

Rancang Bangun Sistem Penerjemah Multibahasa Daerah Maluku Utara dengan Integrasi GPT Sebagai Pelestarian Budaya Lokal

Anggiah Salim^{1*}, Abdul Haris Muhammad², Sakina Sudin³

^{1,2,3} Universitas Muhammadiyah Maluku Utara
Jl. KH. Ahmad Dahlan No. 100, Kel. Sasa, Kec. Ternate Selatan, Kota Ternate, Indonesia.

e-mail korespondensi: anggiahsalim1@gmail.com

Submit: 20-05-2026 | Revisi: 03-06-2026 | Terima: 09-06-2026 | Terbit online: 20-06-2026

Abstrak - Penelitian ini membahas pengembangan sistem penerjemah multibahasa daerah berbasis web yang mengintegrasikan metode Rule-Based Machine Translation (RBMT) dan GPT untuk menerjemahkan Bahasa Indonesia dengan Bahasa Ternate, Makian Dalam, dan Galela menggunakan 5.572 kosakata berbasis kamus digital, dimana RBMT digunakan untuk pencocokan kata berdasarkan aturan dan GPT digunakan untuk memperbaiki struktur kalimat agar lebih natural sesuai konteks, dengan sistem yang dikembangkan menggunakan HTML, CSS, JavaScript, serta integrasi API GPT melalui Vercel, kemudian diuji menggunakan 468 kalimat yang menghasilkan kategori hasil berupa terjemahan sesuai target, berbeda bentuk namun bermakna sama, tidak sesuai akibat keterbatasan kosakata, serta ketidaktepatan karena ambiguitas makna (satu kata yang memiliki banyak arti), sehingga hasil penelitian menunjukkan bahwa kombinasi RBMT dan GPT mampu meningkatkan kualitas terjemahan meskipun masih dipengaruhi oleh ambiguitas kata, kelengkapan kamus, serta penggunaan tanda baca dan pemisahan kata yang dapat mempengaruhi makna dan akurasi hasil terjemahan.

Kata Kunci : Penerjemah Bahasa Daerah, RBMT, GPT, Multibahasa, Kamus Digital

Abstract - This research discusses the development of a web-based regional multilingual translation system that integrates the Rule-Based Machine Translation (RBMT) method and GPT to translate Indonesian into Ternate, Makian Dalam, and Galela languages using 5,572 vocabulary entries based on a digital dictionary, where RBMT is used for rule-based word matching and GPT is used to improve sentence structure to make the translations more natural and contextually appropriate. The system was developed using HTML, CSS, JavaScript, and GPT API integration through Vercel, and was tested using 468 sentences that produced several categories of results, including translations that matched the target exactly, translations with different sentence forms but the same meaning, inaccurate translations due to limited vocabulary, and inaccuracies caused by semantic ambiguity (a single word having multiple meanings). The results of the study indicate that the combination of RBMT and GPT is capable of improving translation quality, although it is still influenced by word ambiguity, dictionary completeness, and the use of punctuation and word separation, which can affect the meaning and accuracy of the translation results.

Keywords : Regional Language Translator, RBMT, GPT, Multilingual, Digital Dictionary

1. Pendahuluan

Menurut berbagai sumber, di Provinsi Maluku Utara terdapat 36 bahasa daerah yang tersebar di sembilan Kabupaten dan Kota". Pelestarian bahasa daerah adalah amanat Undang-Undang Dasar 1945, Pasal 32 (2) bahwa negara menghormati dan memelihara bahasa daerah sebagai kekayaan budaya nasional; Undang-Undang Nomor 24 Tahun 2009 tentang Bendera, Bahasa, dan Lambang Negara, serta Lagu Kebangsaan; dan Peraturan Daerah Provinsi Maluku Utara Nomor 9 Tahun 2009 tentang Pemeliharaan Bahasa dan Sastra Daerah. Menurut perda tersebut, setiap bahasa daerah harus diajarkan agar tetap terpelihara dan lestari[1].

Bahasa Ternate sebagai salah satu dari 36 bahasa yang dituturkan di propinsi Maluku Utara juga mengalami permasalahan yang terkait pergeseran bahasa yang mengarah pada kematian bahasa. Kenyataan ini terlihat dari semakin merendahnya tingkat penggunaan bahasa ini dalam komunikasi antar penuturnya. Hal ini karena tingkat mobilitas masyarakat yang sangat tinggi terjadi di Ternate. Selain Bahasa Ternate, terdapat bahasa Melayu dialek Ternate yang sangat produktif, yang mana sebagian besar masyarakat tutur di sepuluh kabupaten dan kota yang ada di provinsi Maluku Utara dalam komunikasi sehari-hari menggunakan Bahasa Melayu dialek Ternate [2].



Badan Bahasa terus melakukan upaya pelestarian bahasa yang salah satunya melalui kegiatan penyusunan kamus sebagai kodifikasi bahasa-bahasa yang hampir punah. Dengan menyusun kamus, setidaknya ada produk yang dihasilkan dari konservasi bahasa daerah atau bahasa dialek yang sudah atau akan punah. Kamus sebagai hasil kodifikasi kosakata yang disertai batasan makna kata serta contoh pemakaiannya dalam kalimat memberikan tuntunan bagi pemakainya dalam memahami makna kata serta bagaimana menggunakan kata itu secara tepat di dalam kalimat. Di samping itu, kamus dapat membantu memberikan keluasan wawasan serta pengembangan daya nalar. Dengan demikian, kamus itu dapat dijadikan sebagai sarana untuk mencerdaskan seseorang atau dalam jangkauan yang lebih luas sebagai sarana untuk mencerdaskan bangsa serta menjadi acuan bagi penutur yang akan belajar menggunakan suatu Bahasa[2].

Upaya pelestarian bahasa daerah yang relevan di era digital dapat dilakukan melalui pembangunan website penerjemah multibahasa berbasis kamus digital, dengan memilih perwakilan tiga bahasa daerah dari wilayah Maluku Utara, yaitu Bahasa Ternate, Bahasa Makian Dalam, dan Bahasa Galela. Sistem ini menyajikan kosakata ketiga bahasa tersebut yang dilengkapi dengan gambar, suara, serta padanan makna dalam Bahasa Indonesia, sehingga berfungsi sebagai media edukasi yang interaktif dan mudah diakses oleh masyarakat.

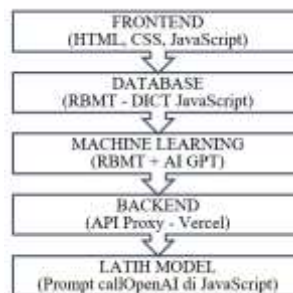
Pengembangan sistem ini menggunakan metode dan algoritma modern yang saling terintegrasi untuk mendukung penerjemahan dan pembelajaran bahasa. Sistem ini memanfaatkan Rule-Based Machine Translation (RBMT) untuk menerjemahkan teks berdasarkan aturan dan pemetaan kamus, Longest Phrase Matching untuk akurasi terjemahan multikata, *Hash Map Lookup* untuk pencarian kosakata cepat, serta GPT-based Post Editing untuk memperbaiki struktur kalimat. Selain itu, sistem juga mendukung *Text-to-Speech* melalui *SpeechSynthesisUtterance()*, *Parsing & JSON Mapping*, *Full-Text Search*, *Prefix Matching* untuk bahasa Indonesia dan multibahasa, serta Lexicographic Sorting untuk menampilkan daftar kata secara abjad. Semua metode ini dibangun berdasarkan kontribusi tokoh dan penelitian terdahulu, seperti Douglas Crockford (JSON & Map), John Larry Kelly Jr. & Louis Gerstman (*Speech Synthesis*), Vaswani et al. & Alec Radford dkk. (GPT), serta Georgetown-IBM Experiment (RBMT).

