



PROMPTING AKADEMIK UNTUK PENGEMBANGAN PERANGKAT PEMBELAJARAN BERBASIS NILAI ISLAM DENGAN CHATGPT

Article history

Received: 14 Juli 2026

Revised: 15 Juli 2026

Accepted: 29 Juli 2026

DOI: [10.35329/jp.v6i3.7436](https://doi.org/10.35329/jp.v6i3.7436)

^{1*}Hamim, ¹Nur Hasaniyah, ¹Nur Arifuddin

¹Universitas Islam Negeri Maulana Malik Ibrahim.

**Corresponding author*

hamym.muchsin@uin-malang.ac.id

Abstrak

Observasi awal di Madrasah Diniyah Gunung Pandak menunjukkan bahwa guru mengenal kecerdasan buatan terutama sebagai sarana pencarian informasi umum, belum memahami struktur prompt, dan belum memiliki pedoman internal untuk memverifikasi keluaran AI. Program pengabdian ini bertujuan meningkatkan penguasaan prompting akademik dan kemampuan guru menghasilkan draf perangkat pembelajaran yang selaras dengan nilai pendidikan Islam. Kegiatan dilaksanakan pada 7 September-6 Oktober 2025 menggunakan Participatory Action Research melalui tahap to know, to understand, to plan, to action, dan to reflection. Peserta terdiri atas seluruh 17 guru aktif mata pelajaran keislaman. Penguasaan peserta diukur dengan pretest dan posttest paralel yang masing-masing memuat 25 butir. Validasi isi oleh satu validator menghasilkan rerata 3,97 dari skala 4 dan seluruh butir berada pada kategori sangat layak. Rata-rata nilai meningkat dari 55,06 menjadi 92,24 atau naik 37,18 poin. Uji t berpasangan menunjukkan perbedaan yang signifikan, $t(16) = 27,30$; $p < 0,001$, dengan rata-rata gain ternormalisasi 0,827 pada kategori tinggi. Selama praktik, peserta beralih dari perintah umum menuju prompt enam elemen, melakukan revisi keluaran, dan menghasilkan draf bahan ajar, metode, monitoring, asesmen, serta evaluasi. Refleksi peserta menunjukkan bertambahnya keberanian menggunakan AI dan persepsi manfaat bagi pekerjaan guru, meskipun dampak non-kognitif tersebut belum diukur dengan skala pretest-posttest. Program menghasilkan modul, contoh prompt, perangkat pembelajaran, serta komitmen tindak lanjut melalui komunikasi komunitas guru dan pengembangan pedoman penggunaan AI.

Kata kunci: *chatgpt; guru madrasah; kecerdasan buatan; pendidikan Islam; prompting akademik.*



Gambar 1. Guru Madrasah sebagai Peserta Pelatihan Prompting Akademik

1. PENDAHULUAN

Kebaruan program pengabdian ini terletak pada penyatuan empat komponen dalam satu alur kerja guru, yaitu struktur prompting akademik enam elemen, verifikasi keluaran berdasarkan kurikulum dan nilai Islam, pendampingan partisipatif, serta praktik penyusunan perangkat pembelajaran. Berbeda dari pelatihan ChatGPT yang umumnya berhenti pada pengenalan fitur atau latihan menghasilkan teks, program ini meminta guru merumuskan kebutuhan pedagogis, menyusun prompt, membaca keluaran, menguji ketepatan isi, memberi instruksi revisi, dan menetapkan kelayakan produk akhir. Kegiatan pengabdian tentang berbagai alat AI menunjukkan bahwa pendampingan tetap diperlukan ketika peserta menerapkan teknologi pada produk akademik yang nyata (Bakri et al., 2024). Pelatihan AI bagi guru juga lebih bermakna ketika materi dimulai dari kebutuhan kerja dan dilanjutkan dengan praktik pengembangan media pembelajaran (Fauziah et al., 2024). Pengalaman klinik guru digital menegaskan pentingnya ruang konsultasi dan latihan yang berhubungan langsung dengan masalah peserta (Askahar & Akbar, 2025). Dengan demikian, perbedaan program ini bukan terletak pada penggunaan ChatGPT sebagai aplikasi, melainkan pada integrasi prompting, verifikasi substansi, nilai Islam, dan produksi perangkat pembelajaran dalam satu siklus kerja akademik.

Madrasah Diniyah Gunung Pandak memiliki 17 guru aktif yang mengampu Fiqih, Hadis, Bahasa Arab, Akhlak, Al-Qur'an, dan Tajwid. Catatan observasi awal menunjukkan bahwa mayoritas guru berlatar pendidikan pesantren tradisional dan memiliki pengalaman teknologi yang terbatas. AI dipahami terutama sebagai sarana mencari informasi umum untuk kebutuhan sehari-hari. Guru belum mengenal istilah prompt, belum dapat membedakan fungsi peran, konteks, tujuan, fokus, format, dan gaya bahasa, serta belum mengetahui bahwa AI dapat membantu penyusunan bahan ajar, metode, monitoring, asesmen, dan evaluasi. Bentuk instruksi awal cenderung berupa perintah satu kalimat, seperti "buatkan materi Fiqih", tanpa menyebut kelas, tujuan pembelajaran, batasan isi, format, dan prosedur verifikasi. Lembaga juga belum memiliki pedoman internal mengenai perlindungan data, pemeriksaan sumber, dan tanggung jawab guru terhadap keluaran AI. Kondisi tersebut menunjukkan bahwa ketersediaan perangkat tidak otomatis menghasilkan pemanfaatan teknologi yang bermutu tanpa penguatan kompetensi dan pengendalian mutu kelembagaan (Gultom et al., 2022).

Kecerdasan buatan generatif dapat membantu guru menyiapkan draf penjelasan, aktivitas belajar, soal, rubrik, dan rancangan evaluasi. Model bahasa besar memiliki kemampuan dasar untuk menghasilkan dan menyesuaikan konten berdasarkan konteks yang diberikan pengguna (Li et al., 2024). Dalam pekerjaan guru, keluaran tersebut sebaiknya diperlakukan sebagai draf awal yang masih harus diperiksa dan dikembangkan kembali (Mustafa et al., 2024). Implementasi model bahasa besar di bidang pendidikan juga menghadapi persoalan akurasi, kesesuaian konteks, dan tata kelola penggunaan (Guizani et al., 2025). Perkembangan sistem multimodal memperluas kemungkinan penggabungan teks dan visual untuk mendukung penyajian materi yang lebih beragam (Xing et al., 2024). Potensi personalisasi tetap bergantung pada keputusan pedagogis guru dan kecukupan informasi tentang karakter peserta didik (Sharma et al., 2025).

