Dimasejati: Jurnal Pengabdian Kepada Masyarakat
ISSN 2716-3377 (print) | ISSN 2721-9364 (online)
https://syekhnurjati.ac.id/jurnal/index.php/dimasejati

OPEN ACCESS

Vol. 8, No. 2, 2026

DOI: 10.24235/dimasejati.51.000

Kritik Prompt dan Validasi Output dalam Pelatihan Internal AI Generatif: Evaluasi Deskriptif bagi Pendidik SMA

Saefurrohman[1], Jeffri Alfa Razaq[2], Purwatiningtyas[3], Felix Andreas Sutanto[4], Sulastri[5]

[1,2,3,4,5] *Universitas Stikubank (UNISBANK), Semarang, Indonesia*

A B S T R A C T

Generative AI use by educators is not always accompanied by structured prompt design and systematic checking of outputs before professional use. This community-engagement program describes an in-house training for educators and education staff at SMA Negeri 1 Kota Mungkid using a Write-Critique-Revise-Execute-Validate cycle with free ChatGPT accounts. Evaluation combined a post-training questionnaire, descriptive categorization of open responses, activity documentation, and an illustrative analysis of prompt transformation during guided practice. The analysis included 38 responses with complete data on 14 numerical items and consent for anonymous reporting. Item means ranged from 4.368 to 4.684 on a five-point scale; 97.4% of respondents rated 4-5 for the need to recheck AI outputs and protect school data, while 92.1% rated 4-5 for understanding prompt structure. Continuing needs remained visible because 22 respondents reported limited prompt-writing fluency, 18 remained concerned about data security, and 11 remained concerned about output accuracy. The findings represent post-training evaluation and perceived understanding rather than objective competency gains or causal effects.

Keywords: AI Literacy; Generative AI; Output Validation; Prompt Engineering; Teacher Professional Development

A B S T R A C T

Pemanfaatan AI generatif oleh pendidik dan tenaga kependidikan belum selalu diikuti kemampuan menyusun instruksi terstruktur dan memeriksa keluaran sebelum digunakan. Kegiatan pengabdian ini mendeskripsikan in-house training bagi PTK SMA Negeri 1 Kota Mungkid yang menerapkan siklus Tulis-Kritisi-Revisi-Eksekusi-Validasi menggunakan ChatGPT akun gratis. Evaluasi program menggunakan kuesioner pascapelatihan, kategorisasi deskriptif respons terbuka, dokumentasi kegiatan, dan analisis ilustratif transformasi prompt pada praktik terbimbing. Analisis mencakup 38 respons dengan data lengkap pada 14 butir numerik dan persetujuan pelaporan anonim. Rerata butir berada pada rentang 4,368-4,684 dari skala 1-5; 97,4% responden memberi skor 4-5 pada kebutuhan memeriksa ulang keluaran AI dan perlindungan data, sedangkan 92,1% memberi skor 4-5 pada pemahaman struktur prompt. Kebutuhan penguatan tetap terlihat karena 22 responden belum terbiasa menyusun prompt, 18 mengkhawatirkan keamanan data, dan 11 mengkhawatirkan akurasi keluaran. Hasil ini merepresentasikan evaluasi pascapelatihan dan pemahaman yang dirasakan, bukan bukti peningkatan kompetensi objektif atau efek kausal pelatihan secara langsung.

Keywords: AI Literacy; Generative AI; Output Validation; Prompt Engineering; Teacher Professional Development

1Corresponding Author: Fakultas Teknologi Informasi dan Industri, Universitas Stikubank (UNISBANK), Semarang, Indonesia; Email: saefur@edu.unisbank.ac.id

Received:	Revised:	Accepted:	Available online:
01.03.2026	01.04.2026	01.06.2026	14.08.2026

Suggested citation:
Saefurrohman, J. A. R., Purwatiningtyas, F. A. S., Sulastri (2026). *Prompt Critique and Output Validation in Generative AI In-House Training: A Descriptive Evaluation For High School Educators.Dimasejati: Jurnal Pengabdian Kepada Masyarakat, 8* (2), 199-213. DOI: 10.24235/dimasejati.51.000

PENDAHULUAN

Ketersediaan large language model (LLM) melalui antarmuka percakapan telah memperluas pemanfaatan AI generatif untuk penyusunan bahan ajar, asesmen, umpan balik, perencanaan pembelajaran, dan pekerjaan administratif. Kasneci et al. (2023) memetakan peluang pemanfaatan LLM dalam pendidikan sekaligus menekankan keterbatasan sistem, bias keluaran, kebutuhan pemeriksaan fakta, dan pengawasan manusia. Lo (2023) mengidentifikasi persoalan akurasi dan reliabilitas, terutama keterbatasan pengetahuan mutakhir serta kemungkinan munculnya informasi yang keliru atau palsu. Tinjauan sistematis pada konteks K-12 menempatkan pengembangan profesional guru sebagai kebutuhan penting ketika penggunaan AI generatif meluas ke perencanaan pembelajaran dan aktivitas kerja sekolah (Alfarwan, 2025).

Literasi AI tidak cukup dimaknai sebagai kemampuan mengakses aplikasi dan meminta sistem menghasilkan teks. Ng et al. (2024) menempatkan literasi AI pada dimensi afektif, perilaku, kognitif, dan etis, sehingga kemampuan menggunakan sistem perlu berjalan bersama evaluasi dan pertimbangan konsekuensi penggunaan. Kerangka kompetensi AI bagi guru UNESCO menekankan human agency, etika AI, fondasi dan aplikasi AI, pedagogi, serta pembelajaran profesional sebagai dimensi yang saling terkait (Miao & Cukurova, 2024). Posisi tersebut memberi dasar bahwa pelatihan PTK perlu menggabungkan keterampilan operasional dengan pemeriksaan keluaran, perlindungan data, dan tanggung jawab pedagogis.

