



REPUBLIC INDONESIA
KEMENTERIAN HUKUM DAN HAK ASASI MANUSIA

SERTIFIKAT PATEN SEDERHANA

Menteri Hukum dan Hak Asasi Manusia atas nama Negara Republik Indonesia berdasarkan Undang-Undang Nomor 13 Tahun 2016 tentang Paten, memberikan hak atas Paten Sederhana kepada:

| | |
|--------------------------------|--|
| Nama dan Alamat Pemegang Paten | : LPPM Universitas Negeri Medan Jl. Willem Iskandar Pasar V – Kotak Pos No. 1589 Medan 20221 |
| Untuk Invensi dengan Judul | : METODE PERCAKAPAN OTOMATIS UNTUK INFORMASI PARIWISATA |
| Inventor | : Hesti Fibriasari Bakti Dwi Waluyo Baharuddin Merdy Roy Sunarya Togatorop Tansa Trisna Astono Putri |
| Tanggal Penerimaan | : 31 Oktober 2022 |
| Nomor Paten | : IDS000006506 |
| Tanggal Pemberian | : 31 Agustus 2023 |

Pelindungan Paten Sederhana untuk invenisi tersebut diberikan untuk selama 10 tahun terhitung sejak Tanggal Penerimaan (Pasal 23 Undang-Undang Nomor 13 Tahun 2016 tentang Paten).

Sertifikat Paten Sederhana ini dilampiri dengan deskripsi, klaim, abstrak dan gambar (jika ada) dari invenisi yang tidak terpisahkan dari sertifikat ini.



a.n MENTERI HUKUM DAN HAK ASASI MANUSIA
DIREKTUR JENDERAL KEKAYAAN INTELEKTUAL
u.b.

Direktur Paten, Desain Tata Letak Sirkuit Terpadu dan
Rahasia Dagang



Drs. YASMON, M.L.S.
NIP. 196805201994031002

KEMENTERIAN HUKUM DAN HAK ASASI MANUSIA RI
DIREKTORAT JENDERAL KEKAYAAN INTELEKTUAL
DIREKTORAT PATEN, DESAIN TATA LETAK SIRKUIT TERPADU DAN RAHASIA DAGANG
Jln. H.R. Rasuna Said, Kaw. 8-9 Kuningan Jakarta Selatan 12940
Phone/Fax: (6221) 57905611, Website: www.dgip.go.id

INFORMASI BIAYA TAHUNAN

Nomor Paten : IDS000006506 Tanggal diberi : 31 Agustus 2023 Jumlah Klaim : 1
Nomor Permohonan : S00202212210 Tanggal Penerimaan : 31 Oktober 2022

Berdasarkan Peraturan Pemerintah Nomor 28 tahun 2019 tentang Jenis dan Tarif Atas Jenis Penerimaan negara Bukan Pajak Yang Berlaku Pada Kementerian Hukum dan Hak Asasi Manusia, biaya tahunan yang harus dibayarkan adalah sebagaimana dalam tabel di bawah.

Perhitungan biaya tahunan yang sudah dibayarkan adalah :

| Biaya Tahunan Ke- | Periode Perlindungan | Batas Akhir Pembayaran | Tgl Pembayaran | Jumlah Pembayaran | Keterangan |
|-------------------|-----------------------|------------------------|----------------|-------------------|-----------------------------------|
| 1 | 31/10/2022-30/10/2023 | 28/02/2024 | undefined | 0 | Klaim 1, Total Klaim: 0, Denda: 0 |
| 2 | 31/10/2023-30/10/2024 | 28/02/2024 | undefined | 0 | Klaim 1, Total Klaim: 0, Denda: 0 |
| 3 | 31/10/2024-30/10/2025 | 02/10/2024 | undefined | 0 | Klaim 1, Total Klaim: 0, Denda: 0 |
| 4 | 31/10/2025-30/10/2026 | 02/10/2025 | undefined | 0 | Klaim 1, Total Klaim: 0, Denda: 0 |
| 5 | 31/10/2026-30/10/2027 | 02/10/2026 | undefined | 0 | Klaim 1, Total Klaim: 0, Denda: 0 |

Perhitungan biaya tahunan yang belum dibayarkan adalah :

| Biaya Tahunan Ke- | Periode Perlindungan | Batas Akhir Pembayaran | Biaya Dasar | Jml Klaim | Biaya Klaim | Total | Terlambat (Bulan) | Total Denda | Jumlah Pembayaran |
|-------------------|-----------------------|------------------------|-------------|-----------|-------------|-----------|-------------------|-------------|-------------------|
| 6 | 31/10/2027-30/10/2028 | 02/10/2027 | 1.650.000 | 1 | 50.000 | 1.700.000 | 0 | 0 | 1.700.000 |
| 7 | 31/10/2028-30/10/2029 | 02/10/2028 | 2.200.000 | 1 | 50.000 | 2.250.000 | 0 | 0 | 2.250.000 |
| 8 | 31/10/2029-30/10/2030 | 02/10/2029 | 2.750.000 | 1 | 50.000 | 2.800.000 | 0 | 0 | 2.800.000 |
| 9 | 31/10/2030-30/10/2031 | 02/10/2030 | 3.300.000 | 1 | 50.000 | 3.350.000 | 0 | 0 | 3.350.000 |
| 10 | 31/10/2031-30/10/2032 | 02/10/2031 | 3.850.000 | 1 | 50.000 | 3.900.000 | 0 | 0 | 3.900.000 |

Biaya yang harus dibayarkan hingga tanggal 02-10-2027 (tahun ke-6) adalah sebesar Rp.1.700.000,-

