

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

2.1.1 Chatbot

Chatbot adalah sebuah program komputer yang memiliki kemampuan untuk menanggapi otomatis pertanyaan pengguna melalui media teks atau audio, biasanya dalam bentuk percakapan singkat (Muhammad et al., 2022).

2.1.2 Openai

Pada Desember 2015, sekelompok tokoh terkemuka, termasuk Sam Altman, Elon Musk, Greg Brockman, Reid Hoffman, Jessica Livingston, dan Peter Thiel, bersatu dalam pembentukan *OpenAI* dengan komitmen lebih dari US\$1 miliar. Mereka bertujuan untuk mengembangkan *AI* yang ramah dan bermanfaat bagi manusia. *OpenAI* berkomitmen untuk kolaborasi terbuka, mengumumkan kantor pusatnya di San Francisco. Pada April 2016, mereka merilis platform pembelajaran penguatan publik, "*OpenAI Gym*", diikuti oleh "*Universe*" pada Desember 2016 untuk melatih kecerdasan umum *AI*. Pada 2018, Elon Musk mengundurkan diri dari dewan direksi karena potensi konflik kepentingan dengan Tesla AI, tetapi tetap sebagai donatur. Pada 2019, *OpenAI* beralih menjadi laba terbatas, bermitra dengan Microsoft dan mengumumkan niat untuk melisensikan teknologinya. Pada 2020, mereka mengumumkan GPT-3, diikuti oleh peluncuran API dan kemampuan multibahasa. Pada 2021, DALL-E diperkenalkan untuk menghasilkan gambar dari teks. Pada Desember

2022, pratinjau *Chatgpt* diluncurkan secara gratis dan mendapat respon besar, dengan proyeksi pendapatan tinggi untuk tahun-tahun mendatang (Amri, 2023).

2.1.2.1 *Chatgpt*

Chatgpt adalah singkatan dari *Generative Pre-trained Transformer*, sebuah model bahasa yang merupakan bagian dari keluarga GPT dari OpenAI. Model ini telah dilatih secara luas menggunakan teknik pembelajaran pengawasan dan penguatan. Dengan antarmuka web, memori, dan fitur yang lebih ramah pengguna, *Chatgpt* memanfaatkan kemampuan dasar GPT-3 dan mengembangkannya untuk berbagai aplikasi (Amri, 2023).

Tujuan utama *Chatgpt* adalah meniru perilaku manusia dalam percakapan alami. *AI* ini mampu belajar dari interaksi dengan berbagai pengguna dan mengembangkan kemampuannya sendiri, termasuk menulis tutorial, kode, musik, dan tugas-tugas lainnya. Penggunaannya sangat beragam, mencakup kreativitas seperti menulis lagu rap, hingga kekhawatiran terhadap keamanan seperti pembuatan kode berbahaya, serta aplikasi bisnis seperti *SEO*, konten pemasaran, dan produktivitas bisnis (Amri, 2023).

2.1.3 *Artificial Intelligence (AI)*

Artificial Intelligence (AI), atau juga dikenal sebagai Kecerdasan Buatan dalam bahasa Indonesia, adalah bidang ilmu komputer yang berfokus pada pembuatan mesin dan sistem yang memiliki kemampuan untuk melakukan fungsi yang biasanya membutuhkan kecerdasan manusia. *AI* menggunakan

model matematika dan algoritma untuk memungkinkan komputer dan sistem lainnya untuk belajar dari data, mengenali pola, dan membuat keputusan yang cerdas (Eriana & Zein, 2023).

Beberapa konsep penting dalam bidang kecerdasan buatan adalah *machine learning* (pembelajaran mesin), *neural networks* (jaringan saraf tiruan), dan *Natural Language Processing* (pemrosesan bahasa alami). Berbagai industri seperti pengenalan suara, pengenalan wajah, mobil otonom, dan pengobatan sangat dipengaruhi oleh kemajuan kecerdasan buatan (Eriana & Zein, 2023).

2.1.3.1 *Natural Language Processing (NLP)*

Natural Language Processing (NLP) yaitu cabang kecerdasan buatan. Bidang ini memungkinkan interaksi antara manusia dan komputer melalui bahasa manusia. Tujuannya adalah untuk membuat komputer dapat memahami, menganalisis, dan menggunakan bahasa manusia dengan cara yang menguntungkan. Proses pengolahan bahasa natural (*NLP*) menangani kompleksitas dan ambiguitas bahasa alami, yang merupakan masalah yang sulit bagi komputer untuk menangani. Berbagai aplikasi *Natural Language Processing (NLP)* termasuk pemrosesan teks, pemahaman bahasa, deteksi sentimen, dan penerjemahan mesin (Eriana & Zein, 2023).