Prompting akademik dalam program ini dipahami sebagai keterampilan menerjemahkan keputusan pembelajaran menjadi instruksi terstruktur. Interaksi guru dengan AI lebih bermakna ketika instruksi bertolak dari tujuan pedagogis, bukan sekadar pertanyaan umum (Xiao et al., 2025). Kejelasan instruksi membantu pengguna mengendalikan ruang lingkup, tingkat kedalaman, bentuk keluaran, dan pilihan bahasa (Leung, 2024). Atas dasar itu, program menggunakan enam elemen berupa peran, konteks, tujuan, fokus topik, format keluaran, dan gaya bahasa; kerangka semacam ini membantu pengguna menguraikan kebutuhan kompleks menjadi spesifikasi tugas yang lebih terarah (Hewing & Leinhos, 2024). Prompt tidak diperlakukan sebagai perintah sekali jadi. Guru membaca keluaran awal, mengenali kekurangan, lalu memberi instruksi lanjutan melalui contoh, pemecahan tugas bertahap, kombinasi pertanyaan terbuka dan tertutup, serta revisi

spesifik (Park & Choo, 2025). Keterampilan tersebut relevan bagi pengembangan profesional karena melatih guru merumuskan masalah, mengevaluasi informasi, dan memperbaiki hasil kerja, bukan hanya mengoperasikan satu aplikasi (Lee & Palmer, 2025).

Kemudahan menghasilkan teks membawa risiko kesalahan fakta, bias, rujukan yang tidak ditemukan, penyalinan mentah, serta penggunaan bahasa yang tidak sesuai dengan pembaca. Kajian mengenai AI dalam pendidikan menempatkan literasi kritis, transparansi, dan tanggung jawab pengguna sebagai bagian penting dari pemanfaatannya (Garzón et al., 2025). Dalam lingkungan akademik, penggunaan AI tanpa pemahaman dan pemeriksaan dapat mendorong pelanggaran kaidah ilmiah serta penyalinan keluaran secara mentah (Pramana et al., 2024). Prompt yang rinci dapat mempersempit ruang jawaban, tetapi tidak menghilangkan kemungkinan halusinasi atau informasi yang tidak dapat diverifikasi (Dannhoferová & Jedlička, 2024). Evaluasi penggunaan chatbot berbasis model bahasa pada konteks profesional juga menunjukkan perlunya penilaian manusia terhadap mutu dan kelayakan keluaran (Silvestri et al., 2024). Penyesuaian bahasa perlu memperhitungkan konteks sosial-budaya dan perkembangan penggunaan bahasa pada media digital, terutama ketika materi berkaitan dengan Bahasa Arab (Rashid, 2020). Oleh karena itu, program menempatkan pemeriksaan silang, perlindungan data, dan tanggung jawab guru sebagai prosedur kerja yang wajib dilakukan.

Tuntutan verifikasi menjadi lebih kuat pada mata pelajaran keislaman karena materi berkaitan dengan dalil, istilah, hukum, dan sumber yang memerlukan ketelitian. Pemanfaatan AI dalam pendidikan Islam tetap harus dikendalikan oleh tujuan pendidikan dan pertimbangan etik pendidik (Yuhana, 2024). ChatGPT dapat berfungsi sebagai asisten belajar, tetapi guru tetap menentukan validitas isi dan kesesuaiannya dengan konteks madrasah (Hasan et al., 2023). Implementasi AI pada pembelajaran pendidikan Islam juga menunjukkan perlunya pengawasan guru ketika teknologi digunakan untuk menyiapkan materi dan aktivitas pembelajaran (Marwa et al., 2024). Program penguatan guru berbasis pendidikan karakter memperlihatkan bahwa integrasi AI perlu diselaraskan dengan nilai yang menjadi dasar lembaga pendidikan (I. S. Kurniawan et al., 2025). Pemanfaatan ChatGPT untuk penyusunan modul ajar memberi manfaat ketika guru mampu menyusun instruksi yang tepat dan memeriksa hasil sebelum digunakan (S. Kurniawan et al., 2024). Dalam program ini, nilai tabayyun, amanah ilmiah, dan tanggung jawab guru diterapkan sebagai prosedur kerja. ChatGPT ditempatkan sebagai alat penyusun draf, sedangkan penentuan kebenaran isi dilakukan guru dengan merujuk kurikulum, buku ajar, kitab, dan sumber ilmiah yang relevan.

Participatory Action Research dipilih karena guru dan pengelola perlu terlibat dalam pemetaan masalah, pemilihan materi, praktik, refleksi, dan perumusan tindak lanjut. Pendekatan partisipatif memberi ruang bagi peserta untuk ikut menentukan masalah serta memperbaiki tindakan berdasarkan temuan selama kegiatan (Tok et al., 2024). Pengembangan profesional berbantuan teknologi lebih bermakna ketika materi berkaitan langsung dengan kebutuhan kerja peserta dan didukung kolaborasi yang berkelanjutan (Huang et al., 2024). Refleksi setelah praktik membantu peserta menilai proses belajarnya dan mengembangkan kemampuan belajar mandiri (Wang et al., 2024). Model pembelajaran berbasis desain juga menunjukkan bahwa integrasi AI oleh calon guru perlu dilatih melalui rangkaian perencanaan, penerapan, dan evaluasi, bukan hanya penyampaian konsep (Saimon et al., 2024).

Berdasarkan kondisi tersebut, kegiatan pengabdian bertujuan meningkatkan penguasaan guru dalam menyusun prompt akademik, menggunakan ChatGPT untuk menghasilkan draf perangkat pembelajaran, dan memeriksa keluarannya berdasarkan kurikulum serta nilai pendidikan Islam. Program juga diarahkan untuk menghasilkan modul, contoh prompt, perangkat pembelajaran, dan mekanisme tindak lanjut yang dapat direplikasi oleh madrasah lain.

2. METODE

Kegiatan pengabdian dilaksanakan di Madrasah Diniyah Gunung Pandak pada 7 September-6 Oktober 2025 menggunakan pendekatan Participatory Action Research. Peserta ditentukan melalui total sampling, yaitu melibatkan seluruh 17 guru aktif sebagai populasi sasaran program. Jumlah tersebut tidak ditetapkan melalui perhitungan sampel statistik karena kegiatan ditujukan untuk menjangkau seluruh guru pada satu lembaga. Keberagaman mata pelajaran yang diampu, meliputi Fiqih, Hadis, Bahasa Arab, Akhlak, Al-Qur'an, dan Tajwid, memungkinkan struktur prompting diterapkan pada kebutuhan perangkat yang berbeda. Satu pengelola yang merangkap sebagai pengasuh pondok dan kepala madrasah terlibat dalam koordinasi, penilaian produk, dan perumusan tindak lanjut, tetapi tidak dimasukkan sebagai peserta tes.