- Pembayaran biaya tahunan untuk pertama kali wajib dilakukan paling lambat 6 (enam) bulan terhitung sejak tanggal diberi paten
- Pembayaran biaya tahunan untuk pertama kali meliputi biaya tahunan untuk tahun pertama sejak tanggal penerimaan sampai dengan tahun diberi Paten ditambah biaya tahunan satu tahun berikutnya.
- Pembayaran biaya tahunan selanjutnya dilakukan paling lambat 1 (satu) bulan sebelum tanggal yang sama dengan Tanggal Penerimaan pada periode pelindungan tahun berikutnya.
- Permohonan penundaan pembayaran biaya tahunan akan diterima apabila diajukan paling lama 7 hari kerja sebelum tanggal jatuh tempo pembayaran biaya tahunan berikutnya, dan bukan merupakan pembayaran biaya tahunan pertama kali.
- Dalam hal biaya tahunan belum dibayarkan sampai dengan jangka waktu yang ditentukan, Paten dinyatakan dihapus



(12) PATEN INDONESIA

(19) DIREKTORAT JENDERAL
KEKAYAAN INTELEKTUAL

(11) IDS000006506 B

(45) 31 Agustus 2023

(51) Klasifikasi IPC[®] : G 06F 16/951(2019.01), G 06Q 50/14(2012.01), H 04L 51/02(2022.01)

(21) No. Permohonan Paten : S00202212210

(22) Tanggal Penerimaan: 31 Oktober 2022

(30) Data Prioritas :

(31) Nomor (32) Tanggal (33) Negara

(43) Tanggal Pengumuman: 09 November 2022

(56) Dokumen Pembedang:

US 9.369.410 B2
US 11.392.647 B2
US 11.451.496 B1

(71) Nama dan Alamat yang Mengajukan Permohonan Paten :
LPPM Universitas Negeri Medan
Jl. Willem Iskandar Pasar V – Kotak Pos No. 1589
Medan 20221

(72) Nama Inventor :
Hesti Fibriasari, ID
Bakti Dwi Waluyo, ID
Baharuddin, ID
Merdy Roy Sunarya Togatorop, ID
Tansa Trisna Astono Putri, ID

(74) Nama dan Alamat Konsultan Paten :

Pemeriksa Paten : Nico E. Soelistyono, ST.

Jumlah Klaim : 1

(54) Judul Invensi : METODE PERCAKAPAN OTOMATIS UNTUK INFORMASI PARIWISATA

(57) Abstrak :

Invensi ini mengenai metode percakapan otomatis untuk informasi pariwisata yang menggabungkan teknik *harvesting* data untuk memberikan jawaban secara *real-time* sesuai pertanyaan pengguna. Metode ini memanfaatkan algoritma *Knuth-morris-pratt*, *web crawling*, dan *scraping* data untuk mengumpulkan informasi tempat wisata, atraksi, akomodasi, dan fasilitas lain yang relevan. Tujuan utama invensi ini adalah untuk mengatasi permasalahan yang ada pada metode percakapan otomatis sebelumnya dalam menjawab pertanyaan pengguna. Metode ini menyediakan sistem percakapan otomatis yang dapat memberikan informasi pariwisata secara cepat dan akurat. Penerapan teknik *harvesting* data melibatkan langkah-langkah seperti pencocokan pola teks menggunakan algoritma *Knuth-morris-pratt* untuk menganalisis pertanyaan pengguna. Selanjutnya, *web crawling* digunakan untuk mencari data dari berbagai halaman web berdasarkan kata kunci yang diambil dari pertanyaan pengguna, dan data yang relevan disimpan dalam basis data. Teknik *scraping* data kemudian digunakan untuk mengekstrak informasi dari basis data dan menghasilkan jawaban yang sesuai dengan pertanyaan pengguna. Jawaban-jawaban tersebut selalu terkini karena bersumber dari informasi *real-time* yang diperoleh dari internet. Dengan penerapan teknik *harvesting* data ini, sistem percakapan otomatis dapat memberikan jawaban yang terstruktur, cepat, dan akurat kepada pengguna mengenai informasi pariwisata. Inovasi ini berpotensi untuk meningkatkan pengalaman wisatawan dalam merencanakan perjalanan dan memperkaya pengetahuan tentang destinasi pariwisata.



Deskripsi

METODE PERCAKAPAN OTOMATIS UNTUK INFORMASI PARIWISATA

Bidang Teknik Invensi

5 Invensi ini mengenai metode percakapan otomatis untuk informasi pariwisata. Lebih khusus lagi, metode percakapan otomatis sesuai invensi ini menggunakan teknik *harvesting data*, dimana teknik ini merupakan cara untuk menggabungkan algoritma *knuth-morris-pratt*, *web crawling*, dan *scraping data*.
10 Teknik *harvesting data* dalam invensi ini diarahkan untuk mengumpulkan informasi tempat-tempat wisata, atraksi, akomodasi, dan fasilitas lain yang relevan sesuai pertanyaan pengguna secara langsung dan terkini.

15 **Latar Belakang Invensi**

Perjalanan wisata telah menjadi salah satu industri yang sangat penting dan berkembang pesat di seluruh dunia. Para wisatawan sering mencari informasi tentang tempat wisata, akomodasi, atraksi, dan fasilitas lain sebelum atau selama perjalanan.
20

Sebelum adanya teknologi percakapan otomatis, pengguna selalu mencari informasi di situs web, menghubungi pusat informasi, atau mengirimkan surat elektronik untuk mendapatkan informasi pariwisata. proses ini dapat memakan waktu lama dan kadang-kadang sulit untuk menemukan informasi yang akurat dan terkini. Dengan adanya perkembangan teknologi, terutama dalam pemrosesan bahasa alami, maka muncullah potensi untuk mengembangkan sistem yang dapat memberikan informasi pariwisata secara otomatis, akurat, dan terkini.
25

30 Invensi teknologi yang berkaitan dengan percakapan otomatis sesuai invensi ini berupa aplikasi *chatbot*, yang mana telah diungkapkan sebagaimana terdapat pada paten Amerika Serikat Nomor US 9,369,410 B2 Tanggal 14 Juni 2016 dengan judul



"CHATBOTS". Dalam invensi ini berkaitan dengan *chatbot*, yaitu agen komputer yang dirancang untuk melakukan percakapan dengan koresponden manusia. Secara khusus, penemuan ini menyangkut metode untuk menerapkan perangkat lunak ke dalam sistem komputer sehingga mampu melakukan percakapan dengan manusia. Namun invensi ini masih terdapat kelemahan, dimana *database* atau basis data yang dimiliki untuk menjawab pertanyaan pengguna masih disimpan dan diperbaharui secara manual oleh pengelola sistem. Sehingga jika pertanyaan dari pengguna tidak cocok dengan jawaban yang tersimpan dalam basis data, *chatbot* tidak akan menjawab dan memberikan instruksi lainnya.