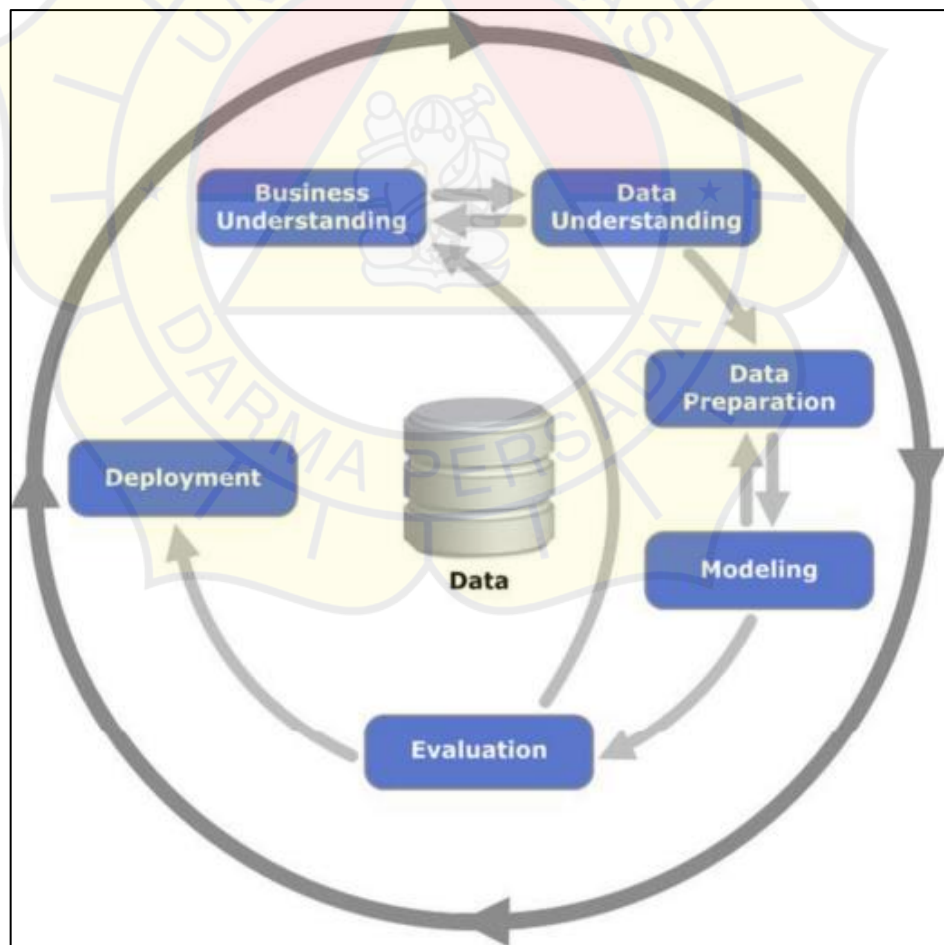
2.1.4 *Data Mining*

Data Mining adalah proses mengekstraksi pengetahuan berharga dari kumpulan data yang besar dan kompleks. Secara teoritis, *Data Mining* adalah

disiplin ilmu yang menggabungkan banyak konsep dan teknik, seperti statistik, ilmu komputer, kecerdasan buatan, dan matematika, untuk mengeksplorasi dan menganalisis data dengan tujuan menemukan pola, hubungan, dan informasi yang tidak terlihat secara langsung (Sulianta, 2024).