Prompt engineering menjadi keterampilan relevan ketika pengguna mengarahkan LLM melalui bahasa alami. Cain (2024) menguraikan prompt engineering melalui tiga komponen penting, yaitu pengetahuan konten, berpikir kritis, dan desain iteratif. Walter (2024) menempatkan literasi AI, kecakapan prompt engineering, dan berpikir kritis sebagai kompetensi yang perlu dikembangkan bersama dalam pendidikan. Tinjauan Lee dan Palmer (2025) menemukan beragam framework prompt engineering dan menekankan keterampilan interaksi AI yang pragmatis, terstruktur, serta selaras dengan tujuan pendidikan yang kontekstual. Qian (2025) membedakan strategi prompting berbasis teknik dan berbasis proses; pendekatan berbasis proses diarahkan untuk mendukung keterlibatan kognitif dan pemikiran kolaboratif dengan AI. Sintesis tersebut mendukung pelatihan yang tidak berhenti pada formula prompt statis.

Program pengabdian di Indonesia telah mulai menjadikan prompt engineering sebagai materi pelatihan bagi guru dan peserta didik. Basuki et al. (2024) menggabungkan penyusunan prompt, pemanfaatan ChatGPT, dan etika LLM bagi guru sekolah Muhammadiyah serta menggunakan pre-test dan post-test sebagai

evaluasi. Ariati et al. (2026) menerapkan praktik prompt engineering untuk penyusunan RPP berbasis project-based learning pada guru SMK dengan evaluasi sebelum dan sesudah pelatihan. Parwati et al. (2026) menggunakan pelatihan AI dan prompt untuk konteks desain komunikasi visual pada siswa SMK. Kegiatan-kegiatan tersebut menguatkan relevansi prompt engineering dalam program pemberdayaan, tetapi perbedaan sasaran, desain, dan instrumen membuat besaran hasil antarkegiatan tidak layak dibandingkan secara langsung.

Penelitian teknis memberikan dasar untuk melampaui formula prompt statis. PE2 menggunakan meta-prompt yang memuat deskripsi tugas, spesifikasi konteks, dan prosedur penalaran untuk membantu model mengidentifikasi kelemahan serta melakukan penyuntingan prompt yang terarah (Ye et al., 2024). Self-Refine memperlihatkan bahwa umpan balik dan penyempurnaan berulang dapat memperbaiki keluaran pada tugas-tugas yang diuji dibandingkan generasi satu langkah (Madaan et al., 2023). Kemampuan model mengkritik dirinya sendiri tetap memiliki batas; Tyen et al. (2024) menunjukkan bahwa LLM dapat kesulitan menemukan lokasi kesalahan penalaran meskipun koreksi menjadi lebih baik setelah lokasi kesalahan diketahui. Landasan tersebut mendukung AI sebagai mitra kritik awal tanpa memindahkan kewajiban verifikasi faktual dan profesional dari pengguna manusia.

Pemetaan kebutuhan mitra di SMA Negeri 1 Kota Mungkid menunjukkan konteks yang tidak berangkat dari ketiadaan akses terhadap AI, melainkan dari kebutuhan menata cara menyusun instruksi, menjaga fokus konteks, menilai keluaran, dan melindungi data sekolah. Tinjauan kritis pengembangan profesional guru berbasis GenAI menemukan keterbatasan pelatihan prompt engineering yang eksplisit serta dominasi hasil berbasis laporan diri dibandingkan pengukuran performa (Shenoy & Saarela, 2026). Literatur juga telah memuat pendekatan Write-Curate-Verify yang mengarahkan pengguna untuk menulis prompt, mengkurasi keluaran, dan memverifikasi hasil pada pengembangan skenario pembelajaran (Bai et al., 2024). Siklus Tulis-Kritisi-Revisi-Eksekusi-Validasi pada program ini tidak diklaim sebagai penemuan konsep kurasi atau verifikasi baru. Perbedaannya berada pada penempatan prompt sebagai objek kritik dan revisi sebelum eksekusi, lalu validasi diarahkan pada hasil yang akan digunakan dalam pekerjaan PTK. Kontribusi artikel dibatasi pada dokumentasi implementasi siklus tersebut dalam IHT bagi PTK dan analisis evaluasi yang tersedia, bukan pada validasi framework baru atau pembuktian superioritas teknik prompting.

Tujuan artikel ini adalah mendeskripsikan implementasi siklus pelatihan, menganalisis evaluasi pascapelatihan dari perspektif responden, menelaah kebutuhan tindak lanjut, dan membahas transformasi prompt pada praktik terbimbing sesuai batas bukti yang tersedia. Fokus tersebut menempatkan keterlacakan data, ketepatan klaim, dan relevansi bagi praktik PTK sebagai dasar interpretasi hasil.