Paten Amerika Serikat Nomor US 11,392,647 B2 Tanggal 19 Juli 2022 dengan judul "*Intent-based question suggestion engine to advance a transaction conducted via a chatbot*". Paten ini menjelaskan tentang sebuah aplikasi *chatbot* yang memiliki kemampuan untuk memberikan saran pertanyaan berikutnya kepada pengguna selama melakukan percakapan langsung (*live chat*). Saat aplikasi *chatbot* berada dalam mode percakapan langsung, terdapat komponen yang menganalisa pertanyaan dari pengguna. Hasil analisis dari pertanyaan tersebut menjadi pertanyaan prediksi yang diberikan *chatbot* kepada pengguna. Sehingga hasilnya, aplikasi *chatbot* memberikan kepada pengguna setidaknya satu saran pertanyaan berikutnya yang mungkin ingin ditanyakan sebagai tanggapan atas pertanyaan pertama pengguna. Namun invensi ini hanya berfokus memberikan panduan bagi pengguna mengenai pertanyaan-pertanyaan selanjutnya, dan *chatbot* dalam invensi ini dalam memberi jawaban kepada pengguna terhubung ke jaringan komputer. Pada invensi ini juga disebutkan bahwa dapat menggunakan jaringan publik atau privat. Sehingga dalam sistem ini tidak tergambar jelas bagaimana jawaban dari pertanyaan pengguna diperoleh, apakah secara manual atau otomatis.



Invensi lainnya sebagaimana pada paten Amerika Serikat Nomor US 11,451,496 B1 Tanggal 20 September 2022 dengan judul "Intelligent, Personalized, and Dynamic Chatbot Conversation", dimana diungkapkan *chatbot* menggantikan fasilitas pertanyaan yang sering diajukan (*frequently asked questions, FAQ*) berbasis teks pada situs web. Konsepnya adalah *chatbot* akan menafsirkan setiap pertanyaan dari pengguna dan memberikan jawaban berdasarkan data teks yang telah disiapkan, namun invensi tersebut masih terdapat kekurangan yaitu jawaban yang diberikan kepada pengguna disimpan dan diperbarui secara manual oleh *programmer* atau admin sistem. Oleh sebab itu, pengguna tidak akan mendapatkan informasi terbaru mengenai pariwisata sebelum admin menambahkan informasi terbaru di dalam basis datanya. Mengingat tempat wisata pada umumnya bertambah dan berkembang, maka kebutuhan akan informasi terbaru sangat dibutuhkan.

Selanjutnya invensi ini diajukan dengan maksud untuk mengatasi permasalahan yang telah dikemukakan sebelumnya, yaitu cara mencari, menyimpan, dan memberikan jawaban. Dalam hal ini jawaban yang dihasilkan berasal dari informasi-informasi yang didapatkan dari internet secara *real-time* berdasarkan pertanyaan dari pengguna. Jawaban yang diberikan kepada pengguna diperoleh dengan cara memanen informasi dari berbagai sumber yang ada di situs web menggunakan teknik *harvesting data*.

Uraian Singkat Invensi

Tujuan utama dari invensi ini adalah untuk mengatasi permasalahan yang telah ada sebelumnya khususnya metode percakapan otomatis yang digunakan untuk menjawab pertanyaan dari pengguna. Dalam invensi ini, sistem percakapan otomatis telah diperbaiki sehingga dapat memberikan informasi pariwisata kepada pengguna secara *real-time*.

1



Untuk mencapai tujuan yang disebutkan diatas, invensi ini menyediakan metode percakapan otomatis untuk informasi pariwisata yang dilengkapi dengan teknik *harvesting data*, yang terdiri dari: pada perangkat pengguna (1): menerima masukan pengguna yang terkait dengan pertanyaan mengenai pariwisata, pertanyaan mengenai pariwisata tersebut diproses oleh server (3) melalui jaringan internet (2); pada server (3): memproses pertanyaan mengenai pariwisata dari perangkat pengguna (1) menggunakan pemrosesan teks (6) untuk meningkatkan relevansi dan format yang sesuai, teks diproses menggunakan algoritma *knuth-morris-pratt* (7) untuk mencari pola pertanyaan dalam teks, algoritma *knuth-morris-pratt* menggunakan tabel lompatan (18) untuk menghindari pengulangan pencarian pola yang tidak perlu; mencatat hasil pencocokan pola (20), dan menghasilkan kata kunci sebagai pola pertanyaan terakhir; memasukkan kata kunci ke proses *web crawling* (10) untuk mengumpulkan data pariwisata dari berbagai situs *web* sesuai dengan kata kunci, proses *web crawling* (10) dilakukan dengan mengirimkan permintaan (*HTML request*) ke situs *web*, menganalisis kode *HTML*, dan mengekstrak data yang sesuai dengan kata kunci; menyimpan data pariwisata ke dalam basis data (11) untuk diproses ke proses *scraping data* (13), proses *scraping data* (13) melibatkan pengambilan data dari basis data (30), mengekstraksi teks dari elemen data, melakukan pembersihan data (32), dan merubah data menjadi informasi terstruktur sesuai kata kunci (33); menampilkan informasi yang dihasilkan sebagai jawaban atas pertanyaan pengguna (34); dan memonitor dan memperbaharui data yang dikumpulkan secara teratur untuk memastikan keakuratan dan keterkinian informasi kepada pengguna.

Peningkatan dalam metode percakapan otomatis ini dilakukan dengan mengintegrasikan teknik *harvesting data*. Teknik ini menggabungkan algoritma *Knuth-morris-pratt*, *web crawling*, dan

PK



scraping data. Melalui teknik *harvesting data* ini, sistem percakapan otomatis dapat memberikan jawaban yang lebih akurat dan relevan terhadap pertanyaan pengguna.

