

**PENGEMBANGAN ASISTEN VIRTUAL BERBASIS LLM DAN
LANGCHAIN UNTUK ANALISIS KOMPARATIF ARTIKEL JURNAL
BERBAHASA INGGRIS DENGAN PENYAJIAN RINGKASAN
KONTEKSTUAL BERBAHASA INDONESIA**

SKRIPSI

Diajukan untuk memenuhi sebagian syarat untuk memperoleh gelar
Sarjana Teknik pada Program Studi Mekatronika dan Kecerdasan Buatan



Oleh:

Nurhikam

NIM 2101216

**PROGRAM STUDI MEKATRONIKA DAN KECERDASAN BUATAN
KAMPUS UPI DI PURWAKARTA
UNIVERSITAS PENDIDIKAN INDONESIA
2025**

LEMBAR HAK CIPTA

PENGEMBANGAN ASISTEN VIRTUAL BERBASIS LLM DAN LANGCHAIN UNTUK ANALISIS KOMPARATIF ARTIKEL JURNAL BERBAHASA INGGRIS DENGAN PENYAJIAN RINGKASAN KONTEKSTUAL BERBAHASA INDONESIA

Ditulis Oleh

Nurhikam

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Teknik pada Program Studi Mekatronika dan Kecerdasan Buatan

© **Nurhikam**

Universitas Pendidikan Indonesia

Juni 2025

Hak Cipta dilindungi undang-undang

Skripsi ini tidak dapat diperbanyak seluruhnya atau sebagian dengan cara apapun
tanpa izin tertulis dari penulis.

LEMBAR PENGESAHAN

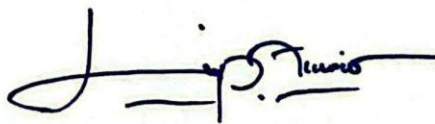
Nurhikam

2101216

**PENGEMBANGAN ASISTEN VIRTUAL BERBASIS LLM DAN
LANGCHAIN UNTUK ANALISIS KOMPARATIF ARTIKEL JURNAL
BERBAHASA INGGRIS DENGAN PENYAJIAN RINGKASAN
KONTEKSTUAL BERBAHASA INDONESIA**

Disetujui dan disahkan oleh pembimbing,

Pembimbing I



Liptia Venica, S.T., M.T.
NIP. 920210919941203201

Pembimbing II



Mahmudah Salwa Gianti, S.Si., M.Eng.
NIP. 920210919960408201

Mengetahui,

**Ketua Program Studi
Mekatronika dan Kecerdasan Buatan,**



Dewi Indriati Hadi Putri, S.Pd., M.T.
NIP. 920190219900126201

LEMBAR KEASLIAN SKRIPSI DAN ANTI PLAGIARISME

Dengan ini saya menyatakan bahwa skripsi yang berjudul "Pengembangan Asisten Virtual Berbasis LLM dan LangChain untuk Analisis Komparatif Artikel Jurnal Berbahasa Inggris dengan Penyajian Ringkasan Kontekstual Berbahasa Indonesia" beserta seluruh isinya adalah benar-benar merupakan hasil karya saya sendiri dan bukan merupakan plagiat atau saduran dari karya orang lain yang saya akui sebagai karya saya.

Adapun bagian-bagian tertentu dalam penulisan skripsi ini yang saya kutip dari karya orang lain telah dituliskan sumbernya secara jelas sesuai dengan norma, kaidah, dan etika penulisan ilmiah yang berlaku. Apabila di kemudian hari ditemukan adanya unsur plagiarisme atau ketidakaslian dalam skripsi ini, maka saya bersedia menerima sanksi akademik sesuai dengan peraturan perundang-undangan yang berlaku.

Purwakarta, 30 Juni 2025



Nurhikam
NIM. 2101216

KATA PENGANTAR: UCAPAN TERIMA KASIH

Bismillāhirrahmānirrahīm.

Puji syukur penulis panjatkan kehadiran Allah SWT., karena atas berkat, rahmat, dan karunia-Nya, penulis dapat menyelesaikan penyusunan skripsi yang berjudul "Pengembangan Asisten Virtual Berbasis LLM dan LangChain untuk Analisis Komparatif Artikel Jurnal Berbahasa Inggris Dengan Penyajian Ringkasan Kontekstual Berbahasa Indonesia" dengan baik. Shalawat serta salam semoga senantiasa tercurah kepada junjungan kita, Nabi Muhammad SAW, beserta keluarga, sahabat, dan para pengikutnya hingga akhir zaman.

Penyusunan skripsi ini merupakan salah satu syarat untuk memperoleh gelar Sarjana Teknik pada Program Studi Mekatronika dan Kecerdasan Buatan, Universitas Pendidikan Indonesia. Penulis menyadari bahwa penyelesaian skripsi ini adalah sebuah proses yang menuntut ketekunan, dedikasi, dan komitmen tinggi dalam menghadapi berbagai tantangan dan hambatan. Selama proses tersebut, banyak pihak yang telah memberikan dukungan, bimbingan, doa, dan semangat yang tak ternilai harganya. Oleh karena itu, dengan segala kerendahan hati, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada::

1. Kedua orang tua tercinta, Bapak Dudi Kusnadi dan Ibu Eti Heryati, S.Pd., serta adik tersayang Kirana Jayasri dan seluruh keluarga besar yang senantiasa memberikan doa restu, dukungan moral dan material, serta kasih sayang yang tak terhingga sebagai sumber motivasi dan kekuatan penulis dalam menyelesaikan studi.
2. Ibu Liptia Venica, S.T., M.T., selaku Dosen Pembimbing I, yang telah memberikan arahan, masukan, dan bimbingan yang sangat berharga dengan penuh kesabaran dan dedikasi tinggi dari tahap awal penelitian hingga penyelesaian skripsi ini.

