

Efektivitas, Strategi *Prompt* , dan Risiko Penggunaan AI Generatif dalam Penyusunan Rencana Pembelajaran: Sebuah Pendekatan *Systematic Literature Review*

Syhabudin^{1,2}, Muhammad Rizal³

¹ MAN Bulungan, Indonesia

² Prodi Magister Pendidikan Agama Islam, Universitas Islam Sultan Agung, Indonesia

³ Kanwil Kemenag Kalimantan Utara, Indonesia

syhabudin@manbulungan.sch.id

Abstract: *This study synthesizes evidence on the effectiveness, prompt strategies, and risks of using generative AI for lesson planning. The review followed PRISMA 2020 and covered Scopus- and SINTA-indexed journal articles published from 2023 to April 2026. Sixteen studies were included and analysed through qualitative thematic synthesis, supported by MMAT-based appraisal for empirical studies and descriptive appraisal for conceptual work. The findings show that generative AI most consistently improves time efficiency, supports the initial structure of lesson plans, helps complete core components, and facilitates revision. The most productive prompt strategies are contextual, role- or rubric-based, iterative, and example-based prompts. The dominant risk is pedagogical misalignment, followed by information hallucination, bias, and ethical concerns. Overall, generative AI is most valuable when positioned as a drafting partner that remains under the teacher's professional judgement.*

Keywords: *generative AI, lesson planning, prompt engineering, instructional planning, systematic literature review*

Abstrak: Penelitian ini mensintesis bukti tentang efektivitas, strategi *prompt*, dan risiko penggunaan AI generatif dalam penyusunan rencana pembelajaran. Review sistematis ini mengikuti PRISMA 2020 atas artikel jurnal Scopus dan SINTA periode 2023 - April 2026. Seleksi menghasilkan 16 studi yang dianalisis melalui sintesis tematik kualitatif dan appraisal kualitas berbasis MMAT serta pembacaan deskriptif untuk artikel konseptual. Hasil menunjukkan bahwa AI generatif paling konsisten meningkatkan efisiensi waktu, membantu struktur awal *lesson plan*, melengkapi komponen dasar, dan memudahkan revisi. Strategi *prompt* yang paling produktif ialah *prompt* kontekstual, berbasis peran atau rubrik, iteratif, dan berbasis contoh. Risiko yang paling menonjol ialah ketidaksesuaian pedagogik, diikuti halusinasi informasi, bias, serta isu etika. Temuan menegaskan bahwa AI generatif paling bernilai ketika ditempatkan sebagai mitra drafting yang tetap dikendalikan oleh penilaian profesional guru.

Kata Kunci: *AI generatif, rencana pembelajaran, prompt engineering, lesson planning, systematic literature review*

A. Pendahuluan

Implementasi kurikulum modern telah mengeskalasi beban guru melalui standarisasi perencanaan yang menuntut rincian mendalam pada aspek diferensiasi, asesmen hingga pada persoalan administrasi yang adaptif. AI generatif menawarkan bantuan nyata pada fase *drafting* karena mampu menghasilkan tujuan, langkah kegiatan, asesmen, dan diferensiasi berbasis *prompt*. Namun, kegunaan teknis ini tidak otomatis menjamin kualitas pedagogik; keluaran AI tetap bergantung pada kualitas *prompt*, verifikasi pengguna, dan konteks

kurikulum. Oleh karena itu, isu AI generatif dalam *lesson planning* perlu dibaca bukan hanya sebagai efisiensi kerja, tetapi juga sebagai persoalan pengetahuan profesional guru. Pada level global, kajian AI dalam pendidikan tumbuh cepat; pada level nasional, kebutuhan modul ajar yang kontekstual dalam Kurikulum Merdeka membuat topik ini semakin relevan. Celah riset terletak pada minimnya sintesis yang mempertemukan tiga dimensi sekaligus efektivitas, strategi *prompt*, dan risiko secara khusus dalam konteks penyusunan rencana pembelajaran. Penggabungan sumber Scopus dan SINTA memberi kebaruan kontekstual karena memungkinkan pembacaan serentak atas kecenderungan internasional dan implementasi Indonesia.