Pertama-tama, sistem menggunakan algoritma *Knuth-morris-pratt* untuk mencocokkan pola dalam teks pertanyaan yang diberikan oleh pengguna. Selanjutnya, dengan menggunakan teknik *web crawling*, pola teks yang teridentifikasi dari pertanyaan pengguna akan digunakan sebagai kata kunci untuk melakukan pencarian dan pengambilan data dari berbagai halaman web yang relevan.

Data yang berhasil diambil dari berbagai halaman web tersebut akan disimpan dalam basis data. Kemudian, sistem menggunakan teknik *scraping data* untuk mengekstrak informasi dari data yang telah disimpan dalam basis data. Informasi ini akan diolah menjadi jawaban yang sesuai dengan pertanyaan yang diajukan oleh pengguna.

Sebelum jawaban akhirnya ditampilkan kepada pengguna, informasi yang ditemukan juga akan disimpan dalam basis data sebagai jawaban terbaru. Hal ini memastikan bahwa jawaban yang diberikan oleh sistem selalu mengandalkan informasi terkini yang diperoleh dari internet secara *real-time*.

Dengan penerapan teknik *harvesting data* ini, sistem percakapan otomatis ini menjadi lebih terstruktur dan responsif dalam memberikan jawaban kepada pengguna. Pengguna dapat yakin bahwa jawaban yang mereka terima selalu didasarkan pada informasi terkini dari sumber-sumber terpercaya. Invensi ini berpotensi memberikan pengalaman berbicara dengan sistem percakapan otomatis yang lebih efisien.

30 **Uraian Singkat Gambar**

Gambar 1 adalah gambaran umum sistem percakapan otomatis untuk informasi pariwisata sesuai dengan invensi ini.

PF



Gambar 2 adalah diagram alir (*flowchart*) metode percakapan otomatis yang menggunakan teknik *harvesting data*.

Gambar 3 adalah cara kerja algoritma *knuth-morris-pratt*.

Gambar 4 adalah cara kerja *web crawling*.

5 Gambar 5 adalah cara kerja *scraping data*.

Uraian Lengkap Invensi

Invensi ini bertujuan untuk menciptakan sebuah metode percakapan otomatis yang menyediakan informasi pariwisata dengan menggunakan teknik *harvesting data*. Teknik *harvesting data* berfungsi untuk menggabungkan tiga algoritma utama yang terdiri dari: algoritma *knuth-morris-pratt*; *web crawling*; *scraping data*. Selanjutnya, masing-masing unit tersebut dijelaskan dengan pencirian sebagai berikut.

15 Algoritma *knuth-morris-pratt* adalah algoritma pencocokan pola yang digunakan untuk mencari kemunculan suatu pola (*pattern*) dalam teks (*text*). Dengan memanfaatkan informasi yang diperoleh selama pencocokan sebelumnya, algoritma ini menghindari pengulangan pencarian pola yang tidak perlu.
20 Algoritma ini bertugas untuk mengubah pertanyaan pengguna menjadi "kata kunci" yang akan diproses oleh tahap berikutnya, yaitu *web crawling*.

25 *Web crawling* adalah proses pengumpulan data dari berbagai situs *web* secara sistematis dan besar-besaran. Data yang dikumpulkan tidak beraturan dan hanya disesuaikan dengan kata kunci. Setelah data terkumpul, informasi yang relevan akan diekstrak dan disimpan dalam basis data untuk tahap selanjutnya, yaitu *scraping data*.

30 Proses *scraping data* membaca dan mengekstrak informasi sesuai dengan kata kunci dari basis data. Proses ini mirip dengan cara manusia membaca dan menyusun informasi. Hasil dari proses *scraping data* akan disimpan dalam basis data sebelum ditampilkan sebagai jawaban kepada pengguna. Jika ada

PK



pertanyaan serupa dari pengguna lain, informasi yang telah disimpan akan dibandingkan dengan data terbaru untuk memastikan jawaban yang diberikan selalu terkini dan *real-time*.

5 Metode percakapan pintar yang dihasilkan dari invensi ini (Gambar 1) melibatkan perangkat pengguna (1), jaringan internet (2), server (3), dan pengelola sistem (4). Pengguna dapat berinteraksi dengan sistem percakapan otomatis ini melalui berbagai perangkat seperti komputer jinjing, komputer meja, atau telepon pintar. Selanjutnya, terdapat server (3) yang menyediakan layanan penyimpanan, pemrosesan data, dan sebagai tempat teknik *harvesting data* dikonfigurasikan. Server (3) dalam invensi ini dikelola oleh pengelola sistem (4), yang bertugas sebagai pengelola, mengkonfigurasi, memelihara, dan menjaga keamanan. Berikutnya untuk dapat berinteraksi dengan pengguna, sistem percakapan otomatis ini harus terhubung dengan jaringan internet (2).

Mengacu pada Gambar 2, diperlihatkan diagram alir (*flowchart*) metode percakapan otomatis yang menggunakan teknik *harvesting data*. Dimana teknik ini dimulai dari pengguna memasukkan pertanyaan mengenai pariwisata melalui perangkat pengguna (1). Pertanyaan dari pengguna berupa teks akan masuk ke pemrosesan teks (6), dalam tahap ini pertanyaan dari pengguna ditingkatkan relevansinya, dan merubah teks menjadi format yang lebih sesuai untuk pemrosesan lebih lanjut. Sebagai contoh, menghilangkan tanda baca, simbol, atau karakter khusus yang tidak memberikan banyak informasi atau bahkan mengganggu analisis teks. Setelah teks selesai dilakukan pemrosesan, selanjutnya teks diproses menggunakan 25 algoritma *knuth-morris-pratt* (7), lebih jelas pada Gambar 3.