3. Ibu Mahmudah Salwa Gianti, S.Si., M.Eng., selaku Dosen Pembimbing II, yang telah meluangkan waktu dan pikiran untuk memberikan bimbingan, koreksi, dan saran-saran konstruktif demi kesempurnaan skripsi ini.
4. Ibu Dewi Indriati Hadi Putri, S.Pd., M.T., Ketua Program Studi Mekatronika dan Kecerdasan Buatan beserta seluruh dosen dan staf yang telah memberikan ilmu pengetahuan, pengalaman, dan wawasan yang luas selama masa perkuliahan, yang menjadi fondasi penting dalam penyusunan penelitian ini.
5. Tim *Expert Panel* yang terdiri dari Bapak Muhammad Ryan Hafiez Nugraha, S.Pd., Bapak Gian Fajar Gemilang, S.Pd., dan Bapak Lisan Shidqi Zul Fahmi, S.Pd., yang telah berkenan memberikan validasi serta masukan yang sangat berharga untuk meningkatkan kualitas dan kredibilitas penelitian ini.
6. Dosen Pembimbing Akademik pada masa awal perkuliahan yang menjadi motivasi penulis untuk tetap berada di Program Studi ini. Meskipun kini sudah tidak mengajar, nasihat dan inspirasi yang diberikan tetap menjadi dorongan berharga dalam perjalanan akademik penulis.
7. Rekan-rekan mahasiswa Program Studi Mekatronika dan Kecerdasan Buatan yang telah bersama-sama menempuh perkuliahan, berbagi ilmu pengetahuan, dan memberikan dukungan moral serta semangat dalam menghadapi berbagai tantangan akademik selama masa studi.
8. Rekan kerja dan kolega dari berbagai institusi tempat penulis pernah bekerja (PT PLN (Persero), Regtech, Braincore, dan Stechoq) yang telah memberikan pengalaman kerja berharga dan wawasan industri yang memperkaya perspektif penulis dalam penelitian ini.
9. Partner terbaik, yang telah memberikan dukungan emosional, motivasi, dan semangat kepada penulis selama proses penyelesaian studi.
10. Semua pihak yang tidak dapat penulis sebutkan satu per satu, yang telah memberikan kontribusi, bantuan, dan dukungan dalam berbagai bentuk selama proses penyusunan skripsi ini.

Penulis menyadari sepenuhnya bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, kritik dan saran yang membangun dari para pembaca sangat penulis harapkan untuk perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat, baik secara teoretis maupun praktis, serta menjadi kontribusi bagi pengembangan ilmu pengetahuan, khususnya di bidang kecerdasan buatan.

Purwakarta, 30 Juni 2025



Nurhikam
NIM. 2101216

**PENGEMBANGAN ASISTEN VIRTUAL BERBASIS LLM DAN
LANGCHAIN UNTUK ANALISIS KOMPARATIF ARTIKEL JURNAL
BERBAHASA INGGRIS DENGAN PENYAJIAN RINGKASAN
KONTEKSTUAL BERBAHASA INDONESIA**

Nurhikam

2101216

ABSTRAK

Keterbatasan akses dan pemahaman terhadap artikel ilmiah berbahasa Inggris menjadi tantangan bagi komunitas akademik di Indonesia, terutama di tengah pertumbuhan pesat jumlah publikasi ilmiah global. Hal ini menghambat proses kajian literatur yang mendalam dan relevan. Penelitian ini bertujuan mengembangkan asisten virtual yang mampu melakukan analisis komparatif otomatis terhadap dua artikel jurnal dan menyajikannya dalam ringkasan kontekstual berbahasa Indonesia. Jenis penelitian ini adalah *Research and Development* (R&D) dengan pendekatan *Waterfall*. Sistem dikembangkan menggunakan model LLM GPT-4o-mini yang diorkestrasi oleh *LangChain*, serta antarmuka *Streamlit*. Evaluasi dilakukan melalui *pengujian performa*, validasi oleh tiga ahli, dan pengukuran akurasi semantik menggunakan BERTScore. Hasil menunjukkan waktu pemrosesan rata-rata 3 menit 50 detik. Validasi ahli memberikan skor 4,47 dari 5 (89,31%), dan akurasi semantik mencapai BERTScore 0,80. *Usability testing* menghasilkan skor 4,15 dari 5 (83,00%), membuktikan sistem mudah digunakan dan akurat. Sistem ini efektif menjembatani kesenjangan bahasa dan kompleksitas analisis, serta mempercepat kajian literatur bagi akademisi di Indonesia.

Kata Kunci: Asisten Virtual, *Large Language Model (LLM)*, *LangChain*, Analisis Komparatif, Jurnal Ilmiah, *Natural Language Processing (NLP)*, GPT-4o mini.

DEVELOPMENT OF AN LLM-BASED VIRTUAL ASSISTANT USING LANGCHAIN FOR COMPARATIVE ANALYSIS OF ENGLISH JOURNAL ARTICLES WITH INDONESIAN CONTEXTUAL SUMMARIES

Nurhikam

2101216

ABSTRACT

Limited access to and understanding of English-language scientific articles remains a significant challenge for the academic community in Indonesia, especially amidst the rapid growth of global scientific publications. This hampers the ability to conduct in-depth and relevant literature reviews. This study aims to develop a virtual assistant capable of automatically performing comparative analysis between two journal articles and presenting the results in contextual summaries in Indonesian. The research adopts a Research and Development (R&D) approach using the Waterfall model. The system is developed using the GPT-4o-mini large language model orchestrated by LangChain, with a Streamlit-based interface. Evaluation was conducted through performance testing, expert validation by three domain experts, and semantic accuracy measurement using BERTScore. Results show an average processing time of 3 minutes and 50 seconds. Expert validation yielded a quality score of 4.47 out of 5 (89.31%), and semantic accuracy reached a BERTScore of 0.80. Usability testing produced a score of 4.15 out of 5 (83.00%), indicating that the system is both user-friendly and accurate. This system has proven effective in bridging the language barrier and analytical complexity, thereby accelerating literature review processes and improving access to scientific knowledge for the academic community in Indonesia.

Keywords: *Virtual Assistant, Large Language Model (LLM), LangChain, Comparative Analysis, Scientific Journal, Natural Language Processing (NLP), GPT-4o mini.*

DAFTAR ISI

LEMBAR HAK CIPTA	ii
LEMBAR PENGESAHAN	iii
LEMBAR KEASLIAN SKRIPSI DAN ANTI PLAGIARISME	iv
KATA PENGANTAR: UCAPAN TERIMA KASIH	v
ABSTRAK	viii
<i>ABSTRACT</i>	ix
DAFTAR ISI	x
DAFTAR TABEL	xv
DAFTAR GAMBAR	xvi
DAFTAR KODE PROGRAM	xvii
DAFTAR LAMPIRAN	xviii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Penelitian	5
1.4 Tujuan Penelitian	6
1.5 Manfaat Penelitian	7
1.5.1 Manfaat bagi Peneliti dan Akademisi	7
1.5.2 Manfaat Bagi Masyarakat Umum	8
1.5.3 Manfaat Bagi Pengetahuan dan Teknologi	8
1.6 Struktur Penulisan Skripsi	9
BAB II TINJAUAN PUSTAKA	10
2.1 Asisten Virtual Berbasis LLM	10
2.1.1 Definisi dan Karakteristik Asisten Virtual	10