Data awal diperoleh melalui observasi, diskusi dengan pengelola dan guru, pemetaan pengalaman menggunakan AI, serta pretest. Observasi memusatkan perhatian pada pengalaman teknologi, pola instruksi yang digunakan, jenis pekerjaan akademik yang paling dibutuhkan, dan cara guru memeriksa keluaran. Selama praktik, setiap peserta menggunakan satu perangkat yang terdiri atas lima laptop dan 12 telepon pintar. Peserta diarahkan untuk tidak memasukkan identitas atau data pribadi siswa ke dalam ChatGPT.

Pelaksanaan mengikuti lima tahap PAR dengan empat pertemuan tatap muka. Setiap pertemuan terdiri atas dua sesi, pukul 08.00-09.45 dan 10.00-11.45, sehingga durasi tatap muka mencapai 210 menit per pertemuan atau 14 jam untuk empat pertemuan. Rincian waktu dan kegiatan disajikan pada Tabel 1.

Tabel 1. Tahapan, durasi, dan kegiatan Participatory Action Research

Tahap PAR	Waktu dan durasi	Kegiatan utama
To know	7-13 September 2025; 7 hari	Observasi awal, koordinasi dengan pengelola, pemetaan pengalaman guru, identifikasi kebutuhan perangkat pembelajaran, dan persiapan pretest.
To understand	14 September 2025; 210 menit	Pretest, diskusi terarah, identifikasi kesulitan menyusun instruksi, membedakan jenis perangkat, dan memeriksa keluaran AI.
To plan	15-20 September 2025; 6 hari	Penyusunan modul, kisi-kisi dan instrumen evaluasi, contoh prompt, skenario praktik, dan pembagian pendampingan.
To action	21 dan 28 September 2025; 420 menit	Pelatihan konsep, praktik bahan ajar, metode, monitoring, asesmen, evaluasi, verifikasi substansi, dan revisi prompt.
To reflection	3 Oktober 2025; 210 menit, dilanjutkan pengolahan data 4-6 Oktober	Posttest, penelaahan produk, refleksi pengalaman, identifikasi kendala, dan perumusan tindak lanjut.

Modul pelatihan memuat prinsip etis penggunaan AI, struktur prompt enam elemen, role prompting, pemberian contoh atau few-shot, pemecahan tugas bertahap, kombinasi pertanyaan terbuka dan tertutup, iterasi, dan pemilihan kata kunci. Praktik diarahkan pada lima kebutuhan guru, yaitu bahan ajar, metode pembelajaran, monitoring, asesmen, dan evaluasi. Fasilitator memberikan contoh perbandingan prompt umum dan prompt terstruktur, kemudian peserta mengembangkan prompt sesuai mata pelajaran masing-masing. Pelatihan yang menggabungkan penjelasan, praktik, pemeriksaan hasil, dan perbaikan produk memungkinkan peserta menghubungkan materi dengan pekerjaan akademik yang nyata (Bakri et al., 2024).

Instrumen utama berupa tes paralel Form A dan Form B yang masing-masing terdiri atas 25 butir, yaitu 20 pilihan ganda dan lima benar-salah. Kedua bentuk menggunakan

indikator dan tingkat kognitif yang sama, tetapi redaksinya berbeda. Materi tes mencakup etika dan peran AI, enam elemen prompt, teknik iterasi, serta penerapan ChatGPT untuk bahan ajar, metode, monitoring, asesmen, dan evaluasi. Jawaban benar diberi skor 1 dan jawaban salah atau kosong diberi skor 0; nilai akhir dihitung dari jumlah jawaban benar dikalikan empat.

Validasi isi dilakukan oleh satu validator terhadap 50 butir pada aspek relevansi, kejelasan, konstruksi, dan bahasa menggunakan skala 1-4. Rerata keseluruhan mencapai 3,97; rerata relevansi 4,00, kejelasan 3,86, konstruksi 4,00, dan bahasa 4,00. Sebanyak 43 butir memperoleh rerata 4,00 dan tujuh butir memperoleh rerata 3,75, sehingga seluruh butir dikategorikan sangat layak. Catatan validator digunakan untuk mengarahkan soal dan praktik pada pengembangan perangkat pembelajaran yang berhubungan langsung dengan pekerjaan guru. Kesepakatan antarvalidator tidak dihitung karena validasi hanya melibatkan satu ahli. Reliabilitas internal tidak digunakan sebagai dasar utama karena 17 peserta merupakan seluruh populasi program, bukan sampel uji psikometrik; pengujian reliabilitas pada sampel yang lebih besar direkomendasikan untuk implementasi berikutnya.

Data praktik diperoleh dari catatan observasi fasilitator, hasil prompt, draf perangkat pembelajaran, diskusi reflektif, dan testimoni peserta. Indikator non-kognitif dibaca secara kualitatif melalui keberanian mengoperasikan AI, kemauan merevisi prompt, rencana menerapkan hasil pelatihan, dan persepsi manfaat. Program tidak menggunakan skala non-kognitif pretest-posttest, sehingga temuan tersebut tidak dinyatakan sebagai peningkatan kuantitatif.

Data tes dianalisis menggunakan nilai rata-rata, simpangan baku, median, minimum, maksimum, selisih, dan normalized gain. Kategori N-Gain adalah tinggi untuk $g \geq 0,70$, sedang untuk $0,30 \leq g < 0,70$, dan rendah untuk $g < 0,30$. Normalitas selisih nilai diperiksa menggunakan Shapiro-Wilk. Perbedaan nilai dianalisis menggunakan uji t berpasangan pada taraf 0,05. Target kognitif dinyatakan tercapai apabila rata-rata posttest lebih tinggi daripada pretest dan seluruh peserta mencapai skor minimal 75. Alur program diringkas pada Gambar 2.



Siklus pelaksanaan program pengabdian berbasis Participatory Action Research

Gambar 2. Alur pelaksanaan program pengabdian berbasis Participatory Action Research

3. HASIL DAN PEMBAHASAN

3.1. Kondisi awal, pelaksanaan, dan respons lapangan

Pemetaan awal menegaskan bahwa persoalan utama mitra bukan ketersediaan perangkat atau internet, melainkan literasi penggunaan AI untuk pekerjaan akademik. Guru mengetahui AI sebagai sarana pencarian informasi umum, tetapi belum mengenal struktur prompt dan belum membedakan secara sistematis kebutuhan bahan ajar, metode, monitoring, asesmen, dan evaluasi. Instruksi awal cenderung generik dan tidak memuat jenjang, tujuan, karakter peserta didik, bentuk keluaran, atau batasan sumber. Pada mata pelajaran keislaman, peserta juga belum memiliki prosedur untuk menguji dalil dan penjelasan hukum yang dihasilkan AI.