Pengembangan website dilakukan menggunakan Waterfall Development Model, sementara proses pengumpulan kosakata mengikuti pendekatan kualitatif deskriptif melalui penelusuran arsip bahasa daerah di Kantor Bahasa Maluku Utara, studi literatur dari buku Cerita Anak Maluku Utara dan kamus Bahasa Daerah (Ternate, Makian Dalam, Galela – Indonesia) [3], serta wawancara dengan penutur asli bahasa daerah. Dengan demikian, sistem ini tidak hanya berfungsi sebagai alat penerjemah, tetapi juga sebagai media edukasi adaptif yang dapat diperluas dan digunakan untuk pelestarian bahasa daerah lain di masa mendatang.

2. Metode Penelitian

2.1. Analisis Sistem

Arsitektur sistem dibangun menggunakan aplikasi web dengan pemrosesan utama di sisi frontend dan integrasi API untuk fitur kecerdasan buatan ditampilkan pada Gambar 1..



Gambar 1 Diagram Arsitektur Sistem

Dari Gambar 1, Frontend (HTML, CSS, JavaScript): digunakan sebagai tampilan sistem sekaligus menjalankan proses penerjemahan. Seluruh kode frontend disimpan dan dikelola melalui GitHub., Backend (API Proxy): Backend berupa API proxy yang dijalankan menggunakan Vercel untuk menghubungkan sistem dengan OpenAI serta mengamankan API key. Database (RBMT - DICT JavaScript): Database berupa kamus yang disimpan dalam JavaScript dan digunakan sebagai dasar penerjemahan berbasis RBMT. Machine Learning (RBMT + AI GPT): Sistem menggunakan RBMT untuk menerjemahkan kata dan GPT untuk memperbaiki struktur kalimat agar lebih sesuai konteks. Latih Model (Prompt pada callOpenAI di JavaScript): Proses pembelajaran dilakukan dengan memberikan instruksi (prompt) pada fungsi *callOpenAIcorrect*, dimana model GPT diakses melalui layanan OpenAI API.

2.2. Kebutuhan Sistem dan Perancangan Sistem

Flowchart ini menggambarkan proses interaksi antara pengguna (user), sistem website (*Translator Multibahasa Daerah*), dan pengembang/admin.



Gambar 2 Flowchart Analisis Kebutuhan Sistem

Dari Gambar 2, penjelasan alurnya adalah:

- 1) Pengguna (User) masuk ke sistem website dan memasukkan teks yang ingin diterjemahkan.
- 2) Sistem menerima input dan mengecek mode penerjemahan (aktifkan GPT atau gunakan kamus digital lokal).
- 3) Jika GPT aktif, sistem melakukan *koreksi kalimat* terlebih dahulu agar struktur bahasa lebih alami.
- 4) Jika GPT nonaktif/gagal, sistem langsung menerjemahkan menggunakan kamus digital.
- 5) Hasil terjemahan kemudian ditampilkan, disertai dengan berbagai fitur tambahan:
 - Fitur pilih bahasa (Indonesia ↔ Ternate/Makian Dalam/Galela).
 - Tombol on/off GPT.
 - Fitur Salin, Bersihkan, Audio, ejaan (phonics) dan Kategori Kosakata.
 - Pencarian kata dengan tampilan gambar dan suara.
- 6) Proses berakhir setelah hasil ditampilkan.
- 7) Pengembang (*admin/developer*) melakukan pemeliharaan berkala seperti update kosakata dan pengembangan fitur baru.

Penelitian ini dilakukan di Kota Ternate, Maluku Utara, selama ±8 bulan mulai Oktober hingga Mei yang mencakup tahap perencanaan penelitian dan pengumpulan referensi dari jurnal ilmiah, kamus bahasa daerah, serta cerita rakyat/CEMARA yang bersumber dari Balai Bahasa Provinsi Maluku Utara [4][5][6], tahap pengumpulan data dan validasi bahasa daerah melalui wawancara, observasi, dan komunikasi langsung maupun melalui WhatsApp dan telepon dengan penutur asli Bahasa Ternate, Makian Dalam, dan Galela dari Ternate, Morotai, Desa Tagalaya, dan Pulau Makian Dalam, tahap pengembangan website penerjemah multibahasa berbasis RBMT dan GPT menggunakan HTML, CSS, JavaScript, Bootstrap, GitHub, Vercel, dan GPT API melalui OpenAI Platform, tahap pengujian fungsi sistem dan akurasi terjemahan di Laboratorium Pemrograman Program Studi Teknik Informatika Universitas Muhammadiyah Maluku Utara yang melibatkan mahasiswa Teknik Informatika dan masyarakat pengguna bahasa daerah Maluku Utara, serta tahap evaluasi hasil sistem berdasarkan pengujian dan kuesioner terbuka.

2.3. Metode Pengembangan Sistem

Berikut dua metode sistem yang digunakan dalam pembangunan website:

1) Rule-Based Machine Translation (RBMT)

Rule-Based Machine Translation (RBMT) adalah metode penerjemahan berbasis aturan dan kamus pasangan kata yang dibuat secara manual tanpa menggunakan pelatihan data atau pembelajaran mesin. Pada penelitian ini, RBMT diimplementasikan menggunakan JavaScript dengan mencocokkan input pengguna ke dalam kamus bilingual Bahasa Indonesia–Bahasa Daerah. Sistem hanya menerjemahkan kata atau frasa yang tersedia di dalam kamus, sedangkan kata yang tidak ditemukan akan menampilkan pesan default seperti “Tidak ditemukan”. Contoh sederhana: "saya" → "ngori". Implementasi inti RBMT: `output = dictionary[input] || "Tidak ditemukan";`. Metode ini bekerja berdasarkan aturan dan kosakata lokal yang telah disusun sebelumnya tanpa memahami konteks makna seperti AI. Pendekatan RBMT pertama kali diperkenalkan melalui Georgetown-IBM Machine

Translation Experiment (1954) dan sejalan dengan penelitian Wan Muhammad Hanif et al. (2025) [7] yang menerapkan kamus bilingual dan aturan linguistik untuk penerjemahan bahasa daerah tanpa pembelajaran mesin.