2.1.2	Asisten Virtual Berbasis LLM	11
2.2	<i>Large Language Model (LLM)</i>	13
2.2.1	Konsep Dasar dan Cara Kerja LLM.....	13
2.2.2	GPT-4o-mini: Arsitektur dan Kemampuan	16
2.2.3	Perbandingan Model LLM.....	18
2.2.4	LLM untuk Pemrosesan Dokumen Ilmiah.....	19
2.2.5	<i>Prompt Engineering</i> pada LLM	20
2.2.5.1	Definisi dan Konsep Dasar	20
2.2.5.2	Komponen Utama Prompt yang Efektif.....	21
2.2.5.3	Teknik-Teknik Prompt Engineering	21
2.3	LangChain.....	23
2.3.1	Pengertian dan Fungsi LangChain	23
2.3.2	Komponen Utama	24
2.3.3	LangSmith.....	26
2.3.4	Integrasi LangChain dengan LLM.....	27
2.3.5	Framework LangChain untuk Analisis Teks	28
2.4	Analisis Komparatif Otomatis	29
2.4.1	Metode Analisis Komparatif Teks dengan LLM.....	29
2.4.2	Parameter Perbandingan Artikel Ilmiah.....	30
2.4.3	Implementasi LLM untuk Analisis Komparatif	31
2.5	Jurnal Ilmiah Bidang Sosial Humaniora Berbahasa Inggris	33
2.5.1	Struktur dan Karakteristik Artikel Jurnal Ilmiah.....	33
2.5.2	Tantangan Pengolahan Jurnal Berbahasa Inggris dalam konteks Sosial dan Humaniora.....	34
2.6	Teknik Evaluasi Sistem Asisten Virtual Berbasis LLM.....	35
2.6.1	Skala Likert sebagai Metode Evaluasi	35

2.6.2	Metrik Evaluasi Kuantitatif.....	36
2.6.3	Metrik Evaluasi Kualitatif.....	37
2.6.4	Evaluasi Efektivitas Analisis Komparatif	38
2.6.5	<i>Benchmarking</i> dan Pengujian Performa.....	38
2.6.6	BERTScore: Evaluasi Berbasis <i>Contextual Embeddings</i>	39
BAB III METODOLOGI.....		41
3.1	Analisis dan Perencanaan (Analysis & Planning).....	42
3.1.1	Kajian Literatur dan Identifikasi Masalah	43
3.1.2	Perencanaan Strategis Pengembangan	43
3.1.3	<i>Expert Validation Panel</i>	44
3.2	Perancangan Sistem (Design System)	45
3.2.1	Alat dan Bahan.....	45
3.2.2	<i>Flowchart</i> Sistem	46
3.2.3	<i>Wireframe</i> Antarmuka Pengguna	48
3.3	Pengembangan dan Implementasi (Development & Implementation).....	50
3.3.1	Rencana Pengembangan Sistem.....	50
3.3.2	Arsitektur Sistem.....	51
3.4	Pengujian dan Validasi (Testing & Validation)	53
3.4.1	Skenario Pengujian.....	53
3.4.1.1	Pemilihan Artikel Uji Coba.....	54
3.4.1.2	Penentuan Tingkat Kompleksitas Artikel.....	54
3.4.1.3	Skenario Pengujian untuk Validasi <i>Expert</i>	56
3.4.2	Prosedur Pengumpulan Data Evaluasi	58
3.4.2.1	Pengumpulan <i>Ground Truth</i> dari <i>Expert</i>	58
3.4.2.2	Pengumpulan Data Validasi <i>Expert</i>	59
3.4.2.3	Pengumpulan Data <i>Usability Testing</i>	59

3.5	Evaluasi dan Penyimpulan (Evaluation & Conclusion).....	60
3.5.1	Evaluasi Hasil Pengujian Sistem.....	61
3.5.2	Evaluasi BERTScore.....	61
3.5.3	Evaluasi Hasil Validasi <i>Expert</i>	62
3.5.4	Evaluasi <i>Usability Testing</i>	62
3.5.5	Penyimpulan dan Perumusan Rekomendasi	62
BAB IV HASIL DAN PEMBAHASAN.....		64
4.1	Implementasi dan Pengembangan Sistem.....	64
4.1.1	Pengembangan Modul <i>Input</i> dan <i>Preprocessing</i>	64
4.1.2	Implementasi Modul Analisis Komparatif Bahasa Indonesia	69
4.1.3	Struktur <i>Output</i>	81
4.1.4	Integrasi Model dengan Antarmuka Berbasis <i>Web</i>	81
4.1.4.1	Tampilan Antarmuka Berbasis <i>Web</i>	81
4.1.4.2	Struktur <i>Output</i> Sistem.....	85
4.2	<i>Monitoring</i> dan <i>Tracking</i> dengan LangSmith	86
4.2.1	Implementasi LangSmith	86
4.2.2	Parameter Optimalisasi	88
4.2.3	Evaluasi Performa Sistem	89
4.2.3.1	Analisis <i>Success Rate</i> dan <i>Error Detection</i>	89
4.2.3.2	Analisis <i>Latency</i> dan <i>Performance</i>	90
4.2.3.3	Analisis <i>Token Usage</i> dan <i>Cost Efficiency</i>	90
4.3	Analisis Hasil Pengujian Sistem	92
4.3.1	Hasil Pengujian Waktu Pemrosesan.....	92
4.3.2	Analisis Performa Berdasarkan Kompleksitas Artikel.....	93
4.3.2.1	Performa pada Kombinasi Setara.....	94
4.3.2.2	Performa pada Kombinasi Beda Satu Tingkat	94

4.3.2.3	Performa pada Kombinasi Beda Dua Tingkat.....	94
4.4	Hasil Evaluasi dan Validasi dari <i>Expert Panel</i>	95
4.4.1	Analisis Akurasi Ringkasan Komparatif dengan BERTScore	95
4.4.1.1	<i>Ground Truth</i> dari <i>Expert</i> secara Manual.....	95
4.4.1.2	Hasil Evaluasi BERTScore	96
4.4.2	Analisis Statistik Hasil Validasi <i>Expert</i>	97
4.4.2.1	Analisis Data <i>Expert</i> 1	97
4.4.2.2	Analisis Data <i>Expert</i> 2	98
4.4.2.3	Analisis Data <i>Expert</i> 3	99
4.4.3	Validasi <i>Expert Panel</i>	100
4.4.3.1	Validasi dari <i>Academic Practitioner</i>	100
4.4.3.2	Validasi dari <i>Graduate Researcher</i>	101
4.4.3.3	Validasi dari <i>Field Practitioner</i>	102
4.4.4	Analisis Hasil Validasi Ketiga <i>Expert</i>	103
4.5	Hasil <i>Usability Testing</i>	105
4.5.1	Analisis Statistik Hasil <i>Usability Testing</i>	105
4.5.2	Analisis Performa Berdasarkan Aspek <i>Usability</i>	106
4.5.3	Perbandingan Hasil <i>Usability Testing</i> dengan <i>Expert Validation</i>	107
BAB V KESIMPULAN DAN SARAN.....		108
5.1	Kesimpulan	108
5.2	Saran	109
DAFTAR PUSTAKA		111
LAMPIRAN.....		124