Mengacu pada Gambar 3, mengenai cara kerja algoritma *knuth-morris-pratt*. Berikut adalah langkah-langkah kerja algoritma *knuth-morris-pratt*: dimulai dari tahap persiapan

PK



(17), dimana terlebih dahulu menentukan pola (*pattern*) pertanyaan yang ingin dicari dalam teks dan teks (*text*) yang sudah dipersiapkan, terutama kata atau kalimat yang berhubungan dengan pariwisata. Selanjutnya, membuat tabel lompatan (18), hal ini tabel dibuat secara otomatis dengan ukuran yang sama dengan panjang pola. Inisialisasi tabel lompatan (18) dengan nilai 0 pada setiap elemennya, kemudian isi tabel lompatan (18) berdasarkan teks yang ingin dicari. Proses mengisi tabel lompatan (18) dapat dilakukan sebelum proses pencocokan dimulai, sehingga hanya perlu dilakukan sekali. Selanjutnya adalah melakukan pencocokan pola (19) dari awal teks dan awal pola, dimana jika karakter pada indeks teks dan indeks pola sama maka lanjutkan pencocokan dengan meningkatkan nilai indeks teks dan indeks pola. Sebaliknya, jika karakter pada teks dan pola berbeda maka periksa tabel lompatan (18) untuk menentukan jumlah karakter yang dapat dilewati pola, geser pola ke kiri sejauh nilai yang ada di tabel lompatan (18) untuk karakter yang cocok terakhir di pola. Langkah tersebut akan diulangi sampai ditemukan pola yang cocok atau mencapai akhir teks. Langkah terakhir adalah hasil pencocokan pola (20), dimana jika ditemukan pola dalam teks, catat posisi atau indeks awal dari kemunculan pola tersebut. Jika tidak ditemukan pola teks, maka pola tidak ada dalam teks. Selanjutnya, apabila terjadi kecocokan maka mengeluarkan hasil pencarian yang menjadi "kata kunci" sebagai pola pertanyaan terakhir

Masih mengacu pada Gambar 2, setelah "kata kunci" dihasilkan, maka langkah selanjutnya adalah kata kunci masuk ke bagian *web crawling* (10). Lebih detail mengenai cara kerja *web crawling* pada Gambar 4. Mengacu pada Gambar 4, algoritma *web crawling* akan menentukan tujuan sesuai kata kunci (21), yaitu menentukan data pariwisata apa yang ingin dikumpulkan. Setelah menetapkan tujuan, selanjutnya mengirimkan permintaan



(HTTP request) ke situs web yang akan dijelajahi (22). Permintaan ini akan dikirimkan ke server situs web, dan server akan merespon dengan mengirimkan halaman web yang diminta. Setelah server merespon permintaan, web crawler akan menerima halaman web tersebut sebagai balasan. Halaman web ini berisi kode HTML yang menyusun tampilan visual situs web (23). Web crawler akan menganalisis kode HTML halaman web untuk mengidentifikasi informasi sesuai kata kunci (24). Setelah menganalisis halaman web, web crawler akan mengekstrak data yang diperlukan sesuai dengan kata kunci (25). Kemudian data ini disimpan ke dalam basis data dalam format teks. Proses web crawling dapat melibatkan banyak halaman web, jadi perlu ada mekanisme penjadwalan untuk mengatur bagaimana halaman-halaman ini akan diakses dan dijelajahi (26). Dalam hal ini, web crawler dapat mengikuti tautan halaman yang telah diambil untuk menjelajahi lebih banyak halaman web. Beberapa situs web mungkin memiliki kebijakan yang melarang web crawler mengakses atau menjelajahi halaman web, oleh karena itu penting untuk mengikuti aturan ini dengan mengizinkan atau menolak akses halaman tertentu sesuai kebijakan situs (27). Setelah data diambil, seringkali diperlukan langkah pembersihan untuk menghapus informasi yang tidak relevan atau duplikat (28). Data yang telah diolah dan disaring kemudian disimpan sesuai dengan format yang dibutuhkan. Proses web crawling bisa menjadi tugas yang berulang dan terus-menerus untuk mengumpulkan data terbaru (29). Monitoring dan pengulangan secara teratur diperlukan untuk memastikan data yang diambil selalu mutakhir sesuai kata kunci.

Kembali mengacu kepada Gambar 2, setelah melewati web crawling (10), data yang dihasilkan dilakukan penyimpanan (11) ke basis data (16). Selanjutnya data dari berbagai situs web berdasarkan kata kunci (12) masuk ke scraping data (13). Lebih jelas mengenai cara kerja scraping data (13) dapat dilihat

NP
-



pada Gambar 5. Dimana langkah pertama dalam proses *scraping* data adalah mengambil data dari basis data (30), langkah ini memerlukan analisis struktur data untuk mengidentifikasi informasi yang ingin diambil sesuai kata kunci. Selanjutnya,
5 data yang sudah diidentifikasi diekstraksi (31) untuk mengambil teks dari elemen data, misalnya, jika ingin mengambil judul pariwisata dari situs pariwisata, maka langkahnya adalah menemukan elemen yang berisi judul tersebut dan kemudian mengambil teks dari elemen tersebut. Langkah
10 selanjutnya adalah melakukan pembersihan data (32), hal ini untuk menghapus karakter khusus, mengkonversi format tanggal, atau melakukan langkah pembersihan lainnya agar data menjadi lebih terstruktur dan mudah diolah. Kemudian data dikumpulkan sesuai kata kunci dan dirubah menjadi informasi (33) dan
15 informasi ditampilkan dalam bentuk terstruktur dalam bentuk jawaban sesuai kata kunci yang diminta pengguna (34).

Kembali kepada Gambar 2, informasi yang dihasilkan dari proses *scraping data* (13), selanjutnya disimpan (14) ke dalam basis data (16). Proses ini diperlukan untuk membandingkan
20 informasi yang telah disimpan sebelumnya dengan informasi baru hal ini dimaksudnya untuk menambah cakupan informasi yang akan diberikan kepada pengguna berikutnya. Informasi ini lah yang menjadi jawaban kepada pengguna mengenai informasi pariwisata yang ditanyakan(15), sehingga jawaban yang diberikan pengguna
25 akan selalu terkini (*up to date*) dan waktu nyata (*real-time*).