Seluruh 17 peserta mengikuti pretest, praktik, dan posttest tanpa kehilangan data. Lima laptop dan 12 telepon pintar memungkinkan setiap peserta menjalankan ChatGPT secara individual. Kegiatan tidak mengalami gangguan internet yang menghambat praktik. Kondisi ini membuat pendampingan dapat difokuskan pada kualitas instruksi, ketepatan keluaran, dan proses revisi. Dokumentasi pelaksanaan ditampilkan pada Gambar 3.



Gambar 3. Pelaksanaan pelatihan dan praktik penyusunan prompt akademik

Kendala utama selama praktik bukan hanya pengoperasian aplikasi. Sebagian peserta masih kesulitan memahami istilah dasar kurikulum dan membedakan fungsi materi, metode, asesmen, dan evaluasi. Kesulitan tersebut membuat peserta belum dapat langsung merumuskan tujuan dan format keluaran yang spesifik. Tim pengabdian mengatasinya melalui penjelasan prasyarat sebelum praktik AI, demonstrasi perbandingan prompt umum dan prompt terstruktur, contoh yang berasal dari mata pelajaran peserta, latihan bertahap, serta diskusi sejawat. Strategi ini mencegah pelatihan berubah menjadi sekadar latihan mengetik perintah.

Perubahan praktik terlihat ketika peserta mulai mengganti perintah satu kalimat dengan prompt yang memuat peran, konteks, tujuan, fokus, format, dan gaya bahasa. Keluaran pertama tidak langsung diterima; peserta diminta menandai bagian yang terlalu luas, tidak sesuai usia siswa, tidak memiliki indikator, atau memuat informasi yang perlu diperiksa. Peserta kemudian menggunakan perintah revisi untuk menyederhanakan istilah, menambah kolom, memperjelas langkah, atau menghapus bagian yang tidak relevan. Perubahan ini menunjukkan pergeseran dari penggunaan AI sebagai mesin jawaban menuju penggunaan AI sebagai alat penyusun draf yang dikendalikan guru.

3.2. Perubahan penguasaan prompting akademik

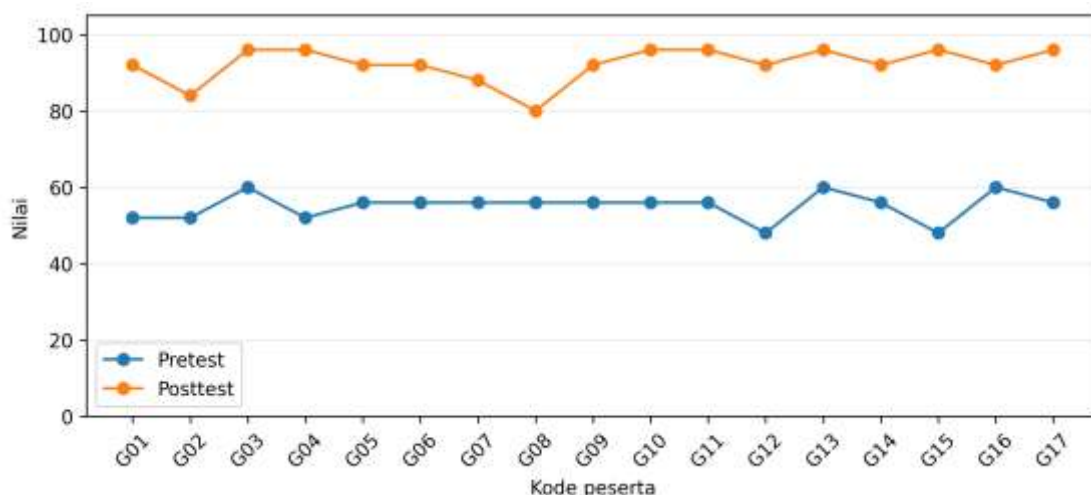
Hasil deskriptif menunjukkan perbedaan yang jelas antara nilai sebelum dan sesudah kegiatan. Rata-rata pretest sebesar 55,06 meningkat menjadi 92,24 pada posttest, atau bertambah 37,18 poin. Nilai minimum meningkat dari 48 menjadi 80, sedangkan nilai maksimum meningkat dari 60 menjadi 96. Statistik lengkap disajikan pada Tabel 2.

Tabel 2. Statistik hasil pretest dan posttest

Statistik	Pretest	Posttest
Jumlah peserta	17	17
Rata-rata	55,06	92,24

Statistik	Pretest	Posttest
Simpangan baku	3,61	4,58
Median	56,00	92,00
Nilai minimum	48,00	80,00
Nilai maksimum	60,00	96,00

Seluruh peserta memperoleh nilai posttest lebih tinggi daripada pretest. Kenaikan individual berada pada rentang 24-48 poin. G08 mengalami kenaikan terendah, dari 56 menjadi 80, sedangkan G15 mengalami kenaikan tertinggi, dari 48 menjadi 96. Pola tersebut menunjukkan bahwa peningkatan terjadi pada peserta dengan nilai awal rendah maupun relatif tinggi. Perbandingan individual ditampilkan pada Gambar 4.



Gambar 4. Perbandingan nilai pretest dan posttest setiap peserta

Uji Shapiro-Wilk pada selisih nilai menghasilkan $W = 0,945$ dan $p = 0,377$, sehingga asumsi normalitas terpenuhi. Uji t berpasangan menunjukkan perbedaan yang signifikan antara pretest dan posttest, $t(16) = 27,30$; $p < 0,001$. Interval kepercayaan 95% menempatkan kenaikan rata-rata pada rentang 34,29-40,06 poin. Hasil ini menunjukkan peningkatan penguasaan segera setelah pelatihan, tetapi tidak membuktikan bahwa program merupakan satu-satunya penyebab karena kegiatan tidak menggunakan kelompok pembandingan dan posttest diberikan segera setelah pelatihan.

Rata-rata N-Gain mencapai 0,827 dan termasuk kategori tinggi. Sebanyak 15 peserta atau 88,24% berada pada kategori tinggi, sedangkan dua peserta atau 11,76% berada pada kategori sedang. Tidak terdapat peserta dalam kategori rendah. Distribusi tersebut disajikan pada Tabel 3.