DAFTAR TABEL

Tabel 2.1 Contoh Penerapan Teknik <i>Prompt Engineering</i>	22
Tabel 3.1 Spesifikasi Alat dan Bahan	46
Tabel 3.2 Kriteria Penentuan Tingkat Kompleksitas Artikel	55
Tabel 3.3 Artikel Pengujian dan Klasifikasi Kompleksitas	55
Tabel 3.4 Skenario Kombinasi Pengujian Sistem	57
Tabel 3.5 Format Pembuatan <i>Ground Truth</i>	58
Tabel 4.1 Spesifikasi Teknis Modul <i>Input</i> dan <i>Preprocessing</i>	65
Tabel 4.2 Contoh Masalah pada Hasil Ekstraksi Teks Mentah	66
Tabel 4.3 Hasil Pembersihan Karakter <i>Non-printable</i>	67
Tabel 4.4 Hasil Normalisasi Teks	68
Tabel 4.5 Distribusi <i>Chunk</i> Sampel Dokumen	69
Tabel 4.6 Struktur <i>Template Prompt</i> dengan Teknik <i>Chain-of-Thought</i>	72
Tabel 4.7 <i>Prompt</i> Penerjemahan untuk <i>System</i> dan <i>Human</i>	74
Tabel 4.8 Hasil <i>Output</i> Modul <i>Summarization</i>	76
Tabel 4.9 Contoh <i>Prompt</i> Permintaan <i>Output</i> dalam Format JSON Terstruktur ..	79
Tabel 4.10 Waktu Pemrosesan	92
Tabel 4.11 Hasil Evaluasi BERTScore antara <i>Output</i> Sistem dan <i>Ground Truth</i> .	96
Tabel 4.12 Hasil Evaluasi <i>Expert Panel 1</i>	97
Tabel 4.13 Hasil Evaluasi <i>Expert Panel 2</i>	98
Tabel 4.14 Hasil Evaluasi <i>Expert Panel 3</i>	99
Tabel 4.15 Perbandingan Skor Rata-rata Keseluruhan Antar <i>Expert</i>	103
Tabel 4.16 Hasil <i>Usability Testing</i>	105
Tabel 4.17 Perbandingan <i>Usability Testing</i> dan <i>Expert Validation</i>	107

DAFTAR GAMBAR

Gambar 2.1 Arsitektur <i>Transformers</i>	14
Gambar 2.2 Struktur dan Proses <i>Multi-Head Attention</i>	15
Gambar 2.3 Ilustrasi Perhitungan BERTScore.....	39
Gambar 3.1 Diagram Pendekatan <i>Waterfall</i>	42
Gambar 3.2 <i>Flowchart</i> Sistem Analisis Komparatif Otomatis	47
Gambar 3.3 <i>Wireframe</i> Tampilan Awal	48
Gambar 3.4 <i>Wireframe</i> Hasil Ringkasan.....	49
Gambar 3.5 <i>Wireframe</i> Tabel Perbandingan	49
Gambar 3.6 Diagram Arsitektur Sistem.....	51
Gambar 4.1 Struktur Layout Multi-kolom dari Sampel Dokumen	65
Gambar 4.2 Hasil Ekstraksi Teks dari Sampel Dokumen Multi-kolom.....	66
Gambar 4.3 Teks Hasil Ekstraksi Sebelum Proses Normalisasi	69
Gambar 4.4 Teks Hasil Ekstraksi setelah Proses Normalisasi	69
Gambar 4.5 <i>Input Prompt</i> dengan Format <i>System-human</i>	71
Gambar 4.6 <i>Dashboard Monitoring</i> LangSmith untuk Proses Ekstraksi.....	73
Gambar 4.7 Contoh <i>Output</i> Ekstraksi <i>Chunk</i> Pertama	73
Gambar 4.8 <i>Output</i> Hasil Ekstraksi Elemen Kunci dan Translasi	76
Gambar 4.9 Output Akhir Tabel Berbentuk JSON.....	80
Gambar 4.10 Tampilan Awal Sistem (Unggah Dokumen).....	82
Gambar 4.11 Tampilan Proses Membandingkan Artikel (Indikator Progres).....	83
Gambar 4.12 Tampilan Hasil Perbandingan Analisis Umum	83
Gambar 4.13 Tampilan Hasil Perbandingan Berdasarkan Area Fokus	84
Gambar 4.14 Tampilan Tabel Perbandingan Artikel	84
Gambar 4.15 Grafik <i>Trace Count Success & Error</i>	89
Gambar 4.16 Grafik <i>Trace Latency</i> P50 dan P99	90
Gambar 4.17 Statistik Komprehensif Performa Sistem untuk Satu Kali Analisis	91
Gambar 4.18 Grafik Total <i>Cost</i> untuk Satu Kali Analisis Komparatif	92

DAFTAR KODE PROGRAM

Kode Program 4.1 <i>Syntax</i> Python untuk <i>Setup</i> OpenAI API	70
Kode Program 4.2 Implementasi <i>memory buffer</i> pada LangChain.....	71
Kode Program 4. 3 <i>Syntax</i> untuk <i>Setup</i> dan Inisialisasi LangSmith.....	87

DAFTAR LAMPIRAN

A. Rincian Perolehan Poin untuk Klasifikasi Artikel Pengujian	124
B. Pertanyaan Formulir Evaluasi untuk <i>Expert Panel</i>	125
C. Hasil <i>Ground Truth</i> Dari <i>Expert 1</i> dan <i>Expert 2</i>	129
D. Hasil Ringkasan Ujicoba Sistem.....	137
E. Pertanyaan Formulir <i>Usability Testing</i>	140