2) GPT-based Post Editing (AI Koreksi Kalimat)

Diterjemahkan oleh kamus kemudian sistem dapat memperbaiki struktur kalimat bahasa Indonesia menggunakan GPT.async function `callOpenAIcorrect (originalText, dictResult, direction){ if(!dictResult) return dictResult;Contoh prompt:role: 'system',content: `Kamu adalah korektor tata bahasa Indonesia. Perbaiki ejaan dan tata bahasa tanpa mengubah makna. Sumber: Model GPT (Generative Pre-trained Transformer) dikembangkan oleh OpenAI, dengan pionir utama seperti Alec Radford, Sam Altman, Ilya Sutskever, dkk. Transformer sendiri pertama diperkenalkan oleh Vaswani et al. (Attention is All You Need) yang menjadi dasar GPT. Metode GPT-based Post-Editing (AI Koreksi Kalimat) pada sistem ini sejalan dengan penelitian yang dilakukan oleh Vikas Raunak.,et al. (2023) dalam jurnal Leveraging GPT-4 for Automatic Translation Post-Editing, yang menunjukkan bahwa GPT-4 efektif digunakan untuk memperbaiki struktur, tata bahasa, dan kelancaran kalimat hasil terjemahan tanpa mengubah makna aslinya, sehingga meningkatkan kualitas dan naturalitas teks terjemahan secara signifikan[8].`

2.4. Algoritma

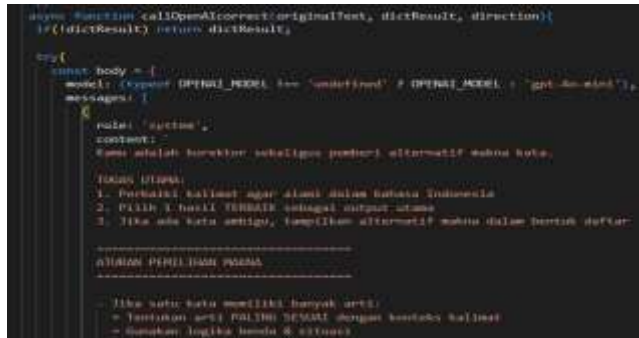
1. Longest Phrase Matching Algorithm: Digunakan untuk mencari pasangan kata/frasa terpanjang agar terjemahan lebih akurat. Program: `for(let len = maxLen; len > 0; len--){`. Konsep: Jika input = "hira sebira" = "saudara" maka sistem memeriksa frasa 2 kata → 1 kata. Metode *Longest Phrase Matching* digunakan untuk meningkatkan akurasi penerjemahan dengan cara mencocokkan frasa terpanjang terlebih dahulu sebelum memecahnya menjadi kata-kata yang lebih kecil. Pendekatan ini memastikan bahwa frasa yang memiliki makna sebagai satu kesatuan, seperti ungkapan atau idiom, dapat dikenali dan diterjemahkan secara utuh, sehingga mengurangi kesalahan terjemahan akibat pemotongan kata secara individual[9].
2. Hash Map Lookup (O(1) Search): Hash Map Lookup adalah metode pencarian data berbasis pasangan key-value, di mana data dapat diakses langsung melalui kunci (key) tanpa harus menelusuri seluruh data satu per satu, sehingga waktu pencariannya sangat cepat (kompleksitas rata-rata O(1)). Contoh sederhana: Key = "saya" → Value = "ngori". Dalam konteks website kamus digital: Key = kata atau frasa bahasa indonesia dan Value = arti kata, gambar, audio, ejaan, atau terjemahan bahasa daerah.[10].
3. Speech Synthesis (Text-to-Speech): Fitur audio pada sistem menggunakan *SpeechSynthesis API* untuk membacakan kosakata dan hasil terjemahan secara otomatis, misalnya kata "ngori" dapat langsung diperdengarkan kepada pengguna. Pendekatan ini membantu pelafalan bahasa daerah dan sejalan dengan penelitian Tacotron oleh Yuxuan Wang et al. (2017) dapat dibacakan dengan fitur *text-to-speech*[11]. Program: `let utter = new SpeechSynthesisUtterance(word); speechSynthesis.speak(utter);`
4. Parsing & JSON Mapping: Metode ini digunakan untuk memetakan dan mengelompokkan data kosakata per kategori menggunakan struktur JSON agar data kamus dapat diproses dan ditampilkan secara responsif pada website.[12].
Contoh program: `const DICT= {"Kata Benda":{"Keluarga":[{"ind":"ibu","ter":"yaya","makian":"ma","galela":"awa","img":"http"}],`
5. Full-Text Search (Substring Matching): Metode ini digunakan untuk melakukan pencarian kosakata berdasarkan kemunculan kata di seluruh isi data, tanpa memperhatikan posisi awal kata. Implementasi dilakukan menggunakan fungsi *includes()* pada JavaScript, sehingga sistem dapat menemukan kosakata meskipun kata kunci berada di tengah atau akhir teks. Contohnya pencarian "yah" tetap dapat menemukan kata "ayah"[13].
6. Prefix Matching Algorithm (Bahasa Indonesia): Metode ini digunakan untuk melakukan pencarian kosakata berdasarkan awalan kata (prefix) dalam Bahasa Indonesia Lebih cepat dan terstruktur dibanding full-text search serta mengurangi hasil yang tidak relevan. Pengguna cukup mengetik beberapa huruf awal, lalu sistem akan menampilkan kosakata yang dimulai dengan awalan tersebut. Misal "mak"="makan", "makanan". Metode ini mendukung fitur *autocomplete* dan pencarian terarah[14].
7. Multilingual Prefix Search (Bahasa Daerah): Metode ini merupakan pengembangan dari prefix matching yang diterapkan pada bahasa daerah, dengan menyesuaikan pilihan bahasa yang dipilih pengguna. Metode ini memungkinkan pencarian kosakata lokal berdasarkan awalan kata dalam berbagai bahasa daerah (misalnya Ternate, Makian Dalam, Galela). Contoh kode: `matchPrefixDaerah(it, prefixLoc, chosen);`[14].
8. Lexicographic (Alphabetical) Sorting: Metode ini digunakan untuk mengurutkan daftar kosakata berdasarkan urutan alfabet agar pengguna lebih mudah menemukan kata dalam kamus digital. Contoh Implementasi Program (JavaScript): `// Data kata kamus: const words = ["ibu","ayah","nenek","kakek","adik","kakak"]; // Sorting A-Z (lexicographic):const sortedWords = words.sort((a, b) =>a.localeCompare(b, 'id', { sensitivity: 'base' }));console.log(sortedWords);` Hasil output: `["adik", "ayah", "ibu", "kakak", "kakek", "nenek"]`
9. Phonics-Based Spelling dengan Text-to-Speech (TTS): Metode ini menerapkan pendekatan fonik (*phonics*), yaitu metode pembelajaran membaca yang menekankan hubungan antara huruf (*grafem*) dan bunyi (*fonem*) melalui proses ejaan huruf demi huruf. Dalam metode ini, sebuah kata dipecah menjadi unit huruf tunggal yang kemudian dieja dan dibacakan secara berurutan sebelum akhirnya dibacakan sebagai satu kata utuh.

Pendekatan fonik bertujuan untuk membantu pembelajar awal, khususnya anak usia dini, dalam memahami struktur bunyi bahasa serta meningkatkan kemampuan membaca dasar secara bertahap. Dalam konteks aplikasi kamus digital berbasis web, metode fonik ini diimplementasikan dengan menggabungkan visualisasi huruf dan audio *Text-to-Speech (TTS)*. Setiap huruf ditampilkan secara individual dan dibacakan satu per satu dengan penekanan visual (highlight), sehingga pengguna dapat mengasosiasikan bentuk huruf dengan bunyinya. Setelah seluruh huruf selesai dieja, sistem membacakan kata secara utuh untuk memperkuat proses blending bunyi[15]. Berikut contoh penerapan metode fonik pada satu kata: const word = "makan"; const letters = word.split(""); letters.forEach(letter => {console.log(letter);}); Hasil ejaan fonik: m – a – k – a – n → makan.