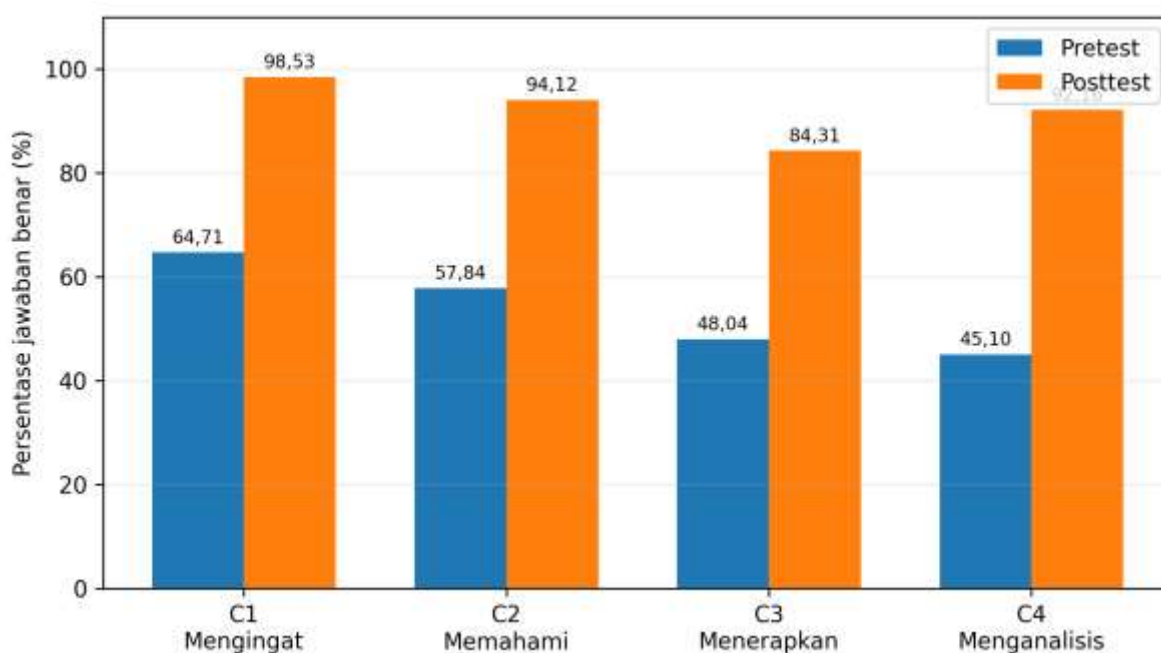
Tabel 3. Distribusi kategori N-Gain

Kategori	Kriteria	Jumlah peserta	Persentase
Tinggi	$g \geq 0,70$	15	88,24%
Sedang	$0,30 \leq g < 0,70$	2	11,76%
Rendah	$g < 0,30$	0	0,00%
Total	-	17	100,00%

Analisis berdasarkan tingkat kognitif menunjukkan bahwa capaian awal tertinggi berada pada C1 sebesar 64,71%, sedangkan capaian terendah berada pada C4 sebesar 45,10%. Setelah pelatihan, seluruh tingkat meningkat: C1 menjadi 98,53%, C2 menjadi 94,12%, C3 menjadi 84,31%, dan C4 menjadi 92,16%. Peningkatan terbesar terjadi pada C4 sebesar 47,06 poin persentase. Tabel 4 dan Gambar 5 menyajikan perubahan tersebut.

Tabel 4. Capaian berdasarkan tingkat kognitif

Tingkat kognitif	Pretest (%)	Posttest (%)	Kenaikan
C1-Mengingat	64,71	98,53	33,82
C2-Memahami	57,84	94,12	36,27
C3-Menerapkan	48,04	84,31	36,27
C4-Menganalisis	45,10	92,16	47,06



Gambar 5. Capaian pretest dan posttest berdasarkan tingkat kognitif

Kenaikan C4 berkaitan dengan latihan membedakan fungsi unsur prompt, memilih format keluaran, dan menilai kecocokan hasil dengan kebutuhan pembelajaran. Struktur enam elemen membantu peserta menguraikan kebutuhan yang kompleks menjadi instruksi yang lebih jelas (Hewing & Leinhos, 2024). Namun, capaian tes tidak identik dengan konsistensi menghasilkan perangkat bermutu dalam praktik sehari-hari. C3 tetap menjadi capaian posttest terendah, sehingga penerapan prompt pada situasi baru masih membutuhkan latihan yang bervariasi dan berulang. Strategi prompting untuk pendidik memang menekankan proses iteratif, yaitu membaca keluaran, menemukan kekurangan, dan menyusun perintah revisi yang lebih spesifik (Park & Choo, 2025). Kemampuan tersebut perlu diperlakukan sebagai bagian dari pengembangan profesional guru agar tidak berhenti pada penguasaan teknis aplikasi (Lee & Palmer, 2025).

3.3. Luaran program dan contoh produk

Program menghasilkan Modul Optimalisasi Prompting untuk Guru Madrasah, contoh prompt, dan draf perangkat pembelajaran. Modul memuat prinsip penggunaan AI sebagai asisten, kewajiban verifikasi untuk mencegah halusinasi, larangan menyalin keluaran secara mentah, integrasi nilai Islam, dan pemeriksaan silang dengan sumber kredibel.

Bagian teknis menjelaskan enam elemen prompt, role prompting, few-shot, step-by-step, kombinasi pertanyaan terbuka dan tertutup, iterasi, serta penerapan pada lima bidang kerja guru.

Contoh konkret dalam modul membandingkan perintah generik “buatkan materi tentang salat” dengan prompt yang menentukan peran guru, jenjang, tujuan, durasi, metode, dan format keluaran. Salah satu bentuk prompt enam elemen yang dikembangkan untuk materi thaharah diringkas pada Tabel 5.

Tabel 5. Contoh prompt akademik enam elemen

Elemen	Isi prompt
Peran	Bertindaklah sebagai guru Fiqih kelas 2 Madrasah Ibtidaiyah.
Konteks	Materi digunakan untuk siswa usia 8-10 tahun dengan kemampuan membaca beragam dan alokasi waktu 60 menit.
Tujuan	Susun draf bahan ajar yang membantu siswa memahami pengertian, rukun, dan urutan wudu.
Fokus	Batasi pembahasan pada rukun wudu dan kesalahan umum. Jangan mencantumkan dalil atau sumber yang tidak dapat diverifikasi.
Format keluaran	Sajikan dalam tabel yang memuat tujuan, penjelasan singkat, aktivitas siswa, dan tiga pertanyaan formatif.
Gaya bahasa	Gunakan bahasa Indonesia sederhana, kalimat pendek, religius, dan sesuai untuk siswa kelas rendah.