DAFTAR PUSTAKA

- [1] Y. Zhou, “Navigating the currents of natural language processing: A comprehensive overview of modern techniques and applications,” *Appl. Comput. Eng.*, vol. 74, no. 1, pp. 213–218, 2024, doi: 10.54254/2755-2721/74/20240475.
- [2] Dr. Chandrika G, “Natural language processing,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 12, no. VI, pp. 1092–1095, 2024, doi: 10.1016/0010-4655(90)90107-C.
- [3] A. Sohail and L. Zhang, “Using large language models to facilitate academic work in the psychological sciences,” *Curr. Psychol.*, pp. 1–17, 2025, doi: 10.1007/s12144-025-07438-2.
- [4] J. Evans, J. D’Souza, and S. Auer, “Large Language Models as Evaluators for Scientific Synthesis,” 2024, doi: 10.1145/3640457.3688075.
- [5] B. Wang, X. Zhang, S. Li, and Y. Wang, “The Practice of Enhancing Learning and Scientific Innovative Abilities Using LLM-Based AI Tools,” *Proc. - 2024 6th Int. Conf. Comput. Sci. Technol. Educ. CSTE 2024*, pp. 166–170, 2024, doi: 10.1109/CSTE62025.2024.00038.
- [6] M. Cavojsky, G. Bugar, T. Kormanik, and M. Hasin, “Exploring the Capabilities and Possible Applications of Large Language Models for Education,” *ICETA 2023 - 21st Year Int. Conf. Emerg. eLearning Technol. Appl. Proc.*, pp. 91–98, 2023, doi: 10.1109/ICETA61311.2023.10344166.
- [7] Devia Mutiara Sari, “Analysis Students Difficulty in Reading English Article,” *Jotika J. Educ.*, vol. 3, no. 1, pp. 27–31, 2023, doi: 10.56445/jje.v3i1.107.
- [8] Hamamah, I. Emaliana, Y. Hapsari, P. D. D. Degeng, and A. C. Fadillah, “Using Nominal Group Technique to Explore Publication Challenges and the Usefulness of AI-Based Writing Technologies: Insights From Indonesian

- Scholars,” *Theory Pract. Lang. Stud.*, vol. 13, no. 8, pp. 2038–2047, 2023, doi: 10.17507/tpls.1308.20.
- [9] J. Rüländ, “When English language matters: Internationalizing PhD research by Indonesian political scientists,” *Englisia J. Lang. Educ. Humanit.*, vol. 10, no. 2, p. 137, 2023, doi: 10.22373/ej.v10i2.17607.
- [10] R. I. Naharuddin, P. I. Santosa, and T. B. Adji, “Systematic Literature Review: Automated Text Summarization for Indonesian Language,” *2023 10th Int. Conf. Inf. Technol. Comput. Electr. Eng. ICITACEE 2023*, pp. 254–259, 2023, doi: 10.1109/ICITACEE58587.2023.10277395.
- [11] G. Sharma, A. Paretkar, and D. Sharma, “Synthesizing Scientific Summaries : An Extractive and Abstractive approach,” pp. 1–10, 2024.
- [12] H. Li and X. Tuo, “Research on an English translation method based on an improved transformer model,” *J. Intell. Syst.*, vol. 31, no. 1, pp. 532–540, 2022, doi: 10.1515/jisys-2022-0038.
- [13] M. R. Raj, “An Article Summarizer Application using OpenAI’s GPT Model,” *Interantional J. Sci. Res. Eng. Manag.*, vol. 08, no. 03, pp. 1–5, 2024, doi: 10.55041/ijrsrem29606.
- [14] Scispace, “Scispace.” <https://scispace.com/>
- [15] SciAIHub, “SciSpace - AI assistant to discover and understand research papers”, [Online]. Available: <https://www.sciaihub.com/tools/scispace>
- [16] J. Y. Kung, “Elicit,” *J. Can. Heal. Libr. Assoc.*, 2023.
- [17] Elicit, “Elicit”, [Online]. Available: <https://elicit.com/>
- [18] University of Michigan Library, “Semantic Scholar — Scientific Literature Search Engine”, [Online]. Available: <https://guides.lib.umich.edu/c.php?g=1209331&p=9938580#s-lg-box-31387188>
- [19] University of Michigan Library, “Research Rabbit — Citation-based

Mapping”, [Online]. Available:
<https://guides.lib.umich.edu/c.php?g=1209331&p=9938580#s-lg-box-31387258>

- [20] Rakshitha Kiran P., “Leveraging LLaMA3 and LangChain for Rapid AI Application Development,” *J. Electr. Syst.*, vol. 20, no. 10s, pp. 2146–2153, 2024, doi: 10.52783/jes.5539.
- [21] A. A. Micheal, A. Prasanth, T. S. Aswin, and B. L. Krisha, “Advancing Educational Accessibility: The LangChain LLM Chatbot’s Impact on Multimedia Syllabus-Based Learning,” pp. 1–16, 2024.
- [22] L. Farinetti and L. Canale, “Chatbot Development Using LangChain: A Case Study to Foster Critical Thinking and Creativity,” *Annu. Conf. Innov. Technol. Comput. Sci. Educ. ITiCSE*, vol. 1, pp. 401–407, 2024, doi: 10.1145/3649217.3653557.
- [23] S. Ghane, R. Sawant, G. Supe, and C. Pichad, “LangchainIQ: Intelligent Content and Query Processing,” *Int. J. Manag. Technol. Soc. Sci.*, vol. 9, no. 3, pp. 34–43, 2024, doi: 10.47992/ijmts.2581.6012.0360.
- [24] A. S. Korat, “AI-Augmented LangChain: Facilitating Natural Language SQL Queries for Non-Technical Users,” *J. Artif. Intell. Cloud Comput.*, vol. 3, no. 3, pp. 1–5, 2024, doi: 10.47363/jaicc/2024(3)335.
- [25] K. Sharma, C. Castellini, E. L. van den Broek, A. Albu-Schaeffer, and F. Schwenker, “A dataset of continuous affect annotations and physiological signals for emotion analysis,” *Sci. Data*, vol. 6, no. 1, pp. 1–13, 2019, doi: 10.1038/s41597-019-0209-0.
- [26] Y. Mo, H. Qin, Y. Dong, Z. Zhu, and Z. Li, “Large Language Model (LLM) AI Text Generation Detection based on Transformer Deep Learning Algorithm,” *Int. J. Eng. Manag. Res. Peer Rev. Ref. J. e*, vol. 14, no. 2, 2024, [Online]. Available:
<https://ijemr.vandanapublications.comhttps://doi.org/10.5281/zenodo.11124>