Dari uraian diatas jelas bahwa hasil invensi ini dapat menjadi penyempurna dari invensi-invensi sebelumnya, khususnya pada teknik pencarian, penyimpanan, dan pemberian jawaban pada sistem percakapan otomatis. Kelebihan invensi ini dimana
30 pencarian, penyimpanan, dan pemberian jawaban tidak perlu dimasukkan manual dan diperbaharui oleh pengelola sistem secara manual. Sistem percakapan otomatis ini dalam memberikan

PC



jawaban kepada pertanyaan pengguna sepenuhnya otomatis, terkini, dan waktunya.

**Klaim**

1. Suatu metode percakapan otomatis untuk informasi pariwisata yang dilengkapi dengan teknik *harvesting data*, yang terdiri dari:
 - 5 pada perangkat pengguna (1):

menerima masukan pengguna yang terkait dengan pertanyaan mengenai pariwisata, pertanyaan mengenai pariwisata tersebut diproses oleh server (3) melalui jaringan internet (2);
 - 10 pada server (3):

memproses pertanyaan mengenai pariwisata dari perangkat pengguna (1) menggunakan pemrosesan teks (6) untuk meningkatkan relevansi dan format yang sesuai, teks diproses menggunakan algoritma *knuth-morris-pratt* (7) untuk mencari pola pertanyaan dalam teks, algoritma *knuth-morris-pratt* menggunakan tabel lompatan (18) untuk menghindari pengulangan pencarian pola yang tidak perlu;
 - 15 mencatat hasil pencocokan pola (20), dan menghasilkan kata kunci sebagai pola pertanyaan terakhir;
 - 20 memasukkan kata kunci ke proses *web crawling* (10) untuk mengumpulkan data pariwisata dari berbagai situs *web* sesuai dengan kata kunci, proses *web crawling* (10) dilakukan dengan mengirimkan permintaan (*HTML request*) ke situs *web*, menganalisis kode *HTML*, dan mengekstrak data yang sesuai dengan kata kunci;
 - 25 menyimpan data pariwisata ke dalam basis data (11) untuk diproses ke proses *scraping data* (13), proses *scraping data* (13) melibatkan pengambilan data dari basis data (30), mengekstraksi teks dari

PK



elemen data, melakukan pembersihan data (32), dan merubah data menjadi informasi terstruktur sesuai kata kunci (33);

5

menampilkan informasi yang dihasilkan sebagai jawaban atas pertanyaan pengguna (34); dan

memonitor dan memperbaharui data yang dikumpulkan secara teratur untuk memastikan keakuratan dan keterkinian informasi kepada pengguna.