Prompt tersebut tidak diperlakukan sebagai produk akhir. Guru memeriksa urutan dan penjelasan rukun wudu menggunakan kurikulum, buku ajar, dan sumber keislaman yang digunakan madrasah. Jika hasil terlalu panjang atau tidak sesuai usia siswa, guru memberi instruksi lanjutan untuk menyederhanakan bahasa, mengubah format, atau memperjelas aktivitas.

Produk praktik mencakup lima kelompok sebagaimana disajikan pada Tabel 6. Contoh yang lebih konkret adalah instrumen monitoring pembelajaran Fiqih yang memuat indikator kesiapan dan keteladanan guru, partisipasi dan sikap siswa, proses pembelajaran, serta kolom refleksi guru. Pada bidang asesmen, modul menampilkan rancangan 20 soal pilihan ganda thaharah dengan sebaran C1-C4, kunci jawaban, dan panduan rubrik. Contoh tersebut memberi gambaran bahwa keluaran program tidak hanya berupa daftar prompt, tetapi draf perangkat yang dapat diperiksa dan disesuaikan sebelum digunakan.

Tabel 6. Produk praktik prompting akademik

Jenis produk	Fokus praktik	Contoh penggunaan
Bahan ajar	Materi sesuai jenjang, tujuan, dan nilai Islam	Draf materi thaharah dan aktivitas kognitif, afektif, serta psikomotorik
Metode pembelajaran	Strategi berdasarkan karakter siswa dan materi	Perbandingan demonstrasi, studi kasus, dan jigsaw beserta langkah penerapan
Monitoring	Indikator guru, siswa, proses, dan refleksi	Instrumen observasi dengan skala 1-4 dan catatan tindak lanjut
Asesmen	Soal, indikator, tingkat kognitif, kunci, dan rubrik	Draf asesmen sumatif 20 soal C1-C4
Evaluasi	Analisis hasil dan rekomendasi perbaikan	Rencana perbaikan materi, metode, dan pendampingan siswa

Pemeriksaan substansi menjadi bagian utama pada setiap produk. Peserta diarahkan untuk tidak menyalin keluaran yang memuat ayat, hadis, istilah fikih, atau penjelasan hukum tanpa membandingkannya dengan sumber yang digunakan lembaga. Prinsip tabayyun juga diterapkan dengan memeriksa rujukan yang tidak dapat ditemukan dan penjelasan yang tampak meyakinkan tetapi belum tentu tepat. Penggunaan ChatGPT sebagai asisten pendidikan Islam tetap menempatkan guru sebagai penanggung jawab akhir atas validitas dan kelayakan materi (Hasan et al., 2023). Praktik penyusunan modul dengan bantuan ChatGPT juga menuntut instruksi yang jelas, pemeriksaan isi, dan penyesuaian sebelum produk digunakan dalam pembelajaran.

3.4. Dampak non-kognitif dan refleksi pengalaman lapangan

Catatan fasilitator dan refleksi peserta menunjukkan tiga dampak non-kognitif yang bersifat kualitatif. Pertama, kepercayaan diri tampak dari keberanian peserta menjalankan ChatGPT pada perangkat masing-masing dan mencoba instruksi lanjutan ketika keluaran belum sesuai. Kedua, tingkat penerapan selama pelatihan terlihat dari keterlibatan seluruh peserta dalam penyusunan prompt dan draf perangkat untuk lima kebutuhan pembelajaran. Ketiga, peserta memandang program bermanfaat karena membantu mengarahkan substansi pendidikan pesantren agar lebih sesuai dengan perkembangan zaman dan memberi bekal untuk belajar mandiri mengembangkan kebutuhan madrasah.

Temuan tersebut perlu dibaca secara terbatas. Program tidak menggunakan skala kepercayaan diri atau persepsi manfaat sebelum dan sesudah kegiatan, sehingga tidak tersedia besaran perubahan non-kognitif. Bukti yang ada menunjukkan respons dan perilaku selama pelatihan, bukan keberlanjutan penggunaan dalam kelas. Pengukuran lanjutan perlu menggunakan skala terstruktur, log penggunaan prompt, observasi kelas, dan penilaian produk beberapa minggu setelah kegiatan.

3.5. Keberlanjutan dan komponen yang dapat direplikasi

Mitra menyatakan komitmen untuk melanjutkan komunikasi dan berbagi hasil melalui grup WhatsApp komunitas guru. Komitmen tersebut menjadi dasar tindak lanjut, tetapi belum cukup untuk menjamin penggunaan rutin. Secara operasional, madrasah perlu menunjuk koordinator, menyelenggarakan klinik AI atau klinik perangkat pembelajaran secara berkala, meminta setiap guru membawa satu produk yang akan digunakan, dan menyimpan prompt serta produk yang telah diverifikasi dalam repositori internal. Setiap berkas perlu mencantumkan mata pelajaran, kelas, tujuan, tanggal revisi, sumber pemeriksaan, dan penanggung jawab verifikasi.

Pedoman internal juga perlu mengatur larangan memasukkan data pribadi, kewajiban memeriksa sumber, penyesuaian dengan kurikulum, pencantuman penggunaan AI bila relevan, dan tanggung jawab guru terhadap produk akhir. Klinik dan repositori sebaiknya dievaluasi setelah satu semester melalui penelaahan produk, wawancara guru, dan tes retensi. Gambar 6 merangkum model implementasi program dan membedakan hasil yang telah terlaksana dari rencana tindak lanjut.

Komponen program yang dapat direplikasi oleh sekolah atau madrasah lain disajikan pada Tabel 7. Komponen tersebut bukan paket yang harus diterapkan secara identik; materi dan contoh perlu disesuaikan dengan karakter guru, mata pelajaran, sumber rujukan, dan kebijakan lembaga



Gambar 6. Model implementasi prompting akademik berbasis nilai Islam

Tabel 7. Komponen inti program untuk replikasi

Komponen	Isi minimum	Indikator pelaksanaan
Pemetaan kebutuhan	Pengalaman AI, jenis tugas guru, risiko, dan sumber rujukan	Masalah dan prioritas mitra terdokumentasi sebelum pelatihan
Modul inti	Etika, enam elemen prompt, iterasi, dan verifikasi	Peserta memiliki panduan dan contoh kontekstual
Praktik individual	Bahan ajar, metode, monitoring, asesmen, dan evaluasi	Setiap peserta mengoperasikan perangkat dan menghasilkan draf
Pendampingan partisipatif	Demonstrasi, latihan bertahap, diskusi sejawat, dan revisi	Kendala peserta dibahas menggunakan produk yang nyata
Pengendalian mutu	Kurikulum, sumber keislaman, bahasa, privasi, dan tanggung jawab	Keluaran diperiksa sebelum digunakan
Evaluasi dan keberlanjutan	Tes, penilaian produk, refleksi, klinik, repositori, dan SOP	Tersedia bukti hasil segera dan rencana pemantauan lanjutan

3.6. Keterbatasan program

Program melibatkan 17 guru dari satu lembaga sehingga hasil tidak dapat digeneralisasikan ke seluruh madrasah. Implementasi pada madrasah dengan jenjang, fasilitas, budaya organisasi, dan tingkat literasi digital yang berbeda diperlukan untuk menguji keberterimaan model. Posttest diberikan segera setelah pelatihan dan tidak terdapat kelompok pembanding, sehingga hasil hanya menunjukkan perubahan yang mengikuti kegiatan, bukan bukti kausal atau retensi jangka panjang.