2.1.5 Metode CRISP-DM

Metode CRISP-DM (*Cross-Industry Standard Process for Data Mining*) adalah kerangka kerja atau metodologi yang digunakan untuk *Data Mining* dan analisis data. CRISP-DM (Sulianta, 2024), terdiri dari enam tahap utama:



Gambar 2. 1 Fase CRISP-DM

Gambar 2.1 dapat dijelaskan sebagai berikut:

1. Pemahaman Bisnis (*Business Understanding*): Tahap pertama melibatkan pemahaman mendalam terhadap tujuan bisnis proyek, kebutuhan pengguna, serta konteks bisnis secara keseluruhan.
2. Pemahaman Data (*Data Understanding*): Langkah ini melibatkan eksplorasi dan pemahaman terhadap data yang tersedia, termasuk jenis data, kualitasnya, serta hubungannya dengan tujuan bisnis.
3. Persiapan Data (*Data Preparation*): Tahap ini mencakup pra-pemrosesan data, termasuk pembersihan, transformasi, dan integrasi data agar siap untuk analisis lebih lanjut.
4. Pemodelan (*Modeling*): Pada tahap ini, model *Data Mining* dibangun dan dievaluasi untuk memahami pola dan tren dalam data.
5. Evaluasi (*Evaluation*): Langkah ini melibatkan evaluasi model yang dibangun, termasuk pengujian kinerja dan validasi terhadap tujuan bisnis yang ditetapkan.
6. Implementasi (*Deployment*): Tahap terakhir adalah implementasi model dan hasilnya ke dalam lingkungan produksi, serta penyajian temuan dan wawasan kepada pemangku kepentingan.

2.1.6 Pemodelan Sistem UML

UML adalah bahasa visual yang digunakan untuk menentukan, menggambarkan, membangun, dan mendokumentasikan berbagai elemen sistem perangkat lunak. Penggunaan UML membantu pengembang dan

pelanggan lebih memahami perangkat lunak atau produk yang akan dikembangkan. Beberapa jenis diagram UML adalah *Use Case Diagram*, *Activity Diagram*, *Class Diagram*, *Sequence Diagram*, *Collaboration* atau *Communication Diagram*, *UML State Machine Diagram*, *Component Diagram*, dan *Deployment Diagram* (Sundaramoorthy, 2022).

UML bukan hanya sebuah bahasa pemrograman visual; itu juga dapat berinteraksi dengan berbagai bahasa pemrograman lainnya, seperti Visual Basic, JAVA, C++, atau bahkan *basis data object-oriented* (Sumirat et al., 2023).

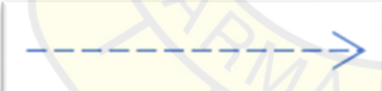
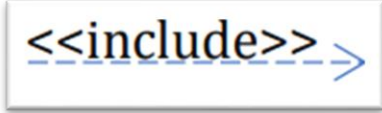
2.1.6.1 Use case Diagram

Menguraikan bagaimana sistem berinteraksi dengan pengguna atau komponen lainnya. Ini sangat berguna untuk menentukan kebutuhan sistem (Sumirat et al., 2023). Komponen dalam *Use Case Diagram* mencakup tiga item utama:

1. Sistem, yang digambarkan dalam bentuk persegi, berfungsi untuk membatasi interaksi *Use Case* dari luar sistem. Meskipun jarang digambarkan dalam diagram karena tidak memberikan makna tambahan, sistem biasanya diberi label yang sesuai.
2. Aktor, meskipun bukan bagian dari sistem, memiliki dampak besar terhadapnya. Aktor memiliki peran penting dalam menggambarkan siapa yang berinteraksi dengan sistem, meskipun mereka tidak memiliki kontrol atasnya dan hanya menerima serta memberikan informasi.
3. *Use Case*, sebagai komponen yang menggambarkan fungsional dalam sistem, membantu pemahaman konsumen dan pengembang terhadap alur