- [27] X. Zheng *et al.*, “KS-LLM: Knowledge Selection of Large Language Models with Evidence Document for Question Answering,” 2024, [Online]. Available: <http://arxiv.org/abs/2404.15660>
- [28] A. Raj, M. Raj, N. Umasankari, and D. Geethanjali, “Document-Based Text Summarization using T5 small and gTTS,” *2024 Int. Conf. Adv. Data Eng. Intell. Comput. Syst. ADICS 2024*, pp. 1–6, 2024, doi: 10.1109/ADICS58448.2024.10533605.
- [29] F. Farhat, B. M. Chaudhry, M. Nadeem, S. S. Sohail, and D. Ø. Madsen, “Evaluating Large Language Models for the National Premedical Exam in India: Comparative Analysis of GPT-3.5, GPT-4, and Bard,” *JMIR Med. Educ.*, vol. 10, pp. 1–12, 2024, doi: 10.2196/51523.
- [30] M. E. Mamalis, E. Kalampokis, F. Fitsilis, G. Theodorakopoulos, and K. Tarabanis, “A Large Language Model Agent Based Legal Assistant for Governance Applications,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 14841 LNCS, pp. 286–301, 2024, doi: 10.1007/978-3-031-70274-7_18.
- [31] Z. Duan, “Application development exploration and practice based on LangChain+ChatGLM+Rasa,” *2023 2nd Int. Conf. Cloud Comput. Big Data Appl. Softw. Eng. CBASE 2023*, pp. 282–285, 2023, doi: 10.1109/CBASE60015.2023.10439133.
- [32] C. Lee, S. Kumar, K. A. Vogt, and S. Meraj, “Improving Clinical Documentation with AI : A Comparative Study of Sporo AI Scribe and GPT-4o mini,” p. 8.
- [33] B. Xu and M. M. Poo, “Large language models and brain-inspired general intelligence,” *Natl. Sci. Rev.*, vol. 10, no. 10, pp. 9–10, 2023, doi: 10.1093/nsr/nwad267.
- [34] C. Chen and A. Leitch, “LLMs as Academic Reading Companions:

- Extending HCI Through Synthetic Personae,” 2024, [Online]. Available: <http://arxiv.org/abs/2403.19506>
- [35] U. H. Khan, M. H. Khan, and R. Ali, “Large Language Model based Educational Virtual Assistant using RAG Framework,” *Procedia Comput. Sci.*, vol. 252, pp. 905–911, 2025, doi: 10.1016/j.procs.2025.01.051.
 - [36] N. P. Mane, A. P. Hanchate, V. H. Kumbhar, A. A. Patil, A. M. Kate, and V. R. Bhivse, “A Review Paper on AI Virtual Assistants,” no. March, 2025.
 - [37] A. Berkol and I. G. Demirtaş, “Advancing Human-Machine Interaction: Speech and Natural Language Processing,” *ISAS 2023 - 7th Int. Symp. Innov. Approaches Smart Technol. Proc.*, 2023, doi: 10.1109/ISAS60782.2023.10391814.
 - [38] P. P. Jagtap, D. Pardeshi, A. Bhombe, V. Nirmal, and V. Gangurde, “Enhancing Human-Computer Interaction : A Survey on Intelligent Virtual Assistants for Accessibility and Contextual Awareness,” pp. 82–85, 2024, doi: 10.48175/IJARSCT-22310.
 - [39] D. Luitse and W. Denkena, “The great transformer: Examining the role of large language models in the political economy of AI,” *Big Data Soc.*, vol. 8, no. 2, 2021, doi: 10.1177/20539517211047734.
 - [40] B. Almarie, P. E. P. Teixeira, K. Pacheco-Barrios, C. A. Rossetti, and F. Fregni, “The Use of Large Language Models in Science: Opportunities and Challenges,” *Princ. Pract. Clin. Res. J.*, vol. 9, no. 1, pp. 1–4, 2023, doi: 10.21801/ppcrj.2023.91.1.
 - [41] Y. Guan *et al.*, “Intelligent Agents with LLM-based Process Automation,” *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 5018–5027, 2024, doi: 10.1145/3637528.3671646.
 - [42] X. L. Dong, S. Moon, Y. E. Xu, K. Malik, and Z. Yu, “Towards Next-Generation Intelligent Assistants Leveraging LLM Techniques,” *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 5792–5793, 2023, doi:

10.1145/3580305.3599572.

- [43] A. Vaswani *et al.*, “Attention is all you need,” *Adv. Neural Inf. Process. Syst.*, vol. 2017-Decem, no. Nips, pp. 5999–6009, 2017.
- [44] B. Ampel, C.-H. Yang, J. Hu, and H. Chen, “Large Language Models for Conducting Advanced Text Analytics Information Systems Research,” *ACM Trans. Manag. Inf. Syst.*, 2024, doi: 10.1145/3682069.
- [45] R. Chew, J. Bollenbacher, M. Wenger, J. Speer, and A. Kim, “LLM-Assisted Content Analysis: Using Large Language Models to Support Deductive Coding,” 2023, [Online]. Available: <http://arxiv.org/abs/2306.14924>
- [46] E. P. Pozzobon and H. Pinheiro, “Implementation of a research assistant for literature review with Large Language Models,” *2024 16th Semin. Power Electron. Control. SEPOC 2024*, pp. 1–5, 2024, doi: 10.1109/SEPOC63090.2024.10747480.
- [47] N. Archanaa, B. Shivanesh, J. D. T. Suwin Kumar, G. Bharathi Mohan, and S. Doss, “Comparative Analysis of News Articles Summarization using LLMs,” *2024 Asia Pacific Conf. Innov. Technol. APCIT 2024*, pp. 1–6, 2024, doi: 10.1109/APCIT62007.2024.10673458.
- [48] M. Mitchell, “Large Language Models,” *MIT Press*, p. 11, 2024, doi: <https://doi.org/10.21428/e2759450.2bb20e3c>.
- [49] Q. Luo, W. Zeng, M. Chen, G. Peng, X. Yuan, and Q. Yin, “Self-Attention and Transformers: Driving the Evolution of Large Language Models,” *2023 IEEE 6th Int. Conf. Electron. Inf. Commun. Technol. ICEICT 2023*, pp. 401–405, 2023, doi: 10.1109/ICEICT57916.2023.10245906.
- [50] I. Jindal, C. Badrinath, P. Bharti, L. Vinay, and S. D. Sharma, “Balancing Continuous Pre-Training and Instruction Fine-Tuning: Optimizing Instruction-Following in LLMs,” 2024, [Online]. Available: <http://arxiv.org/abs/2410.10739>