BAHAN DAN METODE

Desain kegiatan dan khalayak sasaran

Kegiatan menggunakan pendekatan pendidikan masyarakat melalui in-house training, demonstrasi, praktik terbimbing, dan evaluasi deskriptif. Khalayak sasaran

terdiri atas pendidik dan tenaga kependidikan SMA Negeri 1 Kota Mungkid, Kabupaten Magelang, Jawa Tengah. Unit analisis evaluasi dibatasi pada 38 respons yang memiliki data lengkap pada 14 butir numerik dan menyatakan persetujuan agar jawaban digunakan secara anonim dalam evaluasi dan laporan kegiatan. Komposisi responden evaluasi terdiri atas 37 guru dan satu pimpinan sekolah. Identitas personal tidak digunakan dalam analisis atau pelaporan. Daftar hadir dan dokumentasi kegiatan diperlakukan sebagai bukti keterlaksanaan, bukan sebagai ukuran perubahan kompetensi. Informasi mengenai pengalaman penggunaan AI sebelum pelatihan diperoleh secara retrospektif melalui butir pada kuesioner pascapelatihan dan digunakan hanya untuk menggambarkan profil 38 responden. Kegiatan tidak melaksanakan pre-test formal. Konsekuensi metodologisnya, artikel tidak menggunakan istilah peningkatan signifikan, efek intervensi, atau perubahan kausal. Hasil kuantitatif diperlakukan sebagai evaluasi pascapelatihan mengenai relevansi materi, pemahaman yang dirasakan, kepercayaan diri, validasi, etika, dan keamanan penggunaan AI.

Roadmap pelaksanaan

Roadmap program disajikan menggunakan waktu relatif sesuai urutan kegiatan pada proposal enam bulan. Bulan pertama diarahkan pada pemetaan kebutuhan dan konteks kerja mitra. Bulan pertama sampai kedua digunakan untuk penyusunan serta penyesuaian perangkat pelatihan. Bulan kedua memuat IHT inti. Bulan kedua sampai ketiga memuat praktik terbimbing dan pengembangan penerapan pada pekerjaan nyata. Bulan ketiga memuat evaluasi pascapelatihan. Bulan ketiga sampai keempat diarahkan pada pengembangan perangkat keberlanjutan seperti bank prompt dan checklist validasi. Bulan keempat sampai kelima digunakan untuk penyusunan luaran ilmiah. Bulan keenam ditempatkan sebagai ruang follow-up penggunaan perangkat oleh mitra. Waktu relatif tersebut berfungsi sebagai struktur roadmap program dan tidak digunakan untuk merekonstruksi tanggal pada dokumen pendukung. Artikel ini berfokus pada IHT inti, praktik terbimbing, dan evaluasi langsung yang memiliki bukti data.

Materi dan model praktik

Pelatihan menggunakan ChatGPT melalui akun gratis peserta sebagai platform praktik utama. Materi menempatkan AI sebagai mitra penyusunan draf dan eksplorasi awal, bukan sebagai sumber kebenaran. Struktur instruksi yang diajarkan memuat konteks, tujuan, peran, batasan, format, dan validasi. Materi lain mencakup perbandingan prompt langsung dan prompt yang lebih terstruktur, audit prompt, penyusunan materi ajar, diferensiasi pembelajaran, soal dan rubrik, audit soal atau materi, perlindungan data sensitif, serta pemeriksaan klaim yang memerlukan sumber.

Siklus Tulis-Kritisi-Revisi-Eksekusi-Validasi menjadi prosedur praktik utama. Tahap Tulis meminta peserta merumuskan kebutuhan kerja dengan bahasa yang digunakan sehari-hari. Tahap Kritisi menjadikan prompt sebagai objek audit sebelum jawaban substantif diminta. Tahap Revisi memperbaiki informasi yang kabur, terlalu luas, atau belum memiliki kriteria hasil. Tahap Eksekusi menjalankan prompt yang telah direkonstruksi. Tahap Validasi menempatkan pengguna sebagai pemeriksa

akurasi, relevansi, bias, kelayakan pedagogis, dan keamanan data. Permintaan audit dua sampai tiga siklus digunakan sebagai aturan praktis selama pelatihan dan tidak diposisikan sebagai jumlah iterasi yang terbukti optimal.

Instrumen dan analisis data

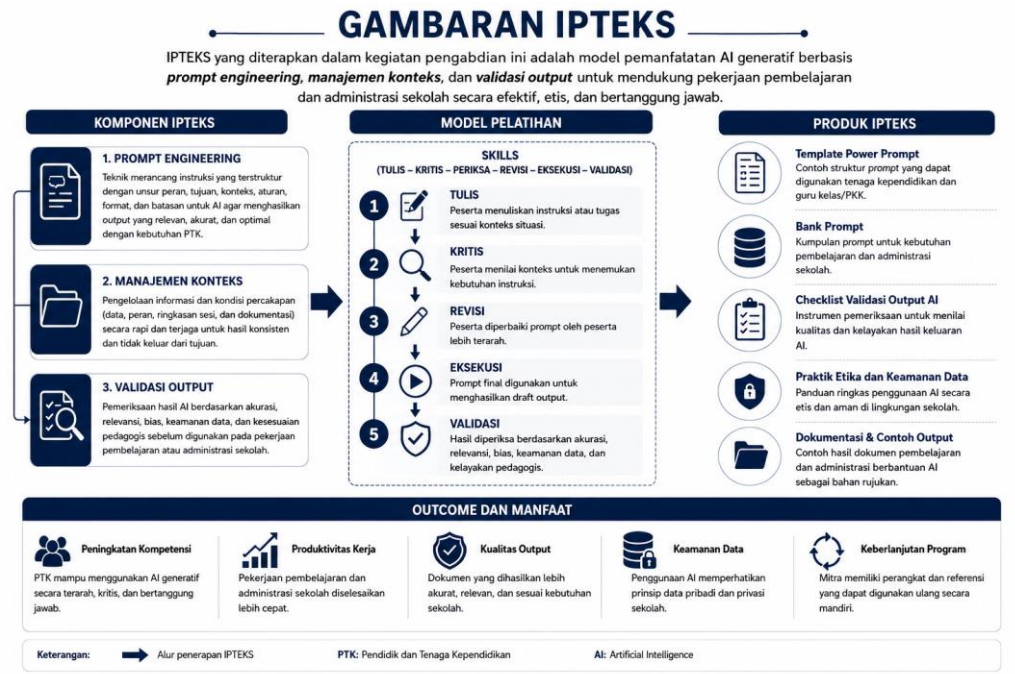
Kuesioner evaluasi pascapelatihan memuat 14 butir numerik skala 1-5, pertanyaan pilihan majemuk mengenai rencana penggunaan dan kendala, pilihan kebutuhan pelatihan lanjutan, pertanyaan terbuka, serta persetujuan penggunaan jawaban secara anonim. Instrumen disusun untuk evaluasi program dan tidak diperlakukan sebagai skala psikometrik tervalidasi. Analisis kuantitatif dilakukan pada tingkat butir dengan menghitung rerata, simpangan baku, frekuensi, dan proporsi respons bernilai 4-5. Penyajian proporsi skor 4-5 digunakan sebagai pendamping rerata agar interpretasi tidak bergantung pada satu statistik ringkas. Skor komposit, skor domain, dan uji inferensial tidak digunakan karena instrumen tidak dirancang sebagai skala konstruk dan kegiatan tidak memiliki baseline formal atau kelompok pembanding.