2.5. Metode Pengujian

Metode pengujian dilakukan dengan menerapkan pola perintah (prompt) pada fungsi callOpenAI yang dipisahkan menjadi dua halaman, yaitu Bahasa Indonesia → Bahasa Daerah dan Bahasa Daerah → Bahasa Indonesia. Alasan Pemisahan Dua Sistem AI karena berdasarkan hasil pengujian penggunaan satu sistem AI untuk dua arah terjemahan menyebabkan perintah menjadi terlalu panjang, AI tidak mengikuti aturan yang diberikan dan hasil terjemahan menjadi tidak konsisten. Sehingga sistem dipisahkan menjadi dua halaman dan dua callOpenAI agar Instruksi lebih focus. Berdasarkan hasil analisis, ketiga bahasa daerah memiliki pola yang lebih fleksibel dibanding Bahasa Indonesia yang umumnya menggunakan struktur Subjek–Predikat–Objek (SPO). Perbedaan utama meliputi posisi angka yang berada di belakang benda, negasi yang berpindah ke belakang kata kerja, pembalikan beberapa frasa, serta struktur kalimat yang lebih bergantung pada konteks. Oleh karena itu, sistem menerapkan dua tahap proses, yaitu RBMT untuk menerjemahkan kata berdasarkan kamus, kemudian GPT untuk menyusun ulang struktur kalimat agar lebih natural. Alur Sistem: Input → hasil kamus (RBMT) → AI memperbaiki struktur kalimat → OUTPUT.

“Pola Bahasa Indonesia → Bahasa Daerah difokuskan pada penyusunan ulang struktur kalimat berdasarkan aturan posisi kata seperti pembalikan angka dan benda (“2 rumah” → “rumah 2”), perubahan posisi negasi (“tidak makan nasi” → “nasi makan tidak”), pembalikan frasa dua kata (“rumah saya” → “saya rumah”), perubahan keterangan tempat (“di belakang rumah” → “di rumah belakang”), serta penyesuaian pola kalimat tanya (“siapa kamu” → “kamu siapa”) agar sesuai dengan struktur bahasa daerah, sehingga aturan tersebut diterapkan pada fungsi callOpenAIcorrect() melalui prompt AI untuk memperbaiki hasil terjemahan kamus menjadi kalimat yang lebih natural dan kontekstual.”



```
async function callOpenAIcorrect(originalText, dictResult, direction) {
  if(!dictResult) return dictResult;

  try {
    const body = {
      model: (direction === 'forward' ? OPENAI_MODEL : 'gpt-4o-mini'),
      messages: [
        {
          role: 'system',
          content: `
            Kamu adalah korrektor sekaligus pembuat alternatif makna kata.

            TUGAS UTAMA:
            1. Perbaiki kalimat agar alami dalam bahasa Indonesia
            2. Pilih 1 hasil TERBAIK sebagai output utama
            3. Jika ada kata ambigu, tambahkan alternatif makna dalam bentuk daftar

            ATURAN PEMERIKSAAN:
            - Jika kata/kata memiliki banyak arti:
              - Tampilkan arti PALING SESUAI dengan konteks kalimat
              - Gunakan logika benda & struktur
          `
        }
      ]
    };
  } catch (error) {
    console.error('Error in callOpenAIcorrect:', error);
  }
}
```

Gambar 3 Pemanggilan fungsi callOpenAI

Pola Bahasa Daerah → Bahasa Indonesia, Sedangkan pada pola Bahasa Daerah → Bahasa Indonesia, AI lebih difokuskan pada pemilihan makna yang sesuai konteks, menghasilkan kalimat yang alami, menangani ambiguitas makna kata serta mengabaikan imbuhan yang tidak dikenali dan fokus pada kata dasar. HASIL:<kalimat terbaik> ALTERNATIF:- <kalimat alternatif 1>, - <kalimat alternatif 2>- dst (urutkan dari paling masuk akal). Berdasarkan pola tersebut, maka disusun perintah pada fungsi callOpenAI sebagai berikut:

Dalam proses pengujian ditemukan bahwa setiap bahasa memiliki imbuhan dengan fungsi yang berbeda dan dapat melekat langsung pada kata dasar. Pada Bahasa Makian, imbuhan seperti “k” digunakan sebagai penanda orang pertama (saya) dan “n” sebagai penanda orang ketiga (dia), sehingga kata dasar attiam (pusing) dapat berubah menjadi kattiam (saya pusing) dan nattiam (dia pusing). Pada Bahasa Ternate, imbuhan seperti “o” dan “i” digunakan untuk menunjukkan waktu atau aspek, misalnya o pake (akan mengenakan) dan i pake (sedang/terjadi mengenakan). Sedangkan pada Bahasa Galela, terdapat imbuhan seperti “o”, “ma”, “i”, dan lainnya yang dapat membentuk kata kerja atau penanda makna, dimana “o” juga dapat muncul di awal kata sebagai penanda tertentu dalam struktur kalimat. Selain itu, terdapat bentuk seperti “se”, “lo”, dan “de” yang dalam penggunaan sehari-hari dapat berfungsi sebagai penghubung seperti “dan” atau “dengan”, namun dalam konteks tertentu seperti bilangan, bentuk tersebut menjadi satu kesatuan dan tidak dapat dipisahkan. Karena perbedaan imbuhan yang cukup kompleks pada ketiga bahasa, sistem menerapkan pendekatan penyederhanaan dengan menghilangkan imbuhan dan memfokuskan pada kata dasar agar mudah dikenali oleh sistem, menginstruksikan AI untuk memahami makna berdasarkan kata dasar yang tersedia, Jika imbuhan merupakan satu kesatuan makna dalam kamus, maka diproses sebagai satu unit dan jika tidak ditemukan, maka kata dikembalikan ke bentuk dasar

untuk dicocokkan dengan kamus. Contoh: “mengikuti” → dicari sebagai “ikut” atau “diikuti” → dicari sebagai “ikut”. Pendekatan ini membantu sistem dalam meningkatkan akurasi pencarian kosakata dan mengurangi kesalahan terjemahan akibat perbedaan bentuk kata.

3. Hasil dan Pembahasan

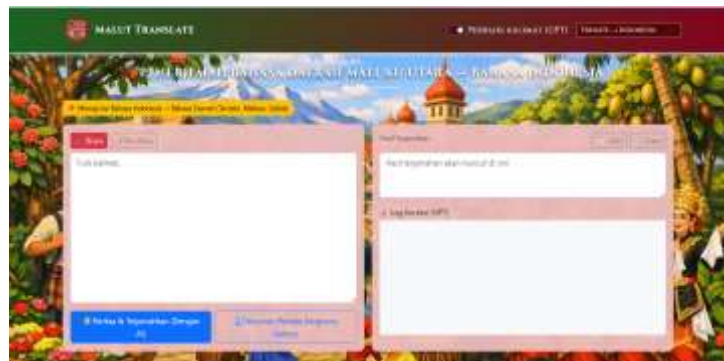
3.1. Tampilan Antarmuka

Antarmuka sistem penerjemah multibahasa daerah Maluku Utara dirancang sederhana, responsif, dan mudah digunakan baik oleh admin maupun pengguna umum. Tampilan sistem terdiri dari halaman admin untuk pengelolaan database dan halaman user untuk proses penerjemahan.