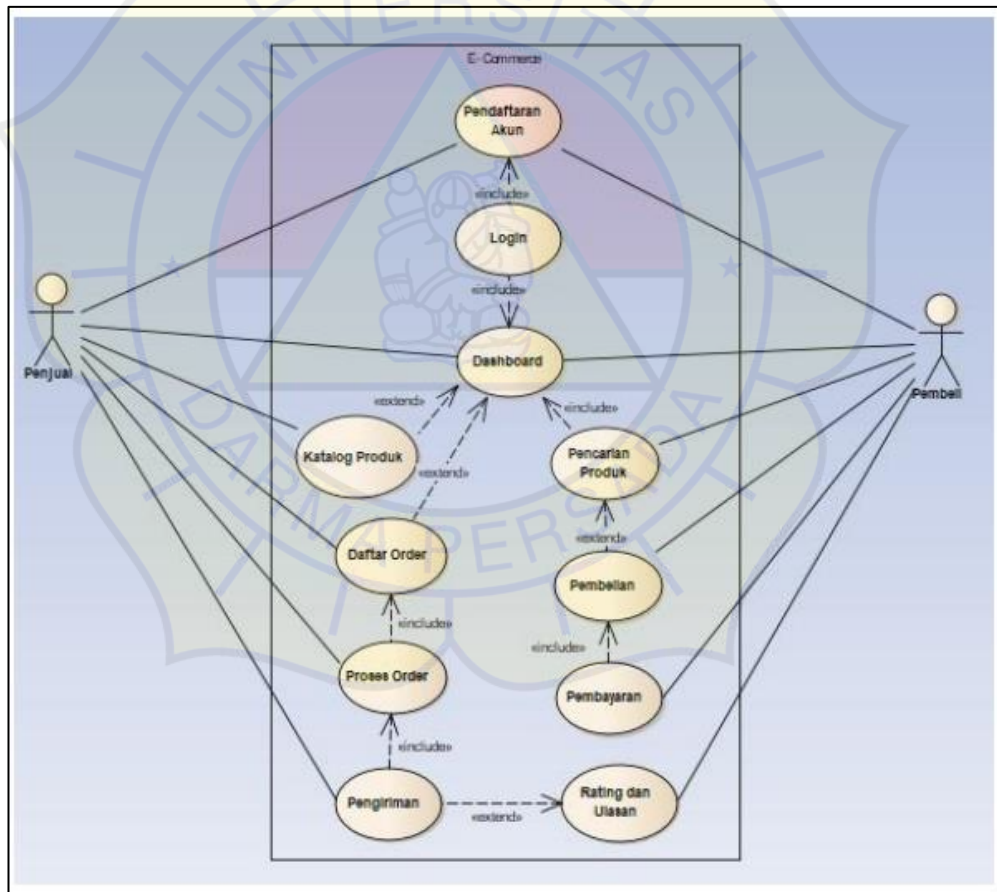
sistem yang akan dibuat. *Use Case* memperjelas bagaimana berbagai fungsi akan berinteraksi di dalam sistem.

Tabel 2. 1 *Komponen Use case*

Simbol	Keterangan
	Aktor: mewakili peran individu, sistem lain, atau alat dalam interaksi dengan <i>Use Case</i> .
	Use case adalah representasi atau hubungan antara sistem dan aktor.
	Association adalah representasi abstrak dari koneksi antara aktor dan <i>Use Case</i> .
	Generalisasi menunjukkan kemampuan aktor untuk berpartisipasi dalam <i>Use Case</i> yang disesuaikan dengan spesialisasi mereka.
	Includes menunjukkan bahwa suatu <i>Use Case</i> sepenuhnya mencakup fungsionalitas dari <i>Use Case</i> lainnya.

Simbol	Keterangan
	Extends menunjukkan bahwa suatu <i>Use Case</i> memberikan tambahan fungsionalitas pada <i>Use Case</i> lainnya jika suatu kondisi tertentu terpenuhi.

Berikut contoh *Use Case Model E-Commerce* yang ditampilkan pada Gambar 2.2 dibawah ini:



Gambar 2. 2 Usecase Model E-Commerce

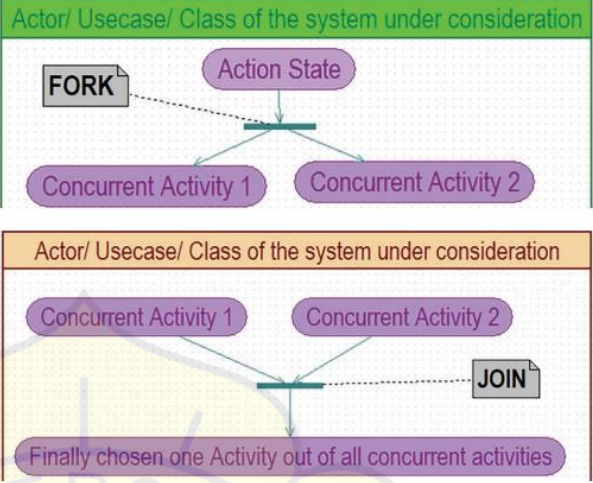
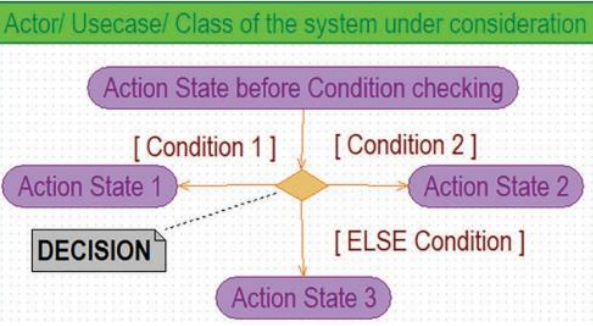
2.1.6.2 Activity Diagram