Tabel 1. Roadmap Pelaksanaan Program Berdasarkan Proposal

Tahap	Waktu Relatif	Kegiatan Utama	Status dalam Artikel
1	Bulan 1	Pemetaan kebutuhan dan konteks kerja mitra	Landasan program
2	Bulan 1–2	Penyusunan dan penyesuaian perangkat pelatihan	Landasan program
3	Bulan 2	IHT inti AI generatif dan *prompting*	Dilaporkan
4	Bulan 2–3	Praktik terbimbing dan penerapan kasus	Dilaporkan
5	Bulan 3	Kuesioner evaluasi pascapelatihan	Dianalisis
6	Bulan 3–4	Pengembangan bank *prompt* dan *checklist* validasi	Roadmap lanjutan
7	Bulan 4–5	Penyusunan luaran ilmiah dan dokumentasi	Roadmap luaran
8	Bulan 6	*Follow-up* pemanfaatan perangkat oleh mitra	Roadmap keberlanjutan

Respons terbuka dianalisis melalui kategorisasi isi deskriptif dengan menandai isu yang disebutkan secara eksplisit, terutama prompt atau prompting, kebutuhan praktik, variasi kasus, keamanan, dan kebutuhan lanjutan. Varian ejaan yang secara jelas merujuk pada prompt, seperti promt, promp, dan prom, dinormalisasi pada saat penghitungan frekuensi. Kategorisasi digunakan sebagai deskripsi isi dan tidak diklaim sebagai analisis tematik dengan validitas antarkoder. Artefak transformasi prompt dianalisis secara ilustratif pada dimensi spesifikasi tugas dan kontrol proses. Analisis artefak tidak digunakan sebagai skor pre-post karena contoh pascapenyampaian materi merupakan modifikasi terbimbing oleh fasilitator. Tabel 1

merangkum tahapan relatif pelaksanaan program dan status pelaporannya dalam artikel. Gambar 1 memvisualisasikan hubungan komponen IPTEKS, siklus praktik, dan sasaran program.



Gambar 1. Model IPTEKS, siklus praktik, dan sasaran program
Sumber: Rancangan program pengabdian; panel outcome merupakan sasaran program, bukan hasil kausal evaluasi

Gambar 1 menempatkan prompt engineering, manajemen konteks, dan validasi output sebagai komponen yang saling mendukung. Panel outcome pada bagian bawah gambar diperlakukan sebagai sasaran program yang direncanakan, bukan sebagai hasil kausal yang dibuktikan oleh desain evaluasi artikel ini.

HASIL DAN PEMBAHASAN

Profil responden dan pengalaman awal

Analisis kuantitatif mencakup 38 respons dengan data lengkap pada 14 butir numerik. Komposisi responden terdiri atas 37 guru dan satu pimpinan sekolah. Pengalaman penggunaan AI sebelum pelatihan yang dilaporkan secara retrospektif menunjukkan 13 responden sering menggunakan AI untuk pekerjaan, 12 responden cukup sering, 12 responden pernah mencoba tetapi belum rutin, dan satu responden belum pernah menggunakan AI. Profil tersebut menunjukkan bahwa sebagian besar responden kuesioner telah memiliki paparan terhadap AI sebelum pelatihan. Data ini tidak digunakan untuk menggambarkan seluruh peserta yang hadir dan tidak

berfungsi sebagai baseline kompetensi. Tabel 2 merangkum profil responden evaluasi dan pengalaman penggunaan AI yang dilaporkan secara retrospektif.

Tabel 2. Profil Responden Evaluasi dan Pengalaman Penggunaan AI

Karakteristik	Kategori	n (%)
Jabatan	Guru	37 (97,4)
Jabatan	Pimpinan sekolah	1 (2,6)
Pengalaman AI	Sering menggunakan untuk pekerjaan	13 (34,2)
Pengalaman AI	Cukup sering	12 (31,6)
Pengalaman AI	Pernah mencoba, belum rutin	12 (31,6)
Pengalaman AI	Belum pernah	1 (2,6)

Evaluasi pascapelatihan

Rerata 14 butir evaluasi berada pada rentang 4,368-4,684 pada skala 1-5, sedangkan simpangan baku per butir berada pada rentang 0,489-0,647. Proporsi responden yang memberikan skor 4-5 berada pada rentang 92,1-100,0%. Pola tersebut menunjukkan evaluasi pascapelatihan yang cenderung positif pada seluruh butir yang diukur. Interpretasi tetap dibatasi pada persepsi responden karena kuesioner tidak mengukur performa aktual dan tidak mempunyai baseline prospektif.