- [51] K. Sun and M. Dredze, “Amuro & Char: Analyzing the Relationship between Pre-Training and Fine-Tuning of Large Language Models,” 2024, [Online]. Available: <http://arxiv.org/abs/2408.06663>
- [52] B. Kumar and V. Murugesan, “A comprehensive overview of LLM-based approaches for machine translation,” vol. 38, no. 1, pp. 344–356, 2025, doi: 10.11591/ijeecs.v38.i1.pp344-356.
- [53] Y. Liu, Y. Yang, and Y. Huang, “Automatic generation of comparative summary for scientific literature,” *Int. J. Performability Eng.*, vol. 14, no. 7, pp. 1570–1579, 2018, doi: 10.23940/ijpe.18.07.p21.15701579.
- [54] A. F. Mohammad, B. Clark, and R. Hegde, “Large Language Model (LLM) & GPT, A Monolithic Study in Generative AI,” *Proc. - 2023 Congr. Comput. Sci. Comput. Eng. Appl. Comput. CSCE 2023*, no. Llm, pp. 383–388, 2023, doi: 10.1109/CSCE60160.2023.00068.
- [55] OpenAI, “OpenAI Platform,” 2025. <https://platform.openai.com/>
- [56] OpenAI, “GPT-4o mini: advancing cost-efficient intelligence,” 2024. <https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/> (accessed May 14, 2025).
- [57] artificialanalysis.ai, “Claude 3 Haiku: Intelligence, Performance & Price Analysis.” <https://artificialanalysis.ai/models/claude-3-haiku>
- [58] artificialanalysis.ai, “Gemini 2.0 Flash (experimental): Intelligence, Performance & Price Analysis.” <https://artificialanalysis.ai/models/gemini-2-0-flash-experimental>
- [59] R. Diaz Milian, P. Moreno Franco, W. D. Freeman, and J. D. Halamka, “Revolution or Peril? The Controversial Role of Large Language Models in Medical Manuscript Writing,” *Mayo Clin. Proc.*, vol. 98, no. 10, pp. 1444–1448, 2023, doi: 10.1016/j.mayocp.2023.07.009.
- [60] M. Schilling-Wilhelmi *et al.*, “From Text to Insight: Large Language Models

- for Materials Science Data Extraction,” pp. 1125–1150, 2024, doi: 10.1039/d4cs00913d.
- [61] D. Scherbakov, N. Hubig, V. Jansari, A. Bakumento, and L. Lenert, “The emergence of Large Language Models (LLM) as a tool in literature reviews: an LLM automated systematic review Dmitry,” 2024, doi: <https://doi.org/10.48550/arXiv.2409.04600>.
- [62] S. Li *et al.*, “SciLitLLM: How to Adapt LLMs for Scientific Literature Understanding,” pp. 1–24, 2024, [Online]. Available: <http://arxiv.org/abs/2408.15545>
- [63] N. Wu, “Expanding Horizons in Prompt Engineering: Techniques , Frameworks , and Challenges,” pp. 1–9, 2025, doi: 10.55041/IJSREM40635.
- [64] S. Tripathi, R. Sukumaran, S. Dheer, and T. Cook, “Promptwise: Prompt Engineering Paradigm for Enhanced Patient-Large Language Model Interactions Towards Medical Education”.
- [65] D. Lamba, “The Role of Prompt Engineering in Improving Language Understanding and Generation,” vol. 6, no. 6, pp. 1–7, 2024.
- [66] X. Amatriain, “Prompt Design and Engineering: Introduction and Advanced Methods,” pp. 1–26, 2024, [Online]. Available: <http://arxiv.org/abs/2401.14423>
- [67] V. Mavroudis, ““? LangChain | ?? LangChain,” no. November, 2024, doi: 10.20944/preprints202411.0566.v1.
- [68] O. Topsakal and T. C. Akinci, “Creating Large Language Model Applications Utilizing LangChain: A Primer on Developing LLM Apps Fast,” *Int. Conf. Appl. Eng. Nat. Sci.*, vol. 1, no. 1, pp. 1050–1056, 2023, doi: 10.59287/icaens.1127.
- [69] T. Wu, M. Terry, and C. J. Cai, *AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts*, vol. 1,

- no. 1. Association for Computing Machinery, 2022. doi: 10.1145/3491102.3517582.
- [70] B. L. S, S. V. R, R. Mahadevan, and R. S. Raman, “Comparative Study and Framework for Automated Summariser Evaluation: LangChain and Hybrid Algorithms,” *arXiv Prepr.*, 2023, [Online]. Available: arxiv:2310.02759
- [71] L. Agarwal, “Comparison and Analysis of Large Language Models (LLMs)”.
- [72] B. K. S. Barbosa and C. E. C. Campelo, “LLMs as Tools for Evaluating Textual Coherence : A Comparative Analysis LLMs as Tools for Evaluating Textual Coherence : A Comparative Analysis,” no. November, 2024, doi: 10.5753/stil.2024.245379.
- [73] L. Zhou, “Qualitative Comparative Analysis (QCA): An Innovative Approach for Integrating Qualitative and Quantitative Analysis,” *Best Evid. Chinese Educ.*, vol. 10, no. 2, pp. 1354–1355, 2022, doi: 10.15354/bece.22.co003.
- [74] L. Yao, “Large Language Models are Contrastive Reasoners,” 2024, [Online]. Available: <http://arxiv.org/abs/2403.08211>
- [75] L. Rettenberger, M. F. Munker, M. Schutera, C. M. Niemeyer, K. S. Rabe, and M. Reischl, “Using Large Language Models for Extracting Structured Information From Scientific Texts,” *Curr. Dir. Biomed. Eng.*, vol. 10, no. 4, pp. 526–529, 2024, doi: 10.1515/cdbme-2024-2129.
- [76] A. Dunn *et al.*, “Structured information extraction from complex scientific text with fine-tuned large language models,” 2022, [Online]. Available: <http://arxiv.org/abs/2212.05238>
- [77] S. Dhaked *et al.*, “Comparative Analysis of Study Design and Statistical Test Utilization in Indian Journal of Community Medicine, Indian Journal of Public Health and Bulletin of the World Health Organization,” no. 1, pp. 1–8, 2019, doi: 10.4103/mamcjms.mamcjms.

- [78] V. Nechakhin, J. D'Souza, and S. Eger, "Evaluating Large Language Models for Structured Science Summarization in the Open Research Knowledge Graph," *Inf.*, vol. 15, no. 6, pp. 1–19, 2024, doi: 10.3390/info15060328.
- [79] M. Brucks and O. Toubia, "Prompt architecture induces methodological artifacts in large language models," *PLoS One*, vol. 20, no. 4 April, pp. 1–13, 2025, doi: 10.1371/journal.pone.0319159.
- [80] J. He, M. Rungta, D. Koleczek, A. Sekhon, F. X. Wang, and S. Hasan, "Does Prompt Formatting Have Any Impact on LLM Performance?," 2024, [Online]. Available: <http://arxiv.org/abs/2411.10541>
- [81] M. Gozzi and F. Di Maio, "Comparative Analysis of Prompt Strategies for Large Language Models: Single-Task vs. Multitask Prompts," *Electron.*, vol. 13, no. 23, 2024, doi: 10.3390/electronics13234712.
- [82] A. Kshirsagar, "Enhancing RAG Performance Through Chunking and Text Splitting Techniques," pp. 151–158, 2024.
- [83] J. Wei *et al.*, "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," *Adv. Neural Inf. Process. Syst.*, vol. 35, no. NeurIPS, pp. 1–43, 2022.
- [84] M. S. Robinson and F. L. Stoller, "Overview of the Journal Article," *Write Like a Chemist: A Guide and Resource*. Oxford University Press, p. 0, Mar. 28, 2023. doi: 10.1093/oso/9780190098940.003.0002.
- [85] E. I. Obeagu *et al.*, "Academic Journal Writing and Types of Journals," *INOSR Exp. Sci.*, vol. 12, no. 2, pp. 205–213, 2023, doi: 10.59298/inosres/2023/2.16.1000.
- [86] C. Gonzalez-Perez, "A conceptual modelling language for the humanities and social sciences," *Proc. - Int. Conf. Res. Challenges Inf. Sci.*, no. January, 2012, doi: 10.1109/RCIS.2012.6240430.
- [87] G. M. Sullivan and A. R. Artino, "Analyzing and Interpreting Data From