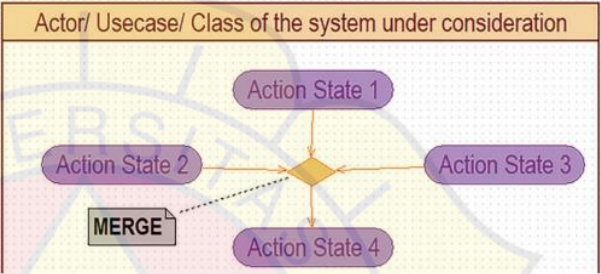
Menggambarkan bagaimana sistem bekerja secara paralel dan sekuensial (Sumirat et al., 2023). Berikut adalah komponen-komponen yang ada pada *Activity Diagram*:

Tabel 2. 2 Komponen *Activity Diagram*

No.	Nama Komponen	Notasi UML dan Tujuan Komponen
1	<i>Initial state</i>	 Ini menggambarkan situasi awal dari sistem yang sedang dipertimbangkan.
2	<i>Final State</i>	 Ini menggambarkan kondisi akhir dari sistem yang sedang dipertimbangkan.
3	<i>Swimlanes</i>	Sebuah swimlane terbagi menjadi dua bagian: bagian atas mewakili entitas seperti aktor, <i>Use Case</i> , atau kelas, sedangkan bagian bawah fokus pada kumpulan aktivitas yang terlibat. Ada dua jenis swimlane: vertikal dan horizontal. <i>Swimlane</i>

No.	Nama Komponen	Notasi UML dan Tujuan Komponen
		vertikal menggambarkan aktivitas paralel dari skenario tertentu dalam sistem yang sedang dipertimbangkan. 
4	Action State	Action State menggambarkan operasi, aktivitas bisnis, atau proses.
5	Object	Entitas yang memindahkan data antara dua status tindakan. 
6	Sinkronisasi	Sinkronisasi adalah representasi dari dua atau lebih aktivitas yang terjadi secara bersamaan atau pada kecepatan yang sama. Sinkronisasi memiliki dua jenis: <i>Fork</i>, yang membagi aliran aktivitas tunggal menjadi dua atau lebih aktivitas secara bersamaan, dan <i>Join</i>, yang menggabungkan dua atau lebih aktivitas bersamaan ke dalam satu aliran dengan memastikan bahwa hanya satu aktivitas berlangsung pada satu waktu.

No.	Nama Komponen	Notasi UML dan Tujuan Komponen
		 Actor/ Usecase/ Class of the system under consideration FORK Action State Concurrent Activity 1 Concurrent Activity 2 Actor/ Usecase/ Class of the system under consideration Concurrent Activity 1 Concurrent Activity 2 JOIN Finally chosen one Activity out of all concurrent activities
7	Decision	Sebuah simpul keputusan memiliki satu <i>input</i> dan dua atau lebih <i>output</i>, tergantung pada kondisi yang diatur. 1. Kondisi yang melekat pada setiap <i>output</i>. 2. Jika suatu kondisi terpenuhi, aliran akan berlanjut bersama dengan <i>output</i> yang sesuai. 3. <i>Output</i> alternatif bisa ditentukan, di mana aliran dapat melanjutkan jika tidak ada kondisi yang terpenuhi.  Actor/ Usecase/ Class of the system under consideration Action State before Condition checking [Condition 1] [Condition 2] Action State 1 Action State 2 DECISION [ELSE Condition] Action State 3

No.	Nama Komponen	Notasi UML dan Tujuan Komponen
8	<i>Merge</i>	Tujuannya adalah untuk menyatukan aliran masukan. Aliran masukan tidak disinkronkan; jika sebuah aliran mencapai simpul tersebut, akan melanjutkan ke <i>output</i> tanpa menunggu aliran lainnya tiba. 
9	<i>Transition</i>	Transisi adalah anak panah yang menunjukkan perpindahan dari satu status aktivitas ke status aktivitas lainnya, dimana perpindahan ini dipicu oleh selesainya aktivitas pada status sebelumnya. 