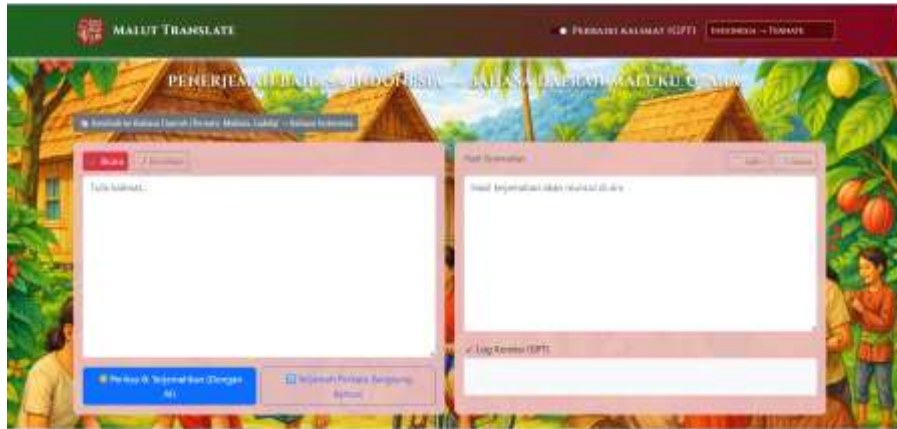
Gambar 4 Tampilan Antarmuka Halaman Admin

Gambar 4, Halaman Admin: Halaman admin digunakan untuk mengelola database kosakata dan source code sistem yang tersimpan pada GitHub serta deployment website melalui Vercel. Admin harus memiliki akun email yang terhubung dengan akun GitHub dan Vercel agar dapat melakukan pengeditan database, perbaikan source code, dan upload ulang sistem. Pada halaman GitHub terlihat file utama seperti HTML, CSS, JavaScript, API proxy, serta database kamus yang tersimpan pada file script.js dengan ribuan baris data kosakata yang digunakan dalam sistem penerjemahan. Halaman ini hanya dapat diakses oleh admin yang memiliki hak akses terhadap repository sistem.



Gambar 5. Tampilan Antarmuka Halaman User

Gambar 5, Halaman User: Halaman user merupakan halaman utama website penerjemah yang dapat diakses secara umum melalui domain bahasamalut-rho.vercel.app. Pada halaman ini pengguna dapat memilih arah bahasa, memasukkan kalimat, menjalankan terjemahan menggunakan AI maupun kamus langsung, melihat hasil terjemahan dan log koreksi GPT, serta menggunakan fitur tambahan seperti suara, salin, dan input bicara. Tampilan dirancang agar pengguna dapat menggunakan sistem secara langsung tanpa perlu login atau konfigurasi tambahan.



Gambar 6. Halaman Utama Input dan Output

Komponen Utama Halaman Input dan Output pada Gambar 6 terdiri dari input kalimat berisi kotak teks untuk menulis kalimat dalam bahasa daerah (misalnya Ternate, Makian, Galela), output terjemahan berisi kotak hasil terjemahan ke Bahasa Indonesia. Log koreksi: Area catatan koreksi otomatis (GPT), tombol periksa dan terjemahkan: Tombol untuk menjalankan terjemahan dengan AI, tombol terjemah per kata: tombol untuk terjemahan langsung per kata dari kamus. Dropdown bahasa: pilihan arah terjemahan: Ternate → Indonesia, Makian → Indonesia, Galela → Indonesia.



Gambar 7 Halaman Utama Kamus Per Kategori

Komponen Utama Halaman Kamus Per Kategori Gambar 7, Halaman ini menyediakan fitur pemilihan bahasa daerah, kategori, dan tema kosakata, dilengkapi kolom pencarian kata dalam Bahasa Indonesia maupun bahasa daerah, ringkasan statistik kosakata, tombol untuk melihat data duplikat, serta kartu kosakata yang menampilkan gambar, terjemahan, dan fitur ejaan kata.

Komponen Utama Halaman Tabel Daftar Semua Kata, Halaman ini menyediakan pilihan kategori kata, tabel yang menampilkan jenis, tema, serta terjemahan kata dalam Bahasa Indonesia, Ternate, Makian, dan Galela, dilengkapi fitur gulir vertikal, tombol kembali ke bagian atas tabel, dan navigasi untuk berpindah antarhalaman data.

3.2. Pengujian Sistem

Hasil uji fungsional (*Black Box Testing*) ditunjukkan pada tabel 1.

Tabel 1 Uji Fungsional (*Black Box Testing*)

No	Fitur yang Diuji	Deskripsi Uji	Hasil yang Diharapkan	Hasil Aktual	Status
1	Penerjemahan Bahasa Daerah → Indonesia	Pengguna memasukkan kalimat bahasa daerah	Sistem menampilkan hasil terjemahan Bahasa Indonesia	Sesuai	OK
2	Penerjemahan Indonesia → Bahasa Daerah	Pengguna memasukkan kalimat Bahasa Indonesia	Sistem menampilkan hasil terjemahan bahasa daerah	Sesuai	OK
3	Koreksi Kalimat AI (GPT)	AI memperbaiki struktur kalimat	Kalimat lebih natural dan sesuai konteks	Sesuai	OK
4	Tombol Terjemah Langsung Kamus	Terjemahan tanpa AI	Sistem menerjemahkan berdasarkan kamus RBMT	Sesuai	OK
5	Log Koreksi GPT	Menampilkan hasil utama dan alternatif	Alternatif kalimat tampil sesuai konteks	Sesuai	OK

No	Fitur yang Diuji	Deskripsi Uji	Hasil yang Diharapkan	Hasil Aktual	Status
6	Fitur Pilih Bahasa	Mengubah arah bahasa terjemahan	Arah bahasa berubah sesuai pilihan	Sesuai	OK
7	Fitur Bicara dan Suara	Input suara dan pembacaan hasil	Sistem dapat membaca dan menerima suara	Sesuai	OK
8	Fitur Salin dan Bersihkan	Menyalin dan menghapus teks	Teks berhasil disalin dan dibersihkan	Sesuai	OK
9	Penanganan Angka	Input angka dalam kalimat	Sistem membaca angka sesuai aturan	Sesuai sebagian	OK
10	Halaman Kosakata	Menampilkan data kategori kata	Data tampil sesuai kategori dan tema	Sesuai	OK
11	Fitur Eja	Membaca huruf per huruf	Huruf dibacakan sesuai teks	Sesuai	OK
12	Akses Domain Website	Website diakses beberapa pengguna	Sistem tetap berjalan tanpa gangguan utama	Sesuai	OK

Analisis Uji Fungsional yang ditampilkan pada Tabel 2 menunjukkan seluruh fitur utama seperti penerjemahan RBMT dan GPT, suara, ejaan, pencarian kata, serta pengelolaan angka dapat berjalan dengan baik. Website juga tetap stabil saat diakses beberapa pengguna melalui domain gratis Vercel pada *bahasamalut-rho.vercel.app*.