- Likert-Type Scales,” *J. Grad. Med. Educ.*, vol. 5, no. 4, pp. 541–542, 2013, doi: 10.4300/jgme-5-4-18.
- [88] M. Koo and S. Yang, “Likert-Type Scale,” pp. 1–11, 2025.
- [89] S. K. Wadkar, K. Singh, R. Chakravarty, and S. D. Argade, “Assessing the Reliability of Attitude Scale by Cronbach’s Alpha,” *J. Glob. Commun.*, vol. 9, no. 2, p. 113, 2016, doi: 10.5958/0976-2442.2016.00019.7.
- [90] Z. Hassani, H. R. Mokhtarinia, A. H. Kahlaee, and C. P. Gabel, “Translation, Validity, and Reliability of the Upper Extremity Fugl-Meyer Assessment (FMA-UE) in Persian Speaking Stroke Patients,” *Iran. Rehabil. J.*, vol. 20, no. 1, pp. 37–46, 2022, doi: 10.32598/irj.20.SpecialIssue.919.3.
- [91] J. P. A. I. Id and Z. Maniadis, “In defense of quantitative metrics in researcher assessments,” pp. 1–5, 2023, doi: 10.1371/journal.pbio.3002408.
- [92] O. Article and D. Enfermagem, “Validation by experts : importance in translation and adaptation of instruments,” vol. 38, no. 3, pp. 1–9, 2017.
- [93] Y. A. Purmala and F. Debora, “A Systematic Literature Review of Benchmarking Implementation in various Industries,” vol. 2, no. 1, pp. 35–51, 2021.
- [94] A. Purwanto, M. Fahlevi, R. Pramono, and J. T. Purba, “BENEFIT OF BENCHMARKING METHODS IN SEVERAL INDUSTRIES: A BENEFIT OF BENCHMARKING METHODS IN SEVERAL INDUSTRIES: A SYSTEMATIC LITERATURE REVIEW,” no. December, 2020.
- [95] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi, “BERTscore: Evaluating Text Generation With Bert,” *8th Int. Conf. Learn. Represent. ICLR 2020*, pp. 1–43, 2020.
- [96] GenisLab, “Monitoring Text-Based Generative AI Models Using Metrics Like Bleu Score.”

<https://www.genislab.com/exchange/usecases/agents/monitoring-text-based-generative-ai-models-using-metrics-like-bleu-score/>

- [97] M. Kaster, W. Zhao, and S. Eger, “Global Explainability of BERT-Based Evaluation Metrics by Disentangling along Linguistic Factors,” *EMNLP 2021 - 2021 Conf. Empir. Methods Nat. Lang. Process. Proc.*, pp. 8912–8925, 2021, doi: 10.18653/v1/2021.emnlp-main.701.
- [98] N. Rachma and I. Muhlas, “Comparison Of *Waterfall* And Prototyping Models In Research And Development (R&D) Methods For Android-Based Learning Application Design,” *J. Inov. Inov. Teknol. Inf. dan Inform.*, vol. 5, no. 1, p. 36, 2022, doi: 10.32832/inova-tif.v5i1.7927.
- [99] S. Pargaonkar, “A Comprehensive Research Analysis of Software Development Life Cycle (SDLC) Agile & *Waterfall* Model Advantages, Disadvantages, and Application Suitability in Software Quality Engineering,” *Int. J. Sci. Res. Publ.*, vol. 13, no. 8, pp. 120–124, 2023, doi: 10.29322/ijsrp.13.08.2023.p14015.
- [100] Z. N. Ghafar, “The evaluation research: A comparative analysis of qualitative and quantitative research methods,” *J. Lang. Lit. Soc. Cult. Stud.*, vol. 2, no. 1, pp. 1–10, 2024, doi: 10.58881/jllscs.v2i1.122.
- [101] M. H. Maltesdotter, “What was Killing Babies in Sundsvall ? A Study of Infant Mortality Patterns Using Individual Level,” vol. 15, pp. 1–27, 2025.
- [102] A. Kliber, B. Łęt, and P. Řezáč, “Fear or Price? Vulnerability of the Interest in Green Transport To Covid Dynamics and Fuel Prices in V4 Economies,” *Econ. Sociol.*, vol. 18, no. 1, pp. 70–89, 2025, doi: 10.14254/2071-789X.2025/18-1/4.
- [103] M. Keil, “Climbing the academic ladder: how social class and habitus mediate academic careers,” *Int. Stud. Sociol. Educ.*, vol. 00, no. 00, pp. 1–22, 2025, doi: 10.1080/09620214.2025.2494099.
- [104] C. Briggs, “the Popular Classes and Royal Justice in Medieval England:

Evidence From the Derbyshire Eyre of 1330-1331,” *Stud. Hist. Hist. Mediev.*, vol. 43, no. 1, pp. 11–34, 2025, doi: 10.14201/shhme.31922.

- [105] K. Łuczaj, “Navigating Recognition: The Symbolic Struggles in the Biographies of Young Polish Internet Content Creators,” *Qual. Sociol. Rev.*, vol. 21, no. 2, pp. 74–92, 2025, doi: 10.18778/1733-8077.21.2.05.
- [106] M. A. Pantea, “The Role of the ”Petru Maior” and ”România Jună” Societies in the Formation of the Romanian Elite,” *Plur. Hist. Cult. Soc.*, vol. 12, no. 1, pp. 23–40, 2024, doi: 10.37710/plural.v12i1_2.
- [107] J. Zhang and A. College, “Navigating Exclusion Across Space and Time : Chinese Seafarers in the United States , Norway , and Britain Before , During , and Beyond World War II,” pp. 138–144.