Instrumen Form A dan Form B disusun berdasarkan kisi-kisi paralel, tetapi kesetaraan tingkat kesulitannya belum diuji pada sampel terpisah. Validasi hanya melibatkan satu ahli,

sehingga kesepakatan antarvalidator tidak tersedia. Dampak non-kognitif didasarkan pada observasi dan refleksi, bukan skala kuantitatif. Skor rubrik individual untuk seluruh perangkat juga belum tersedia, sehingga peningkatan tes tidak boleh disamakan dengan mutu semua produk. Penelitian pengabdian berikutnya perlu melibatkan beberapa validator, sampel lintas lembaga, tes retensi, skala non-kognitif, observasi penerapan di kelas, dan penilaian produk secara sistematis.

4. SIMPULAN

Program pengabdian meningkatkan penguasaan prompting akademik pada seluruh 17 guru Madrasah Diniyah Gunung Pandak. Rata-rata nilai meningkat dari 55,06 menjadi 92,24, dengan kenaikan 37,18 poin; uji t berpasangan menunjukkan perbedaan yang signifikan, $t(16) = 27,30$; $p < 0,001$. Rata-rata N-Gain sebesar 0,827 berada pada kategori tinggi, dengan 15 peserta pada kategori tinggi dan dua peserta pada kategori sedang. Peningkatan terbesar terjadi pada kemampuan menganalisis, sedangkan kemampuan menerapkan prompt tetap memerlukan latihan berulang.

Dampak program tidak hanya terlihat pada skor tes. Selama praktik, peserta beralih dari prompt generik menuju prompt enam elemen, melakukan verifikasi dan revisi, serta menghasilkan draf bahan ajar, metode, monitoring, asesmen, dan evaluasi. Modul dan contoh produk memberi bentuk konkret pada luaran pengabdian. Refleksi peserta menunjukkan bertambahnya keberanian menggunakan AI dan persepsi manfaat, tetapi dampak non-kognitif dan penggunaan rutin belum diukur secara kuantitatif atau melalui observasi kelas.

Keberlanjutan program perlu dibangun melalui grup komunikasi guru, klinik AI, repositori prompt terverifikasi, penunjukan koordinator, dan SOP penggunaan AI. Hasil program masih terbatas pada satu lembaga dan hasil segera setelah pelatihan. Implementasi lintas madrasah, tes retensi, penilaian produk, dan pemantauan praktik kelas diperlukan untuk menguji keberterimaan serta ketahanan model.

UCAPAN TERIMA KASIH

Tim pengabdian menyampaikan terima kasih kepada pengasuh sekaligus kepala Madrasah Diniyah Gunung Pandak yang telah memberikan izin, mendukung pemetaan kebutuhan, serta terlibat dalam penilaian perangkat pembelajaran dan perumusan tindak lanjut. Penghargaan juga disampaikan kepada 17 guru peserta yang mengikuti seluruh rangkaian kegiatan, mulai dari observasi awal, pretest, pelatihan, praktik menggunakan ChatGPT, refleksi, hingga posttest. Terima kasih disampaikan kepada Universitas Islam Negeri Maulana Malik Ibrahim Malang atas dukungan akademik selama penyusunan modul, pelaksanaan kegiatan, pengolahan data, dan penulisan artikel ini.

DAFTAR PUSTAKA

- Askahar, A., & Akbar, M. (2025). Klinik Guru Digital: Pendekatan Inovatif Untuk Transformasi Kompetensi Guru Dalam Pembelajaran Abad 21. *Jurnal Inovasi Penelitian Dan Pengabdian Masyarakat*, 5(1), 67–76. <https://doi.org/10.53621/jippmas.v5i1.490>
- Bakri, R., Hasbiyadi, H., & Astuti, N. P. (2024). Pelatihan dan Pendampingan Penyusunan Tugas Akhir menggunakan Berbagai Tools AI. *CARADDE: Jurnal Pengabdian Kepada Masyarakat*, 7(1), 84–88. <https://doi.org/10.31960/caradde.v7i1.2519>
- Dannhoferová, J., & Jedlička, P. (2024). Protect Yourself From Ai Hallucinations: Exploring Origins and Best Practices. 26th Annual International Conference Economic Competitiveness and Sustainability 2024 Proceedings, 49–59. <https://doi.org/10.11118/978-80-7509-990-7-0049>
- Fauziah, A. N. M., Widodo, W., Hidayati, S. N., Aulia, E. V., Permata Sari, D., & Zahro, F. (2024). Optimalisasi Kurikulum Merdeka melalui Pelatihan Guru untuk Mengimplementasikan Teknologi AI dalam Pengembangan Media Pembelajaran.