Tabel 2 Kategori Penilaian

Kategori Penilaian	Keterangan	Nilai
Persis	Kalimat sama 100% dengan target (termasuk pada alternatif)	1
Bentuk Berbeda tapi Makna Sama	Susunan berbeda namun memiliki arti yang sama (termasuk pada alternatif)	2
Bentuk dan Makna Berbeda	Hasil tidak sesuai atau salah arti (termasuk pada alternatif) - Karena kosa kata yang belum di masukkan atau adanya kata yang memiliki banyak arti sehingga AI keliru	3

Uji Coba Pengguna (*Usability Test*), Uji coba pengguna dilakukan untuk mengevaluasi hasil penerjemahan sistem berdasarkan kesesuaian makna dan struktur kalimat pada tiga bahasa daerah yaitu Ternate, Makian, dan Galela. Pengujian dilakukan menggunakan kalimat uji yang dibandingkan dengan target terjemahan asli dari kamus, CEMARA, dan penutur Bahasa. Penilaian kesesuaian hasil terjemahan dilakukan dengan kategori yang ditampilkan pada Tabel 3.

Tabel 3 Uji Coba Pengguna (*Usability Test*)

No	Usename	Arah Bahasa	Kalimat Uji	Target	Hasil Sistem AI	Kesesuaian	Hasil Kamus
1	Responden 1	Indonesia → Ternate	Saya melihat 21 ikan merah di sungai	ngori hida nyao roriha nyagi romdidi se rimoi toma ake lamo	ngori hida nyao roriha nyagi romdidi toma ake lamo	1	ngori hida nyagi romdidi se rimoi nyao roriha toma ake lamo
2	Responden 2	Ternate → Indonesia	Ana fodi namo kokotu kara piara toma fala	Mereka membeli ayam hitam untuk dipelihara di rumah.	Alternatif: Mereka membeli ayam hitam untuk dipelihara di rumah.	1	Mereka membeli ayam hitam usik asuh kepada rumah
3	Responden 3	Indonesia → Makian	mereka mengumpulkan cangkang kerang yang berserakan	n lutin suat ni liko yang tabalai	n lutin suat ni liko yang tabalai	1	n lutin suat ni liko yang tabalai
4	Responden 4	Indonesia → Makian	wom umli la bersih ni liko	Sampai di rumah, cangkangnya dibersihkan.	Sampai di rumah, cangkangnya dibersihkan.	1	sampai rumah la dibersihkan cangkangnya
5	Responden 5	Indonesia → Galela	Kami tidak bisa pergi karena jalan rusak	to ngomi tagi dadi hiwa sababu ngeko biau	to ngomi tagi dadi hiwa sababu ngeko biau	1	to ngomi dadi wa tagi sababu ngeko biau
6	Responden 6	Galela → Indonesia	o boki mangopa dha tataro ngaimoi I mote-mote ngohi	Seekor anak kucing hitam mengikuti aku.	Seekor anak kucing hitam mengikuti aku.	1	anak kucing hitam seekor mengikuti aku
7	Responden 7	Indonesia → Makian	adik duduk di kursi paling depan	thani nbatalon kurusi soda li	thano batalon ni kurusi sodali paling sodali	2	thano batalon ni kurusi paling sodali
8	Responden 8	Ternate → Indonesia	toma ternate bank dofu ma	diternate sudah banyak bank	Kepada Bank Ternate, banyak yang datang."	3	kepada ternate bank banyak ma
9	Responden 9	Ternate → Indonesia	ternate bank dofu ma	ternate sudah banyak bank	Banyak bank di Ternate.	2	ternate bank banyak ma

Analisis Usability Test, Berdasarkan hasil uji coba, sistem mampu menghasilkan terjemahan yang sesuai ketika kosakata tersedia lengkap dalam kamus dan konteks kalimat dapat dikenali AI. Pada beberapa kasus, hasil terjemahan memiliki bentuk berbeda tetapi tetap memiliki makna yang sama karena AI melakukan penyusunan ulang kalimat agar lebih natural. Namun, beberapa kesalahan masih terjadi akibat kosakata yang belum tersedia atau ambiguitas makna pada kata tertentu. Secara keseluruhan, kombinasi RBMT dan GPT mampu menghasilkan terjemahan yang lebih fleksibel dibanding kamus biasa karena mempertimbangkan struktur dan konteks kalimat.

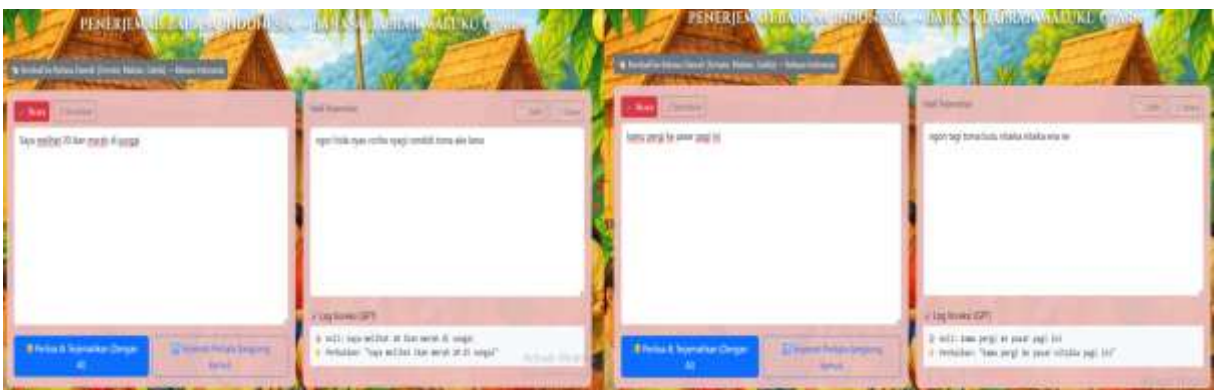
3.3. Pembahasan

Pada Pengujian Penerjemah Bahasa Daerah – Bahasa Indonesia pada Gambar 8 terlihat bahwa sistem mampu menangani kata yang memiliki lebih dari satu arti melalui bantuan konteks kalimat. Pada data kamus ditemukan bahwa satu kata dapat memiliki beberapa makna, seperti kata “tego” yang dapat diartikan sebagai “huni”, “tinggal”, atau “duduk”. Pada pengujian pertama, ketika diberikan input dengan konteks “tego kursi”, sistem cenderung menghasilkan makna “duduk” karena kata “kursi” mengarah pada aktivitas duduk. Namun, pada beberapa alternatif masih terdapat kemungkinan makna lain seperti “tinggal di kursi” yang secara struktur masih terbentuk tetapi kurang tepat secara konteks. Sedangkan, pengujian kedua, ketika diberikan kalimat yang lebih lengkap seperti “fajaru tego toma ma fala fira manyira” (dengan “fala” berarti rumah), sistem secara konsisten memilih makna “tinggal” karena konteks kalimat mengarah pada tempat tinggal. Hal ini menunjukkan bahwa semakin lengkap konteks kalimat yang diberikan, maka semakin tepat pula AI dalam menentukan makna kata yang ambigu. Dengan demikian, sistem tidak hanya menerjemahkan berdasarkan kamus, tetapi juga memanfaatkan konteks kalimat untuk memilih arti yang paling sesuai, meskipun dalam beberapa kondisi masih dapat menghasilkan alternatif yang kurang tepat akibat ambiguitas makna kata.