Indikator dengan rerata tertinggi adalah pemahaman pentingnya menjaga data siswa, nilai individual, identitas pribadi, dan dokumen internal sekolah dengan nilai 4,684. Relevansi contoh terhadap pekerjaan pembelajaran atau administrasi memperoleh rerata 4,632. Pemahaman bahwa hasil AI perlu diperiksa kembali memperoleh rerata 4,553, sedangkan pemahaman struktur prompt yang memuat tujuan, konteks, batasan, dan format keluaran memperoleh rerata 4,368. Nilai terakhir tetap berada pada sisi positif skala, tetapi menjadi item terendah di antara 14 butir sehingga pembiasaan menyusun prompt layak diprioritaskan pada tindak lanjut. Tabel 3 menyajikan statistik deskriptif pada keempat belas indikator evaluasi.

Tabel 3. Hasil Evaluasi Pascapelatihan Pada 14 Indikator (N = 38)

Pernyataan evaluasi (ringkas)	Mean	SD	Skor 4-5 (%)
Materi jelas dan mudah dipahami	4,500	0,604	94,7
Materi sesuai kebutuhan PTK	4,605	0,547	97,4
Penyampaian informatif dan runtut	4,526	0,506	100,0
Contoh relevan dengan pekerjaan sekolah	4,632	0,489	100,0
Materi menjawab kebutuhan informasi GenAI	4,447	0,602	94,7
Kepuasan terhadap kualitas pemateri	4,553	0,555	97,4

Pernyataan evaluasi (ringkas)	Mean	SD	Skor 4-5 (%)
Praktik prompting membuat penggunaan lebih konkret	4,447	0,602	94,7
Langkah praktik runtut dan mudah diikuti	4,500	0,647	92,1
Memahami tujuan, konteks, batasan, format	4,368	0,633	92,1
Praktik membantu solusi awal dokumen kerja	4,447	0,602	94,7
Lebih percaya diri menggunakan AI untuk pekerjaan	4,474	0,603	94,7
Memahami keluaran AI perlu diperiksa ulang	4,553	0,555	97,4
Memahami perlindungan data sekolah	4,684	0,525	97,4
Pelatihan membantu penggunaan AI lebih aman, etis, dan bertanggung jawab	4,526	0,557	97,4

Rencana Penerapan Dan Kebutuhan Tindak Lanjut

Rencana penggunaan AI pascapelatihan paling banyak mengarah pada penyusunan perangkat ajar, dipilih oleh 33 dari 38 responden. Pemanfaatan untuk asesmen atau soal dipilih oleh 23 responden, bahan presentasi 20 responden, kisi-kisi soal 17 responden, rubrik penilaian 15 responden, refleksi pembelajaran 14 responden, laporan kegiatan 13 responden, serta surat resmi atau notulen 10 responden. Pertanyaan ini memperbolehkan lebih dari satu pilihan. Satu responden menyatakan belum berencana menggunakan AI. Data tersebut menggambarkan intensi penerapan dan bukan bukti penggunaan aktual setelah kegiatan.

Kendala yang masih diperkirakan muncul memberi gambaran mengenai batas satu kali IHT. Butir kendala memperbolehkan lebih dari satu pilihan. Sebanyak 22 responden menyebut belum terbiasa menyusun prompt, 18 responden mengkhawatirkan keamanan data, 11 responden mengkhawatirkan akurasi hasil AI, dan sembilan responden masih memerlukan contoh kasus tambahan. Seluruh responden membuka peluang mengikuti pelatihan lanjutan; 22 menjawab bersedia dan 16 memilih kemungkinan mengikuti sesuai tema dan waktu. Butir pilihan mengenai bagian paling bermanfaat menempatkan penyampaian pemateri secara keseluruhan pada 16 responden, prompt engineering pada sembilan responden, dan praktik penyusunan prompt berbasis kasus pada delapan responden. Respons terbuka memberikan pola yang searah karena 22 dari 36 jawaban yang terisi menyebut prompt atau varian ejaan yang jelas merujuk pada prompt sebagai bagian yang membantu. Tabel 4 merangkum rencana penggunaan dan kendala yang masih diperkirakan peserta.

Tabel 4. Rencana Penggunaan dan Kendala yang Masih Diperkirakan Peserta (N = 38; Pilihan Majemuk)

Rencana penggunaan	n	Kendala	n
Perangkat ajar	33	Belum terbiasa menyusun prompt	22
Asesmen/soal	23	Khawatir keamanan data	18
Bahan presentasi	20	Khawatir akurasi hasil AI	11
Kisi-kisi soal	17	Perlu contoh kasus tambahan	9
Rubrik penilaian	15	Lainnya	2
Refleksi pembelajaran	14	Belum ada panduan internal	1
Laporan kegiatan	13	Keterbatasan internet/perangkat	1
Surat/notulen	10		

Transformasi Prompt Pada Praktik Terbimbing

Praktik awal meminta peserta menuliskan kebutuhan kerja menggunakan pola yang biasa mereka gunakan. Salah satu contoh meminta pembuatan modul ajar Sejarah kelas XII tentang rangkaian peristiwa Proklamasi 1945 dengan pendekatan pembelajaran mendalam dan pembobotan joyful, meaningful, serta mindful. Proporsi pembobotan pada contoh tersebut berasal dari formulasi peserta dan tidak diperlakukan dalam artikel sebagai ketentuan pedagogis atau regulasi resmi. Prompt awal sudah mengandung mata pelajaran, kelas, materi, dan pendekatan, sehingga tidak tepat dikategorikan sebagai prompt tanpa struktur.