- Sasambo: Jurnal Abdimas (Journal of Community Service), 6(4), 702–714. <https://doi.org/10.36312/sasambo.v6i4.2206>
- Garzón, J., Patiño, E., & Marulanda, C. (2025). Systematic Review of Artificial Intelligence in Education: Trends, Benefits, and Challenges. *Multimodal Technologies and Interaction*, 9(8), 84. <https://doi.org/10.3390/mti9080084>
- Guizani, S., Mazhar, T., Shahzad, T., Ahmad, W., Bibi, A., & Hamam, H. (2025). A systematic literature review to implement large language model in higher education: Issues and solutions. *Discover Education*, 4(1), 35. <https://doi.org/10.1007/s44217-025-00424-7>
- Gultom, R., Nababan, M., Sihombing, G., Manulang, E., & Purba, E. (2022). Analisis Implementasi Manajemen Mutu Di Sekolah Menengah Teologi Kristen. *Jurnal Dinamika Pendidikan*, 15, 184–197. <https://doi.org/10.51212/jdp.v15i3.152>
- Hasan, M. ferdi, Patty, A., Sani, A., & Taufiq, F. (2023). Chat GPT as A Learning Assistant for Islamic Education: A Study In Madrasah Ibtidaiyah. *Nadwa: Jurnal Pendidikan Islam*, 17(2), 203–230. (2023). <https://doi.org/10.21580/nw.2023.17.2.26167>
- Hewing, M., & Leinhos, V. (2024). The Prompt Canvas: A Literature-Based Practitioner Guide for Creating Effective Prompts in Large Language Models (arXiv:2412.05127). arXiv. <https://doi.org/10.48550/arXiv.2412.05127>
- Huang, L., Liang, M., Xiong, Y., Wu, X., & Lim, C. P. (2024). A systematic review of technology-enabled teacher professional development during COVID -19 pandemic. *Computers & Education*, 223, 105168. <https://doi.org/10.1016/j.compedu.2024.105168>
- Kurniawan, I. S., Wahidah, U., Lada, S. B., Nisa, A. F., & Megarani, S. (2025). Integration of Artificial Intelligence in Character-Based Education: Strengthening Teachers with the Tamansiswa Approach in Indonesian Schools Kota Kinabalu. *International Journal Of Community Service*, 5(1), 57–65. <https://doi.org/10.51601/ijcs.v5i1.267>
- Kurniawan, S., Herlambang, S., Sari, N., Fadian, F., Suratman, B., Nurhidayah, V. A., Naffati, A. K., & Torikoh. (2024). Making Peace with Change: The Effect of GPT Chat Utilization on the Performance of Islamic Religion Teachers in Creating Teaching Modules. *Jurnal Pendidikan Agama Islam*, 21(2), 492–509. <https://doi.org/10.14421/jpai.v21i2.9767>
- 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(1), 7. <https://doi.org/10.1186/s41239-025-00503-7>
- Leung, C. H. (2024). Promoting optimal learning with ChatGPT: A comprehensive exploration of prompt engineering in education. *Asian Journal of Contemporary Education*, 8(2), 104–114. <https://doi.org/10.55493/5052.v8i2.5101>
- Li, Q., Fu, L., Zhang, W., Chen, X., Yu, J., Xia, W., Zhang, W., Tang, R., & Yu, Y. (2024). Adapting Large Language Models for Education: Foundational Capabilities, Potentials, and Challenges (arXiv:2401.08664). arXiv. <https://doi.org/10.48550/arXiv.2401.08664>
- Marwa, N. K., Anwar, S., & Firmansyah, M. I. (2024). The Implementation of Artificial Intelligence in Islamic Education Learning at Elementary School. *Edu Cendikia: Jurnal Ilmiah Kependidikan*, 4(03), 1676–1686. <https://doi.org/10.47709/educendikia.v4i03.5413>
- Mustafa, M. Y., Tlili, A., Lampropoulos, G., Huang, R., Jandrić, P., Zhao, J., Salha, S., Xu, L., Panda, S., Kinshuk, López-Pernas, S., & Saqr, M. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): A roadmap to a future research agenda. *Smart Learning Environments*, 11(1), 59. <https://doi.org/10.1186/s40561-024-00350-5>
- Park, J., & Choo, S. (2025). Generative AI Prompt Engineering for Educators: Practical Strategies. *Journal of Special Education Technology*, 40(3), 411–417. <https://doi.org/10.1177/01626434241298954>
- Pramana, P., Utari, P., Rochimah, H. A. I. N., Setianingrum, V. M., Purwati, E., & Aziz, R. A. (2024). Mitigasi Pelanggaran Etik: Lokakarya Penguatan Kaidah Ilmiah Era

- Artificial Intelligence pada Mahasiswa Universitas Muhammadiyah Ponorogo. *CARADDE: Jurnal Pengabdian Kepada Masyarakat*, 7(1), 158–168. <https://doi.org/10.31960/caradde.v7i1.2517>
- Rashid, H. S. (2020). The Arabic language in social medias' era. *Utopia y Praxis Latinoamericana*, 25(Extra1), Article Extra1. <https://doi.org/10.5281/zenodo.3784811>
- Saimon, M., Mtenzi, F., Lavicza, Z., Fenyvesi, K., Arnold, M., & Diego-Mantecón, J. (2024). Applying the 6E learning by design model to support student teachers to integrate artificial intelligence applications in their classroom. *Education and Information Technologies*, 1–18. <https://doi.org/https://doi.org/10.1007/s10639-024-12795-9>
- Sharma, S., Mittal, P., Kumar, M., & Bhardwaj, V. (2025). The role of large language models in personalized learning: A systematic review of educational impact. *Discover Sustainability*, 6(1), 243. <https://doi.org/10.1007/s43621-025-01094-z>
- Silvestri, C., Roshal, J., Shah, M., Widmann, W. D., Townsend, C., Brian, R., L'Huillier, J. C., Navarro, S. M., Lund, S., & Sathe, T. S. (2024). Evaluation of a novel large language model (LLM)-powered chatbot for oral boards scenarios. *Global Surgical Education - Journal of the Association for Surgical Education*, 3(1), 112. <https://doi.org/10.1007/s44186-024-00303-z>
- Tok, Ş., Dolapçioğlu, S., & Acar, S. P. (2024). Participatory Action Research for Professional Development / Mesleki Gelişim İçin Katılımcı Eylem Araştırması. *E-Uluslararası Eğitim Araştırmaları Dergisi*, 15(2), 145–161. <https://doi.org/10.19160/e-ijer.1590067>
- Wang, C., Li, Z., & Bonk, C. (2024). Understanding self-directed learning in AI-Assisted writing: A mixed methods study of postsecondary learners. *Computers and Education: Artificial Intelligence*, 6, 100247. <https://doi.org/10.1016/j.caeai.2024.100247>
- Xiao, R., Hou, X., Ye, R., Kazemitabaar, M., Diana, N., Liut, M., & Stamper, J. (2025). Improving Student-AI Interaction Through Pedagogical Prompting: An Example in Computer Science Education (arXiv:2506.19107). <https://doi.org/10.48550/arXiv.2506.19107>
- Xing, W., Zhu, T., Wang, J., & Liu, B. (2024). A Survey on MLLMs in Education: Application and Future Directions. *Future Internet*, 16(12), 467. <https://doi.org/10.3390/fi16120467>
- Yuhana, Y. (2024). Dampak penggunaan kecerdasan buatan dalam proses pendidikan islam. *Al-kainah: Journal of Islamic Studies*, 3(1), 167–176. <https://doi.org/10.69698/jis.v3i1.149>