10

14

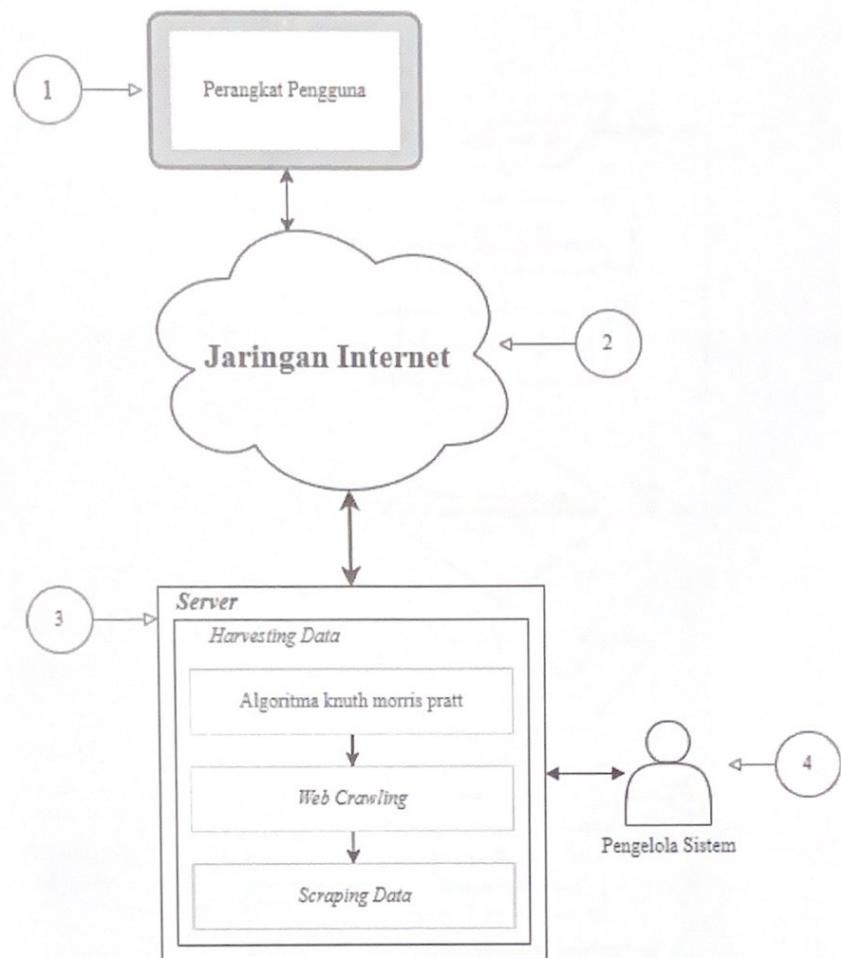


Abstrak

METODE PERCAKAPAN OTOMATIS UNTUK INFORMASI PARIWISATA

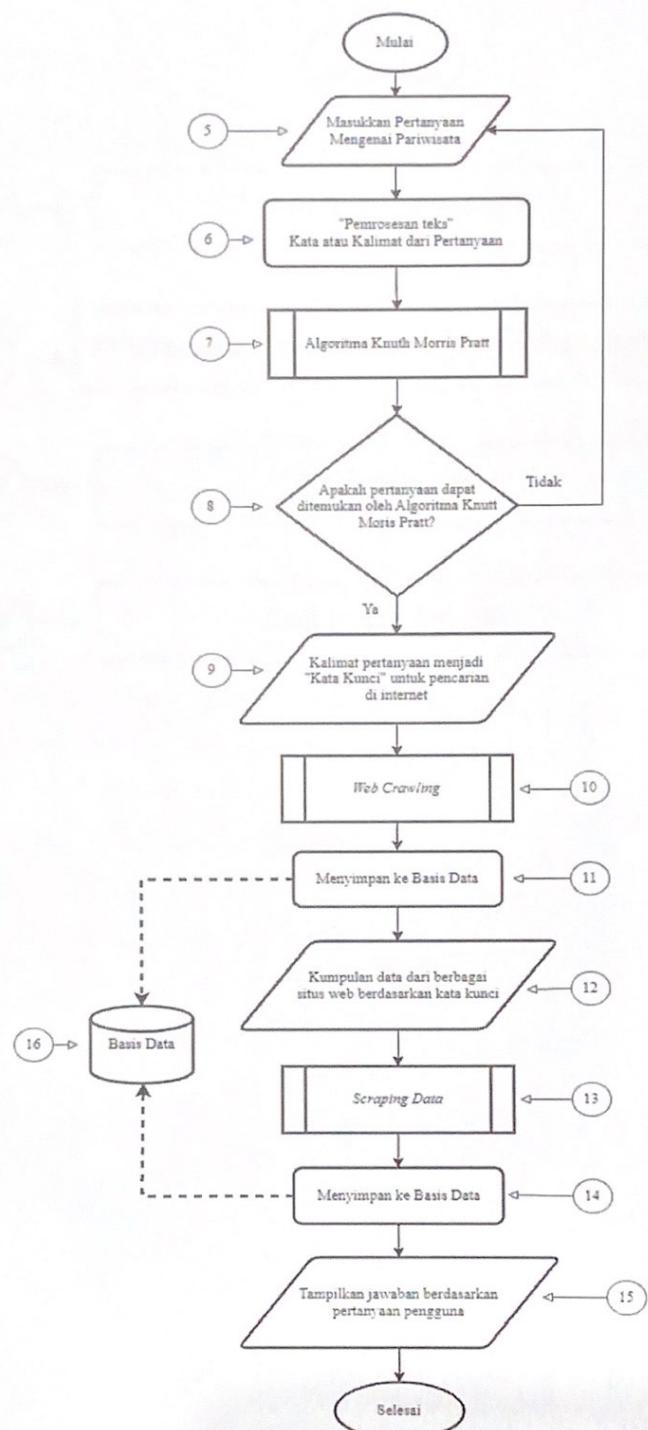
5 Invensi ini mengenai metode percakapan otomatis untuk informasi pariwisata yang menggabungkan teknik *harvesting data* untuk memberikan jawaban secara *real-time* sesuai pertanyaan pengguna. Metode ini memanfaatkan algoritma *Knuth-morris-pratt*, *web crawling*, dan *scraping data* untuk mengumpulkan
10 informasi tempat wisata, atraksi, akomodasi, dan fasilitas lain yang relevan. Tujuan utama invensi ini adalah untuk mengatasi permasalahan yang ada pada metode percakapan otomatis sebelumnya dalam menjawab pertanyaan pengguna. Metode ini menyediakan sistem percakapan otomatis yang dapat memberikan informasi pariwisata secara cepat dan akurat.
15 Penerapan teknik *harvesting data* melibatkan langkah-langkah seperti pencocokan pola teks menggunakan algoritma *Knuth-morris-pratt* untuk menganalisis pertanyaan pengguna. Selanjutnya, *web crawling* digunakan untuk mencari data dari
20 berbagai halaman *web* berdasarkan kata kunci yang diambil dari pertanyaan pengguna, dan data yang relevan disimpan dalam basis data. Teknik *scraping data* kemudian digunakan untuk mengekstrak informasi dari basis data dan menghasilkan jawaban yang sesuai dengan pertanyaan pengguna. Jawaban-jawaban
25 tersebut selalu terkini karena bersumber dari informasi *real-time* yang diperoleh dari internet. Dengan penerapan teknik *harvesting data* ini, sistem percakapan otomatis dapat memberikan jawaban yang terstruktur, cepat, dan akurat kepada pengguna mengenai informasi pariwisata. Inovasi ini berpotensi
30 untuk meningkatkan pengalaman wisatawan dalam merencanakan perjalanan dan memperkaya pengetahuan tentang destinasi pariwisata.