Modifikasi terbimbing mengubah fokus dari permintaan langsung menjadi proses audit terhadap prompt. Fasilitator menambahkan instruksi agar AI mengkurasi dan menelaah kelemahan prompt, mengidentifikasi informasi yang hilang, menyusun ulang instruksi, lalu mengompilasi prompt final sebelum tugas dijalankan. Perubahan paling nyata berada pada kontrol proses, bukan sekadar pertambahan panjang teks. Permintaan agar hasil menjadi "valid" diperlakukan sebagai pemicu pemeriksaan dan bukan sebagai jaminan epistemik bahwa model telah membuktikan kebenaran isi. Tabel 5 membandingkan unsur prompt awal, modifikasi terbimbing, dan batas interpretasinya. Gambar 2 mendokumentasikan penyampaian materi, demonstrasi ChatGPT, dan praktik terbimbing selama IHT.

Tabel 5. Analisis Ilustratif Transformasi Prompt Pada Praktik Terbimbing

Dimensi	Prompt awal peserta	Modifikasi terbimbing	Interpretasi
Spesifikasi tugas	Modul Sejarah kelas XII; Proklamasi 1945; pendekatan pembelajaran mendalam	Spesifikasi awal dipertahankan sebagai isi tugas	Prompt awal sudah memiliki konteks dasar

Dimensi	Prompt awal peserta	Modifikasi terbimbing	Interpretasi
Kontrol proses	Permintaan langsung menghasilkan modul	Prompt diminta diaudit sebelum dieksekusi	Perubahan utama terjadi pada proses interaksi
Deteksi kekurangan	Tidak eksplisit	AI diminta menemukan kekurangan dan kesalahan instruksi	Prompt menjadi objek kritik
Penyempurnaan	Tidak eksplisit	Review 2-3 siklus sebagai aturan praktis pelatihan	Iterasi dipakai untuk pembiasaan, bukan jumlah optimum empiris
Rekonstruksi final	Tidak eksplisit	AI diminta menyusun ulang prompt utuh	Output antara berupa prompt yang telah direvisi
Validasi	Belum dinyatakan	Ada permintaan kesesuaian kriteria/regulasi	Tetap memerlukan sumber eksternal dan verifikasi manusia

Gambar 2. Penyampaian materi, demonstrasi ChatGPT, dan praktik terbimbing bersama PTK

Sumber: Dokumentasi kegiatan pengabdian

Pembahasan

Evaluasi positif terhadap struktur prompt dan praktik kasus sejalan dengan literatur yang menempatkan penggunaan, evaluasi, dan pertimbangan etis sebagai bagian dari literasi AI. Ng et al. (2024) menunjukkan bahwa literasi AI mencakup dimensi kognitif dan etis selain aspek afektif serta perilaku. Walter (2024) menempatkan prompt engineering dan berpikir kritis sebagai kemampuan yang perlu dikembangkan bersama dalam pendidikan. Proporsi 92,1% responden yang memberi skor 4-5 pada butir pemahaman tujuan, konteks, batasan, dan format menunjukkan evaluasi diri yang positif terhadap struktur instruksi yang diajarkan. Nilai tersebut tetap berasal dari laporan diri sehingga tidak dapat disamakan dengan uji performa penyusunan prompt. Pola program memiliki kedekatan dengan pelatihan prompt engineering bagi guru yang telah dilaporkan di Indonesia, tetapi desain evaluasi artikel ini berbeda. Basuki et al. (2024) menggunakan pre-test dan post-test untuk menilai pelatihan guru, sedangkan Ariati et al. (2026) juga memakai evaluasi sebelum dan sesudah kegiatan. Program ini tidak memiliki baseline formal sehingga persentase perubahan dan efektivitas kausal tidak dihitung. Posisi artikel lebih tepat sebagai evaluasi implementasi yang mendokumentasikan penerimaan, pemahaman yang dirasakan, kebutuhan lanjutan, dan karakter proses praktik.

Butir perlindungan data memperoleh rerata tertinggi sebesar 4,684, sedangkan butir pemahaman perlunya pemeriksaan ulang keluaran AI memperoleh rerata 4,553. Pola tersebut memberi sinyal bahwa responden menilai penggunaan AI yang bertanggung jawab sebagai bagian penting dari materi pelatihan. Kerangka UNESCO menempatkan human agency dan etika sebagai bagian dari kompetensi guru dalam ekosistem AI (Miao & Cukurova, 2024). Tinjauan sistematis pada pendidikan K-12 mengidentifikasi privasi, bias, transparansi, akuntabilitas, dan perlindungan peserta didik sebagai isu yang perlu ditangani dalam praktik sekolah (Gouseti et al., 2025). Zhu et al. (2025) juga mengelompokkan invasi privasi, kebocoran data, bias algoritmik, dan kesalahan algoritmik sebagai risiko pada dimensi teknologi. Keberadaan 18 responden yang masih mengkhawatirkan keamanan data menunjukkan bahwa penilaian positif tidak menghapus kebutuhan prosedur operasional dan kebijakan internal sekolah.

Siklus Kritisi-Revisi mempunyai korespondensi konseptual dengan prompting berbasis proses yang oleh Qian (2025) dibedakan dari strategi berbasis teknik. Bai et al. (2024) lebih dahulu menunjukkan proses Write-Curate-Verify untuk penulisan skenario, sehingga unsur kurasi dan verifikasi tidak dapat diposisikan sebagai kebaruan program ini. Pembeda implementatif yang didokumentasikan pada kegiatan ini terletak pada audit prompt sebelum jawaban substantif dihasilkan dan pemisahan tahap revisi prompt dari validasi keluaran. PE2 memperlihatkan bahwa model dapat diarahkan untuk menelaah prompt dan melakukan perubahan spesifik ketika meta-prompt menyediakan deskripsi, konteks, dan prosedur penalaran yang memadai (Ye et al., 2024). Self-Refine menunjukkan manfaat umpan balik dan penyempurnaan berulang pada tugas yang diuji (Madaan et al., 2023). Bukti teknis tersebut tidak memberi dasar untuk menetapkan dua atau tiga iterasi sebagai aturan universal. Penggunaan dua sampai tiga siklus dalam IHT dipertahankan sebagai aturan praktis yang mudah diingat, bukan parameter optimal yang telah dibuktikan.