2.1.7 Software dan Pemograman Terkait

2.1.7.1 Visual Studio Code

Alat pengembangan lintas platform pertama dalam keluarga Microsoft Visual Studio, Visual Studio Code adalah alat gratis dan *open source* yang berjalan di Windows, Linux, dan macOS. Ini adalah alat yang berpusat pada kode yang membuatnya lebih mudah untuk mengedit file kode dan sistem proyek berbasis folder, serta menulis aplikasi web dan seluler lintas platform

melalui platform yang paling populer, seperti Node.js dan .NET Core, desain aplikasi web dan seluler lintas platform, desain aplikasi desktop, desain aplikasi web, desain aplikasi (Del Sole, 2023).

2.1.7.2 NodeJs

Node.js berdiri di atas *engine* javascript yang dirancang untuk melakukan kompilasi dan eksekusi kode javascript, dan berfungsi sebagai runtime javascript yang membaca kode javascript. Google V8 JavaScript *engine* dibuat oleh tim pengembang Google menggunakan bahasa pemrograman C++ (Bakti et al., n.d.).

2.1.7.3 Python

Python, yang pertama kali dirilis pada 1990-an, telah menjadi salah satu bahasa pemrograman paling populer di dunia, digunakan untuk mengembangkan berbagai macam aplikasi, game, dan situs web. Dalam Python, programmer menulis kode menggunakan campuran kata-kata, singkatan, angka, dan simbol, yang memungkinkan instruksi untuk diketik menggunakan keyboard komputer. Dengan pendekatan berbasis teks ini, Python memungkinkan pembacaan, penulisan, dan pemahaman kode yang mudah, menjadikannya bahasa yang sangat digunakan dan disukai dalam pengembangan perangkat lunak (J. T. Santoso, n.d.).

2.1.7.4 Google COLAB

Colaboratory, atau disingkat "Colab," adalah produk dari Google Research yang memungkinkan siapa saja untuk menulis dan menjalankan kode Python di mana saja melalui browser. Produk ini sangat ideal untuk pendidikan, pembelajaran mesin, dan analisis data (Naik et al., 2021).

(Naik et al., 2021) Google Colab, juga dikenal sebagai "kolaborasi", adalah layanan *cloud* gratis yang dibuat oleh Google untuk mendorong penelitian tentang pembelajaran mesin dan kecerdasan buatan, yang seringkali menjadi penghalang untuk belajar dan keberhasilan iklan. Jika Anda adalah seorang ahli dalam pembelajaran mesin dan menghadapi kesulitan untuk melakukan simulasi pada dataset yang sangat besar, Google Colab adalah solusi yang tepat. Google Colab adalah platform yang bagus jika Anda ingin membuat model pembelajaran mesin tetapi tidak memiliki komputer yang kuat.

2.1.7.5 Firebase

Pada tahun 2014, Google menyelesaikan akuisisi perusahaan San Francisco, Firebase, Inc., yang menyediakan solusi pengembangan untuk mengintegrasikan fitur-fitur berbasis *cloud* ke dalam aplikasi seluler dan web. Setelah akuisisi tersebut, Google menggabungkan layanan Firebase dengan fitur-fitur tambahan yang sebelumnya ada di Google *Cloud* Platform, membentuk platform gabungan yang sekarang dikenal sebagai Firebase (Neil Smyth, 2017).

2.1.7.6 HTML

HTML (*Hypertext Markup Language*) adalah bahasa markup yang digunakan untuk membuat dokumen web yang dapat diakses melalui peramban web. WHATWG adalah kelompok orang yang berfokus pada pengembangan teknologi web dan telah mengembangkan HTML versi 5.2, yang memasukkan konsep semantik dan serangkaian API untuk membangun aplikasi web yang kompleks. Karena semua mesin peramban saat ini mendukung standar HTML 5.2, versi terbaru ini dirancang untuk berfungsi dengan baik pada perangkat desktop dan mobile (Bakti et al., n.d.).