JK



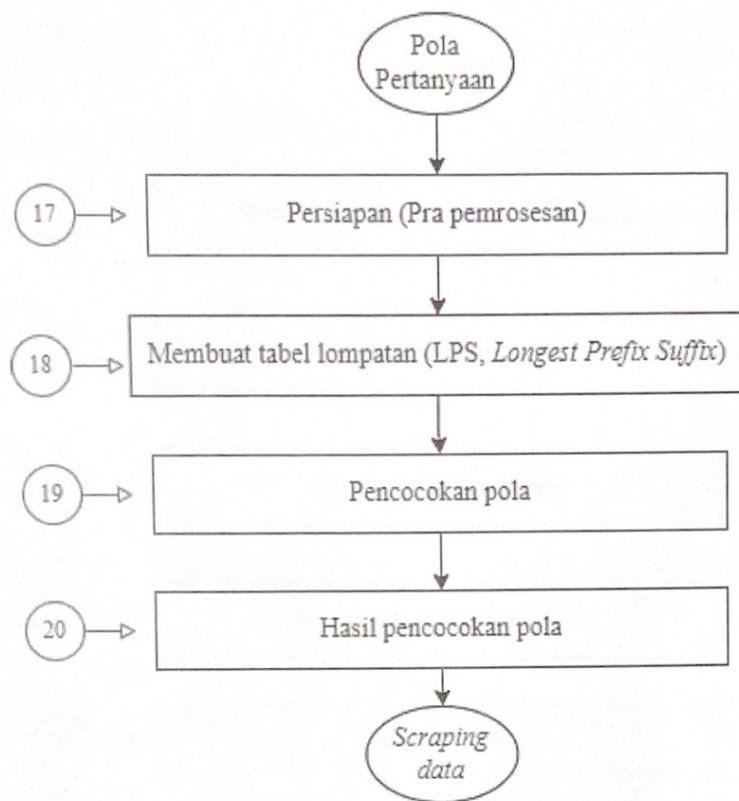
Gambar 1

11

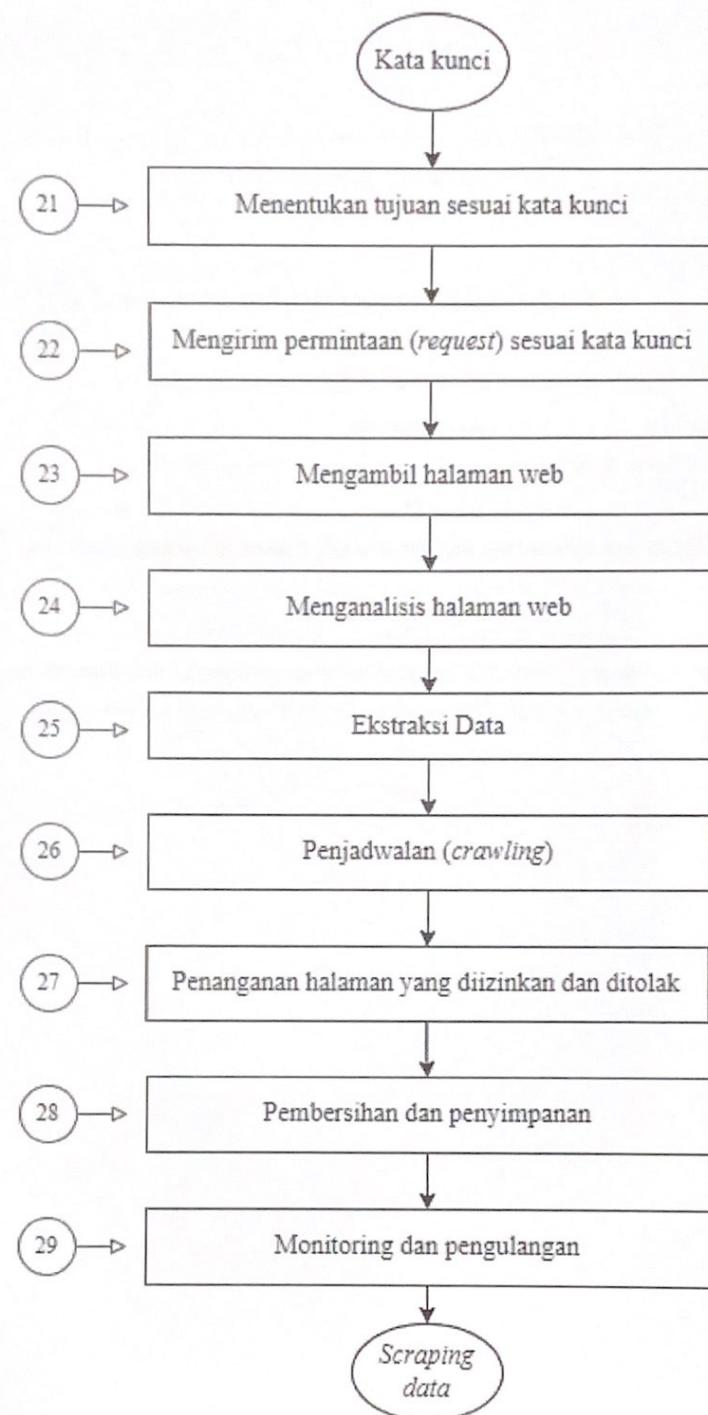


Gambar 2

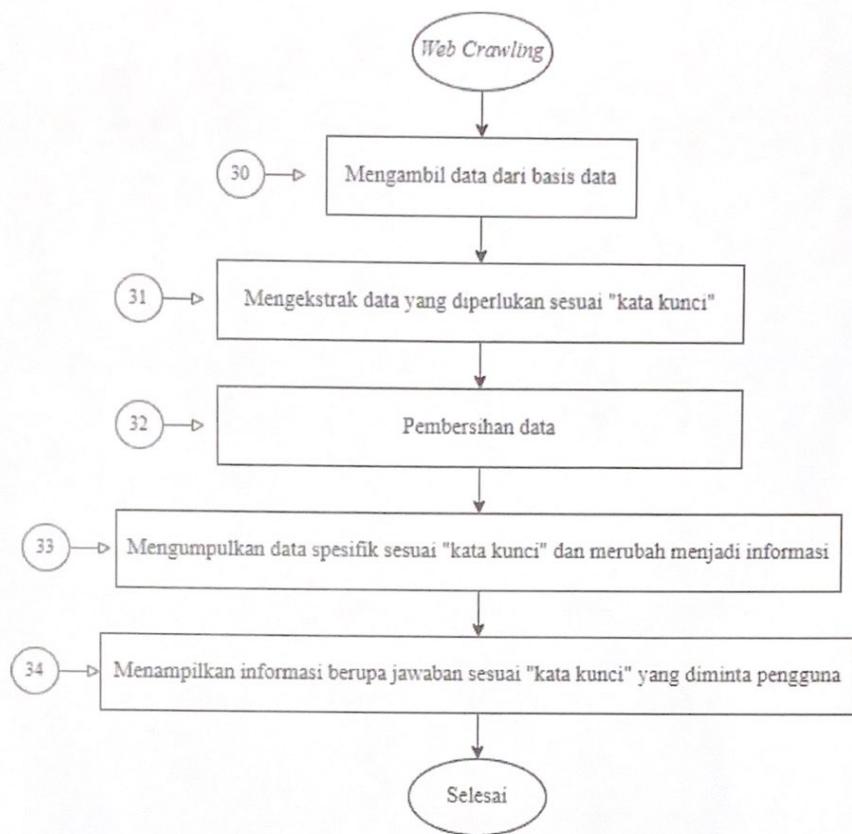
15



Gambar 3



Gambar 4



Gambar 5