Verifikasi oleh pengguna tetap penting karena LLM tidak selalu berhasil menemukan kesalahan pada proses penalarannya sendiri. Tyen et al. (2024)

menunjukkan perbedaan antara kemampuan menemukan kesalahan dan kemampuan memperbaiki kesalahan setelah lokasinya diketahui. Prinsip tersebut mendukung pemisahan antara audit prompt oleh AI dan validasi keluaran oleh pengguna. PTK tetap perlu memeriksa fakta, sumber, kesesuaian kurikulum, regulasi, dan keamanan data sebelum hasil digunakan dalam konteks profesional.

Kekuatan evidensi program perlu dibaca relatif terhadap intervensi yang menggunakan pengukuran prospektif. Woo et al. (2026), misalnya, mengumpulkan kuesioner sebelum dan sesudah workshop, kumpulan prompt sebelum dan sesudah workshop, serta refleksi tertulis untuk menilai intervensi prompt engineering. Moorhouse et al. (2024) menggunakan intervensi sebelas minggu dengan data kuantitatif dan kualitatif untuk menilai kompetensi GenAI calon guru. Perbedaan desain tersebut menegaskan bahwa evaluasi satu kali tanpa baseline pada program ini tidak menyediakan tingkat bukti yang setara. Nilai artikel ini berada pada dokumentasi implementasi, evaluasi pascapelatihan, dan identifikasi kebutuhan tindak lanjut, bukan pada estimasi besar efek pelatihan. Keberadaan 22 responden yang masih merasa belum terbiasa menyusun prompt menjadi dasar substantif untuk merancang tindak lanjut. Shenoy dan Saarela (2026) mengidentifikasi keterbatasan pelatihan prompt engineering, instruksi evaluasi kritis, dan integrasi ethical reasoning dalam pengembangan profesional guru. Moorhouse et al. (2024) melaporkan perkembangan pada seluruh aspek kompetensi GenAI profesional calon guru bahasa setelah intervensi 11 minggu, tetapi derajat perubahan berbeda antar-aspek. Program lanjutan perlu menggeser sebagian evaluasi dari persepsi menuju artefak kinerja berupa prompt independen, produk kerja, dan kemampuan memverifikasi keluaran.

Penggunaan ChatGPT akun gratis memberi keuntungan aksesibilitas karena peserta dapat mengulang praktik tanpa komitmen biaya langganan. Konfigurasi model pada layanan gratis tidak dikontrol secara identik pada setiap peserta, sehingga reproduksibilitas keluaran tidak diasumsikan tetap. Kebutuhan dokumentasi prompt, prosedur, dan kriteria validasi menjadi lebih penting daripada mengandalkan satu contoh respons sebagai representasi performa sistem. Keterbatasan artikel mencakup tidak tersedianya pre-test formal, tidak adanya kelompok pembanding, penggunaan laporan diri sebagai sumber utama evaluasi, serta analisis pada satu sekolah. Kuesioner disusun untuk evaluasi program dan belum divalidasi sebagai instrumen psikometrik. Pengisian dilakukan setelah IHT dan fasilitator merupakan bagian dari tim penulis, sehingga courtesy bias dan social desirability bias tidak dapat dikesampingkan. Kategorisasi respons terbuka bersifat deskriptif dan tidak menggunakan pengujian reliabilitas antarkoder. Artefak transformasi prompt bersifat ilustratif dan dipandu fasilitator sehingga tidak digunakan sebagai bukti peningkatan kemampuan individual. Pengukuran transfer jangka panjang belum tersedia, sedangkan model ChatGPT pada akun gratis tidak dikontrol pada versi yang identik. Batas-batas tersebut mengarahkan simpulan pada implementasi, penerimaan, pemahaman yang dirasakan, dan kebutuhan penguatan, bukan pada efek kausal atau superioritas pendekatan

SIMPULAN

Pelatihan penggunaan AI generatif bagi PTK SMA Negeri 1 Kota Mungkid diimplementasikan melalui siklus Tulis-Kritisi-Revisi-Eksekusi-Validasi dengan ChatGPT akun gratis sebagai media praktik. Evaluasi 38 responden memperlihatkan penilaian pascapelatihan yang cenderung positif; rerata per butir berada pada rentang 4,368-4,684, dan proporsi skor 4-5 berada pada rentang 92,1-100,0%. Kebutuhan penguatan tetap nyata karena 22 responden belum terbiasa menyusun prompt, sementara keamanan data dan akurasi keluaran masih disebut sebagai kendala. Transformasi prompt pada praktik terbimbing memperlihatkan perubahan dari permintaan langsung menuju audit, penyempurnaan, rekonstruksi, dan validasi, tetapi contoh tersebut bukan pengukuran pre-post. Literatur menunjukkan bahwa kurasi dan verifikasi telah digunakan dalam pendekatan prompting sebelumnya; kontribusi artikel ini berada pada dokumentasi penerapan audit prompt sebelum eksekusi serta kontekstualisasinya untuk pekerjaan PTK. Bukti yang tersedia memungkinkan siklus tersebut diposisikan sebagai struktur praktik awal, bukan sebagai framework tervalidasi atau teknik yang terbukti lebih unggul. Program lanjutan perlu menambahkan evaluasi berbasis performa, artefak prompt independen, verifikasi sumber, dan follow-up penggunaan agar kompetensi serta transfer praktik dapat dinilai lebih objektif.