Gambar 8. Pengujian (Kelengkapan Konteks Kalimat yang Diberikan)

Sedangkan pada proses penerjemahan Bahasa Indonesia ke Bahasa Daerah yang ditampilkan pada Gambar 9, sistem menerapkan aturan khusus agar struktur kalimat Bahasa Indonesia menyesuaikan dengan pola Bahasa Daerah. Hal ini dilakukan karena setiap bahasa daerah memiliki susunan kalimat yang berbeda, sehingga diperlukan penyesuaian posisi kata melalui bantuan AI.



Gambar 9. Pengulangan Output Kata

Hasil pengujian menunjukkan bahwa sistem mampu menghasilkan terjemahan yang sesuai dengan aturan yang telah ditentukan, khususnya pada penempatan angka di belakang benda, seperti pada kalimat “20 ikan merah” yang diubah menjadi “ikan merah 20”. Hal ini menandakan bahwa aturan RBMT yang dikombinasikan dengan AI

berjalan dengan baik dalam menyusun ulang struktur kalimat sesuai pola bahasa daerah. Namun, berdasarkan hasil pengujian lain, ditemukan bahwa ketika aturan yang diberikan terlalu banyak, sistem terkadang menghasilkan kalimat yang berulang atau tidak stabil, seperti munculnya pengulangan kata. Contohnya pada kata “pagi” yang dalam kamus diterjemahkan menjadi “nitaika”, namun pada beberapa hasil terjemahan dapat muncul dua kali menjadi “nitaika nitaika”, meskipun input hanya satu kata. Untuk mengatasi hal tersebut, solusinya dengan menggunakan tombol “terjemah langsung kamus” terlebih dahulu, kemudian dilanjutkan dengan proses AI. Pendekatan ini membantu menstabilkan hasil karena AI akan memproses ulang berdasarkan hasil kamus yang sudah lebih terstruktur. Kondisi ini terjadi karena alur sistem bekerja secara bertahap yaitu input → hasil kamus (RBMT) → perbaikan struktur oleh AI → output, sehingga jika hasil awal dari kamus sudah jelas, maka AI dapat menghasilkan kalimat yang lebih sesuai dengan aturan yang telah ditentukan.

Pengujian sistem juga dilakukan dengan menguji kecocokan sebanyak 468 kalimat pada tiga bahasa daerah yaitu Ternate, Makian, dan Galela, dimana data kalimat tersebut berasal dari kamus, CEMARA, dan penutur asli. Pada Gambar 8 hasil pengujian menunjukkan bahwa sistem mampu menghasilkan kalimat yang sama persis dengan target, menghasilkan kalimat yang berbeda bentuk namun masih memiliki makna yang sama, serta menghasilkan kalimat yang tidak sesuai ketika kosakata belum tersedia dalam kamus atau ketika AI tidak memilih makna yang tepat sesuai konteks kalimat. Selain itu, hasil juga menunjukkan bahwa sistem dapat menampilkan alternatif kalimat yang memiliki kesamaan makna, baik dalam bentuk yang sama maupun berbeda, serta tidak menampilkan hasil yang sesuai apabila kosakata yang digunakan belum dimasukkan ke dalam sistem atau terjadi ketidaktepatan dalam pemilihan makna oleh AI.

A	B	C
1 DATA LATIH KALIMAT		
2		
3 Indonesia	Ternate	Hasil sistem bahasa ternate - Indonesia
4 satu abad	taang ratu mai	satu abad
5 dua abad	taang ratu rambidi	dua abad
6 dia mengatakan perintah guru	ura u master ua guru na parenta	dia seorang laki-laki yang mengatakan perintah
7 sebelum sakit kita harus membersihkan diri	kala san aaba haru sigifi nadri	jika melakukan kesalahan, harus bersihkan diri
8 rakyat mengabdikan pada sultan	kala ge co'oa se kolono	titik-titik pengabdian berilah sultan
9 sakit batuk abnormal	kala igogota nyahi ua	sakit mengalami sakit yang tidak normal
10 guru itu selalu abun	guru ura ge selalu urama ua	Guru laki-laki itu selalu abun
11 jalan baru itu banyak abu	ngoko sangi ge defu fira	Jalan baru itu banyak abu
12 rumah itu cukup abu-abu	fala ge ma vet fira fira	Rumah itu abu-abu
13 main domino meski agak acak hasilnya	tomu dorono ge masi sirawaha mamali	mainan meski acak hasilnya
14 laki-laki meski pake ada	gafaha aar masi pake ada	meski acak harus pakai ada
15 siapa yang mengantar anak itu	ngapa ana pang laro ana na	siapa yang mengantar anak itu?
16 anak itu di angkat guru tapi dia tidak acak	ngapa ngapa guru kokoro tapi abdu'ak itu ananggi guru, tetapi tidak mengacaknya	anak itu diangkat guru, tetapi tidak mengacaknya
17 siapa yang tahu jawab angkat tangan	ngapa ana pang soro sango tede ga	siapa yang mengetahui apa mengangkat tangan?
18 tidak bereslah	kama adaha ua	tidak bereslah
19 kadang kadang ada kadang kadang tidak	walisyo'ko sama walisyo'ko maio	Kadang ada, kadang tidak
20 ada rumah	sama fira	ada rumah
21 ada sang	sama pipi	ada sang
22 dalam rumah itu adam sekal	tomu fala ge madaha ala ala faki	di rumah, sangat sempit dan sempit sekali
23 ada kandang	ngogru madhuta	ada kandang
24 ada itu	ngogru fira	ada itu
25 jua fedi adalah adiral dawa' serapan ternate	jua fedi ge kapita bu bolato kanga	Tuan Fandi adalah pemimpin kerajaan Ternate
26 rumah itu alamatnya salt	fala ge ma alamat usah	Rumah itu adalah alamat yang salt
27 abu apam	atobelo namu	abu apam
28 jangan abu donda	afa si maio dudu	jangan talang seawali

Gambar 10. Pengujian Kecocokan Kalimat

Dari pengujian yang telah dilakukan menunjukkan bahwa hasil sistem terbagi menjadi empat kemungkinan utama yaitu hasil yang sesuai dengan target, hasil dengan bentuk berbeda namun memiliki makna yang sama, hasil yang tidak sesuai karena kosakata belum tersedia dalam kamus, serta hasil yang tidak tepat akibat AI tidak memilih makna yang paling sesuai dengan konteks kalimat.

4. Kesimpulan

Berdasarkan hasil penelitian dan implementasi sistem penerjemah bahasa daerah berbasis web dengan pendekatan Rule-Based Machine Translation (RBMT) dan model GPT, maka dapat disimpulkan sistem penerjemah bahasa daerah yang dibangun mampu melakukan proses penerjemahan dua arah, yaitu Bahasa Indonesia ke Bahasa Daerah (Ternate, Makian Dalam, Galela) dan Bahasa Daerah ke Bahasa Indonesia dengan memanfaatkan kombinasi RBMT sebagai dasar pencocokan 5.572 kosakata dalam kamus, sedangkan GPT berfungsi menyusun ulang struktur kalimat agar hasil terjemahan lebih natural dan sesuai konteks bahasa tujuan. Hasil pengujian terhadap 468 kalimat menunjukkan bahwa sistem menghasilkan empat kategori hasil utama, yaitu: hasil sesuai dengan target, hasil berbeda bentuk namun memiliki makna yang sama, hasil tidak sesuai karena kosakata belum tersedia dalam kamus, serta hasil tidak tepat akibat kesalahan pemilihan makna oleh model GPT pada kata yang memiliki lebih dari satu arti.

