didapatkan sebuah tag spesifik yang merujuk kepada bagian skrip yang menampilkan informasi tentang lokasi dan jasa yang terdapat di *website* tersebut. Namun data yang didapatkan masih berupa data kotor karena masih bercampur dengan skrip HTML lainnya yang tidak berhubungan dengan informasi yang diperlukan. Maka dari itu proses selanjutnya adalah *Data Cleansing* agar data-data yang tidak diperlukan dapat dibuang.

### 3.3.2 Metode Perancangan Sistem

**Tabel 3.5 Perbandingan metode perancangan sistem (Lukman,2013)**

<b>Model</b>	<b>Sistem Kerja</b>	<b>Penekanan</b>	<b>Konsep</b>	<b>Pendekatan</b>
<b>Waterfall</b>	Proyek dibagi menjadi beberapa fase yang berurutan, dengan beberapa tumpang tindih dan <i>splashback</i> diterima antar fase.	Penekanan pada perencanaan, jadwal waktu, tanggal target, dan implementasi dari keseluruhan sistem pada satu waktu.	Kontrol ketat dijaga selama umur proyek melalui dokumentasi tertulis yang luas, <i>review</i> normal dan persetujuan oleh pengguna dan manajemen informasi teknologi yang terjadi pada akhir fase yang belum memulai tahap berikutnya.	Menerapkan pendekatan tradisional.
<b>RAD</b>	Model <i>RAD</i> hampir sama dengan model waterfall, bedanya siklus pengembangan yang ditempuh model ini sangat pendek dengan penerapan teknik yang cepat.	<i>RAD</i> menekankan pada siklus pembangunan pendek, singkat, dan cepat..	<i>Rapid application development</i> menggunakan metode iteratif (berulang) dalam mengembangkan system	Proses pembangunan perangkat lunak yang tergolong dalam teknik incremental bertingkat.

Pada Tabel 3.5 dapat disimpulkan bahwa salah satu jenis sistem yang mungkin sangat cocok menggunakan metode *waterfall* sebagai metode pengembangannya adalah sebuah sistem operasi komputer. Hal ini disebabkan karena sistem operasi komputer memiliki fungsi yang sangat kompleks, sehingga dalam pengembangannya membutuhkan analisa yang penuh mengenai kebutuhan user akan sistem operasi, hingga fitur-fitur penting yang harus dimiliki oleh sistem operasi tersebut. Maka dari itu metode *waterfall* tidak cocok karena tahapannya tidak dapat berulang dan memakan biaya yang tidak sedikit karena waktu pengembangannya yang lama. Sedangkan model *RAD* cocok dengan skala yang besar karena menggunakan metode *iteratif* (berulang) dan biasanya dapat menghemat biaya.



**Gambar 3.3 Tahapan RAD (Kendall, 2010)**

**Sumber: [www.researchgate.net](http://www.researchgate.net)**

Gambar 3.3 merupakan tahapan dari RAD, berikut ini adalah tahap-tahap pengembangan aplikasi dari tiap-tiap fase pengembangan aplikasi.

- 1) *Requirements Planning* (Perencanaan Syarat-Syarat)

Dalam fase ini, peneliti akan melakukan perencanaan bagaimana sistem akan dibuat, melakukan observasi dan mencari *website – website* yang dianggap relevan dengan perancangan aplikasi yang dilakukan. Observasi yang dilakukan adalah untuk mendapatkan *website* salon yang nantinya dari *website* tersebut dapat dipakai sebagai sumber data dalam pembuatan aplikasi.

2) *RAD Design Workshop*

Fase ini adalah fase untuk merancang dan memperbaiki sistem. Melakukan *web crawler, web scraping*, lalu diolah di *pentaho data integration*, merancang *user interface*, serta melakukan *testing*.

3) *Implementation* (Implementasi)

Pada fase implementasi ini, aplikasi yang sudah dibuat akan dipublikasikan secara *free* terlebih dahulu di *google play store*.

### 3.4 Variabel

#### 3.4.1 Variabel *independen* (Variabel – X)

Dalam penelitian ini, variabel *independen* yang berhubungan dengan masalah yang diteliti adalah Salon yang memiliki website di daerah Banten

#### 3.4.2 Variabel *Dependen* (Variabel – Y)

Dalam penelitian ini, variabel *dependen* yang berkaitan dengan masalah yang diteliti antara lain:

X<sub>1</sub> = Lokasi salon di daerah Banten

X<sub>2</sub> = Jasa yang disediakan oleh salon