Ucapan Terimakasih

Tim pengabdian menyampaikan terima kasih kepada pimpinan, guru, dan tenaga kependidikan SMA Negeri 1 Kota Mungkid atas partisipasi, ruang praktik, umpan balik, dan dukungan pelaksanaan kegiatan. Apresiasi juga disampaikan kepada Direktorat Penelitian, Pengabdian Masyarakat dan Publikasi Universitas Stikubank atas dukungan kelembagaan terhadap program.

REFERENSI

Alfarwan, A. (2025). Generative AI use in K-12 education: A systematic review. *Frontiers in Education, 10*, 1647573. doi:10.3389/feduc.2025.1647573

Ariati, N., Astuti, L. W., Dhamayanti, D., Antony, F., & Di Kesuma, H. (2026). Pelatihan prompt engineering dalam penyusunan RPP berbasis Project-Based Learning bagi guru SMK. *Reswara: Jurnal Pengabdian Kepada Masyarakat, 7(1)*, 281-289. doi:10.46576/rjpkm.v7i1.7811

Bai, S., Gonda, D. E., & Hew, K. F. (2024). Write-Curate-Verify: A case study of leveraging generative AI for scenario writing in scenario-based learning. *IEEE Transactions on Learning Technologies, 17*, 1313-1324. doi:10.1109/TLT.2024.3378306

Basuki, S., Faiqurrahman, M., Marthasari, G. I., Indrabayu, R., Zachra, F., & Effendy, N. A. (2024). Assistance in preparing engineering prompts for Muhammadiyah school teachers to optimize the use of ChatGPT in the world of education. *Dinamisia: Jurnal Pengabdian Kepada Masyarakat, 8(3)*, 721-729. doi:10.31849/dinamisia.v8i3.19522

Cain, W. (2024). Prompting change: Exploring prompt engineering in large language model AI and its potential to transform education. *TechTrends, 68(1)*, 47-57. doi:10.1007/s11528-023-00896-0

Gouseti, A., James, F., Fallin, L., & Burden, K. (2025). The ethics of using AI in K-12 education: A systematic literature review. *Technology, Pedagogy and Education, 34(2)*, 161-182. doi:10.1080/1475939X.2024.2428601

Kasneci, E., Sessler, K., Kuechemann, S., Bannert, M., Dementieva, D., Fischer, F., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. doi:10.1016/j.lindif.2023.102274

Lee, D., & Palmer, E. (2025). Prompt engineering in higher education: A systematic review to help inform curricula. *International Journal of Educational Technology in Higher Education, 22*, 7. doi:10.1186/s41239-025-00503-7

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13(4)*, 410. doi:10.3390/educsci13040410

Madaan, A., Tandon, N., Gupta, P., Hallinan, S., Gao, L., Wiegreffe, S., ... Clark, P. (2023). Self-Refine: Iterative refinement with self-feedback. *Advances in Neural Information Processing Systems, 36*, 46534-46594. doi:10.52202/075280-2019

Miao, F., & Cukurova, M. (2024). *AI competency framework for teachers*. UNESCO. doi:10.54675/ZJTE2084

Moorhouse, B. L., Wan, Y., Wu, C., Kohnke, L., Ho, T. Y., & Kwong, T. (2024). Developing language teachers' professional generative AI competence: An intervention study in an initial language teacher education course. *System, 125*, 103399. doi:10.1016/j.system.2024.103399

Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2024). Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology, 55(3)*, 1082-1104. doi:10.1111/bjet.13411

Parwati, H., Pane, C. F., Nur, A., Agustin, G. U., Riziq, H., Silpyana, P., ... Nissa, I. I. (2026). Pelatihan teknologi Artificial Intelligence dan prompt untuk peningkatan keterampilan siswa Desain Komunikasi Visual di SMKS Mulia Hati Insani. *Jurnal Pengabdian Masyarakat Terapan, 3(1)*, 44-54. doi:10.20884/1.jupiter.3.1.106

Qian, Y. (2025). Prompt engineering in education: A systematic review of approaches and educational applications. *Journal of Educational Computing Research, 63(7-8)*, 1782-1818. doi:10.1177/07356331251365189

Shenoy, P., & Saarela, M. (2026). A critical review and actionable framework for integrating generative AI into teacher professional development. *Discover Education, 5*, 570. doi:10.1007/s44217-026-01579-7

Tyen, G., Mansoor, H., Carbune, V., Chen, P., & Mak, T. (2024). LLMs cannot find reasoning errors, but can correct them given the error location. In *Findings of the*

Association for Computational Linguistics: ACL 2024 (pp. 13894-13908). Association for Computational Linguistics. doi:10.18653/v1/2024.findings-acl.826

Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education, 21*, 15. doi:10.1186/s41239-024-00448-3

Woo, D. J., Wang, D., Yung, T., & Guo, K. (2026). Effects of a prompt engineering intervention on undergraduate students' AI self-efficacy, AI knowledge and prompt engineering ability: A mixed methods study. *British Educational Research Journal, 52(2)*, 1442-1469. doi:10.1002/berj.70087

Ye, Q., Ahmed, M., Pryzant, R., & Khani, F. (2024). Prompt engineering a prompt engineer. In *Findings of the Association for Computational Linguistics: ACL 2024* (pp. 355-385). Association for Computational Linguistics. doi:10.18653/v1/2024.findings-acl.21

Zhu, H., Sun, Y., & Yang, J. (2025). Towards responsible artificial intelligence in education: A systematic review on identifying and mitigating ethical risks. *Humanities and Social Sciences Communications, 12*, 1111. doi:10.1057/s41599-025-05252-6