Sistem mampu menangani variasi struktur bahasa daerah yang berbeda dengan Bahasa Indonesia, seperti perbedaan posisi angka, negasi, serta struktur frasa, namun penggunaan tanda baca dan pemisahan kata sangat berpengaruh terhadap hasil terjemahan, karena perbedaan struktur penulisan dapat menghasilkan makna yang berbeda, terutama pada kata yang memiliki bentuk serupa namun makna berbeda. Oleh karena itu, apabila model GPT belum mampu menentukan makna yang tepat pada kata yang memiliki ambiguitas, maka pengembang perlu melakukan penyesuaian pada struktur penulisan dan tanda baca dalam basis data agar sistem dapat menghasilkan output yang lebih akurat dan sesuai konteks. Sebagai contoh, kata “hate” dapat diberi penanda berbeda, seperti

hate”=“batang”, hate’=“kayu”, dan hate’’=“pohon”, dengan perbedaan tanda kutip, sehingga AI dapat mengenali perbedaan makna berdasarkan struktur penulisannya. Berdasarkan proses implementasi dan pengujian, sistem mampu mengelola 5.572 kosakata dengan kapasitas script mencapai ± 7.000 baris per file tanpa mengalami gangguan. Proses deployment melalui *GitHub* dan *Vercel* juga berjalan stabil, serta website dapat diakses menggunakan domain gratis Vercel pada *bahasamalut-rho.vercel.app*. Selain itu, ketika token API AI habis, sistem tetap dapat digunakan dalam mode kamus (RBMT), sehingga fitur pencarian kata, suara, ejaan, dan tampilan tetap berjalan normal meskipun koreksi kalimat berbasis AI tidak aktif. Secara keseluruhan, sistem yang dibangun termasuk dalam kategori hybrid translation system, karena menggabungkan pendekatan berbasis aturan (RBMT) dan kecerdasan buatan (GPT) untuk meningkatkan kualitas hasil terjemahan.

Pengembangan penelitian selanjutnya disarankan untuk memperluas jumlah kosakata pada kamus Bahasa Ternate, Makian, dan Galela agar cakupan terjemahan menjadi lebih lengkap dan akurat, serta meningkatkan kemampuan model AI melalui fine-tuning atau optimasi prompt untuk mengurangi kesalahan pada kata ambigu dan struktur kalimat yang kompleks. Sistem juga dapat dikembangkan dalam bentuk aplikasi mobile agar lebih mudah diakses oleh masyarakat, terutama di daerah dengan keterbatasan perangkat komputer. Selain itu, penelitian berikutnya dapat menambahkan metode Natural Language Processing (NLP) yang lebih kompleks seperti Named Entity Recognition (NER) atau sequence-to-sequence model untuk meningkatkan kualitas struktur bahasa dan konteks terjemahan. Pengujian sistem dengan jumlah responden, variasi kalimat, dan data bahasa yang lebih luas juga diperlukan agar evaluasi akurasi dan performa sistem menjadi lebih mendalam.

Referensi

- [1] N. Fokaaya, A. Lestari, M. Taha, F. Febriningsih, K. Bahasa, and P. Maluku, “Bahasa-Bahasa Daerah Bahasa-Bahasa Daerah,” 2014.
- [2] A. U. Sulami Sibua, Erviani La Ane Mbongo, Widya Astuti Bilo, Mandaian Fitrianti M, Dewilsta Bunga, “Desiminasi Kamus Praktis Bahasa Melayu Dialek Ternate untuk Wisatawan Lokal dan Internasional,” *J. Ilm. Wahana Pendidik.*, vol. 8, no. July, pp. 1–23, 2022.
- [3] D. S. Rainu, Kamus Bahasa Indonesia–Galela. Ternate, Indonesia: Kantor Bahasa Maluku Utara, Kementerian Pendidikan dan Kebudayaan, 2016.
- [4] R. M. Hamzah, Gamam Ela-Ela. Jakarta, Indonesia: Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi, 2024.
- [5] I. H. Malan, Ada Pipis Suat Ni Liko Li. Jakarta, Indonesia: Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi, 2024.
- [6] N. Saman, Kaka Madoto. Jakarta, Indonesia: Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi, 2023.
- [7] W. M. Hanif, M. Fikry, F. Yanto, and S. K. Gusti, “Bulletin Of Computer Science Research Implementasi Metode RBMT dalam Penerjemahan Bahasa Indonesia ke Bahasa Makassar,” vol. 6, no. 1, pp. 409–417, 2025, doi: 10.47065/bulletincsr.v6i1.935.
- [8] V. Raunak, A. Sharaf, Y. Wang, H. H. Awadalla, and A. Menezes, “Leveraging GPT-4 for Automatic Translation Post-Editing,” *Find. Assoc. Comput. Linguist. EMNLP 2023*, pp. 12009–12024, 2023, doi: 10.18653/v1/2023.findings-emnlp.804.
- [9] X. Song, A. Salcianu, Y. Song, D. Dopson, and D. Zhou, “Fast WordPiece Tokenization,” *EMNLP 2021 - 2021 Conf. Empir. Methods Nat. Lang. Process. Proc.*, no. Section 3, pp. 2089–2103, 2021, doi: 10.18653/v1/2021.emnlp-main.160.
- [10] D. Liu and S. Xu, “Comparison of hash table performance with open addressing and closed addressing: An empirical study,” *Int. J. Networked Distrib. Comput.*, vol. 3, no. 1, pp. 60–68, 2015, doi: 10.2991/ijndc.2015.3.1.7.
- [11] Y. Wang *et al.*, “Tacotron: Towards end-To-end speech synthesis,” *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 2017-Augus, pp. 4006–4010, 2017, doi: 10.21437/Interspeech.2017-1452.
- [12] J. Keiser and D. Lemire, “On-demand JSON: A better way to parse documents?,” *Softw. - Pract. Exp.*, vol. 54, no. 6, pp. 1074–1086, 2024, doi: 10.1002/spe.3313.
- [13] J. M. Bintang, M. F. Ashshidiq, and H. F. Dzakwan, “Penerapan Algoritma String Matching dan Regular Expression pada Aplikasi Kamus Besar Bahasa Indonesia (KBBI),” *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 4, no. 1, pp. 34–41, 2023, doi: 10.37148/bios.v4i1.57.
- [14] B. Ferreira, E. S. de Moura, and A. da Silva, *Applying burst-tries for error-tolerant prefix search*, vol. 25, no. 4. Springer Netherlands, 2022. doi: 10.1007/s10791-022-09416-9.
- [15] H. Yang, “The Application of Game-based Phonics Learning Apps in Lower Elementary Teaching,” *J. Educ. Humanit. Soc. Sci.*, vol. 8, pp. 1939–1946, 2023, doi: 10.54097/ehss.v8i.4618.