2.1.8 Kajian Penelitian terdahulu

Dibawah ini adalah beberapa Penelitian terdahulu:

Tabel 2. 3 Daftar Penelitian Terkait

No	Penulis	Tahun	Judul	Metode	Publikasi
.					
1.	Aji Wahyu Budimana , Agus Setiawanb , Setiya Nugroho	2023	Pengembangan Sistem Layanan Informasi Berbasis Web dengan Memanfaat	<i>metode prototype</i>	Jurnal Teknologi Dan Sistem Informasi Bisnis Vol. 5 No. 4 Oktober 2023.

			kan AI Pada <i>Chatgpt</i>		
Kesimpulan: Mengembangkan sistem layanan informasi berbasis web yang menggunakan <i>Chatgpt</i> dapat membantu pemerintah Desa Tegalrejo memberikan layanan yang lebih efisien dan efektif. <i>Chatgpt</i>, <i>chatbot</i> AI, dapat menjawab pertanyaan pengguna dengan cepat dan akurat. Metode Prototype digunakan dalam pengembangan sistem untuk memenuhi kebutuhan pengguna, dan hasil penelitian menunjukkan bahwa sistem yang dikembangkan memenuhi harapan pengguna dengan baik (Budiman et al., 2023).					
Kelemahan dari penelitian: Terletak pada keterbatasan dalam pengujian sistem, seperti pengujian yang hanya dilakukan dengan metode black box dan waktu, serta penggunaan kuisioner <i>system usability scale</i> (SUS) yang mungkin tidak mencakup semua aspek kegunaan sistem.					
No	Penulis	Tahun	Judul	Metode	Publikasi
.					
2.	Gerald Santoso, Johan	2023	Developme nt of OpenAI	<i>Metode prototipe</i>	G-Tech: Jurnal Teknologi Terapan

	Setiawan, Agus Sulaiman		API-Based <i>Chatbot</i> to Improve <i>User</i> Interaction on the JBMS Website	<i>terstruktur</i>	
Kesimpulan: <i>Chatbot</i> yang terintegrasi dengan OpenAI API untuk situs web jurnal dapat meningkatkan pengalaman pengguna dengan memberikan solusi yang efektif dan tingkat kepuasan yang tinggi. Evaluasi kinerja <i>chatbot</i> dibantu oleh metodologi pengembangan yang terstruktur dan pengujian penerimaan pengguna. Selain itu, integrasi OpenAI API telah terbukti meningkatkan ekstraksi informasi dan rekomendasi artikel yang lebih personal (G. Santoso et al., 2023).					
Kelemahan dari penelitian: Salah satu kelemahan penelitian ini mungkin terletak pada fakta bahwa cakupan pengujian hanya terbatas pada situs web jurnal tertentu; akibatnya, hasil dapat digeneralisasikan hanya dalam konteks tertentu.					
No	Penulis	Tahun	Judul	Metode	Publikasi
.					

2.	Dodi Setiawan , Emilia Ayu Dewi Karuniawati , Saksia Imelda Janty	2023	Peran <i>Chat</i> Gpt (Generative Pre- Training Transforme r) Dalam Implementa si Ditinjau Dari Dataset	<i>Systematic Literature Review</i>	INNOVATIVE : Journal Of Social Science Research
Kesimpulan: Penelitian ini menunjukkan Pentingnya dataset dalam pengembangan model <i>Chatgpt</i> sangatlah vital untuk memastikan kualitas dan keakuratan sistem chat. Dataset yang baik dapat membantu mengurangi kesenjangan dan bias dalam model, sehingga meningkatkan kualitas interaksi antara manusia dan <i>Chatgpt</i>. Oleh karena itu, perlu dilakukan pengawasan dan evaluasi yang cermat terhadap dataset yang digunakan dalam pengembangan <i>Chatgpt</i> untuk meminimalkan risiko terjadinya kesalahan atau distorsi dalam interaksi. (Setiawan et al., 2023).					
Kelemahan dari penelitian:					

	1. Jurnal tidak memberikan informasi yang cukup mendalam mengenai proses pengumpulan, pengolahan, dan validasi dataset yang digunakan dalam pengembangan <i>Chatgpt</i>, sehingga kurangnya transparansi dalam sumber data. 2. Tidak disebutkan apakah dataset yang digunakan telah diuji untuk mengidentifikasi bias atau kesenjangan yang mungkin memengaruhi kualitas model <i>Chatgpt</i>, sehingga potensi masalah dalam representasi data tidak terungkap secara jelas. 3. Tidak ada pembahasan yang mendalam mengenai metode evaluasi yang digunakan untuk mengukur keakuratan dan kualitas model <i>Chatgpt</i> berdasarkan dataset yang digunakan, sehingga kurangnya pemahaman tentang validitas hasil yang diperoleh dari penggunaan dataset tersebut.
--	---