Kajian ini berangkat dari urgensi untuk memetakan peran AI generatif sebagai alat bantu penyusunan draf perencanaan pembelajaran, mulai dari modul ajar hingga rencana pelaksanaan pembelajaran (RPP). Urgensi tersebut dikuatkan oleh bukti empiris bahwa AI generatif mampu mengurangi beban kerja guru dan mendukung personalisasi perencanaan melalui *tailored queries*,¹ sekaligus oleh temuan bahwa kualitas rencana pembelajaran yang dihasilkan sangat bergantung pada kemampuan guru berinteraksi secara reflektif dengan sistem AI tersebut.² Fokus utama penelitian ini diarahkan pada identifikasi bentuk efektivitas praktis yang dihasilkan, apakah lebih dominan pada efisiensi waktu, kualitas struktur, personalisasi, kelengkapan komponen, atau kemudahan revisi (RQ1). Selain itu, kajian ini bertujuan untuk mengevaluasi produktivitas berbagai strategi *prompt ing* baik yang bersifat umum, kontekstual, berbasis peran, berbasis rubrik, iteratif, maupun berbasis contoh guna menentukan pendekatan yang paling optimal bagi pendidik (RQ2). Secara kritis, analisis ini juga menimbang risiko dominan yang menyertai penggunaan teknologi tersebut, seperti halusinasi informasi, bias, ketidaksesuaian pedagogis, hingga isu etika dan ketergantungan guru (RQ3). Melalui penelitian ini diharapkan mampu memberikan kontribusi berupa pemetaan konseptual serta panduan praktis yang esensial bagi guru, peneliti, dan lembaga pendidikan dalam mengintegrasikan AI generatif secara bertanggung jawab.

B. Metode Penelitian

Penelitian ini menggunakan pendekatan *Systematic Literature Review* (SLR) dengan acuan PRISMA 2020. Kajian diposisikan sebagai studi kepustakaan sistematis, sehingga unit analisisnya adalah artikel jurnal yang relevan dengan penggunaan AI generatif untuk *lesson planning*. Review ini tidak diregistrasikan pada PROSPERO karena berada pada bidang pendidikan non-klinis. Meski demikian, protokol kerja yang memuat RQ, string pencarian, kriteria seleksi, instrumen ekstraksi, dan langkah sintesis ditetapkan sebelum pencarian; sepanjang proses tidak ada perubahan substantif, kecuali penyeragaman terminologi *lesson plan*, modul ajar, RPP, dan instructional planning.

Sumber informasi ialah Scopus dan SINTA karena mewakili horizon internasional dan nasional yang paling relevan. Pencarian terakhir dilakukan pada April 2026. String pencarian menggabungkan istilah AI generatif dan perencanaan pembelajaran dalam bahasa Inggris dan Indonesia dengan operator Boolean. Dua string Scopus menghasilkan 24 rekaman awal (15 + 9), sedangkan dua string SINTA menghasilkan 8 rekaman (5 + 3).

¹Hashem et al., "AI to the rescue" (2024).

²Tran, "Preparing preservice teachers for generative AI in lesson planning" (2025).

Tabel 1. Strategi Pencarian Literatur dan Rekaman Awal

Basis data	String inti pencarian	Filter operasional	Rekaman awal
Scopus-1	TITLE-ABS-KEY(("generative AI" OR ChatGPT OR "large language model") AND ("lesson plan*" OR "lesson planning" OR "instructional planning" OR "curriculum planning"))	Artikel jurnal; 2023–April 2026; Inggris	15
Scopus-2	TITLE-ABS-KEY((ChatGPT OR "generative AI") AND (prompt OR "prompt engineering" OR risk OR ethics) AND ("lesson plan*" OR "teaching module*" OR "instructional planning"))	Artikel jurnal; 2023–April 2026; Inggris	9
SINTA-1	("AI generatif" OR ChatGPT) AND ("rencana pembelajaran" OR "modul ajar" OR RPP)	Artikel jurnal; 2023–April 2026; Indonesia/Inggris	5
SINTA-2	(ChatGPT AND guru AND ("modul ajar" OR "perangkat pembelajaran" OR "lesson plan"))	Artikel jurnal; 2023–April 2026; Indonesia/Inggris	3
Total	Empat string inti yang mewakili fokus efektivitas, prompt, dan risiko pada kegiatan perencanaan pembelajaran.	Seleksi awal pada Scopus dan SINTA	32

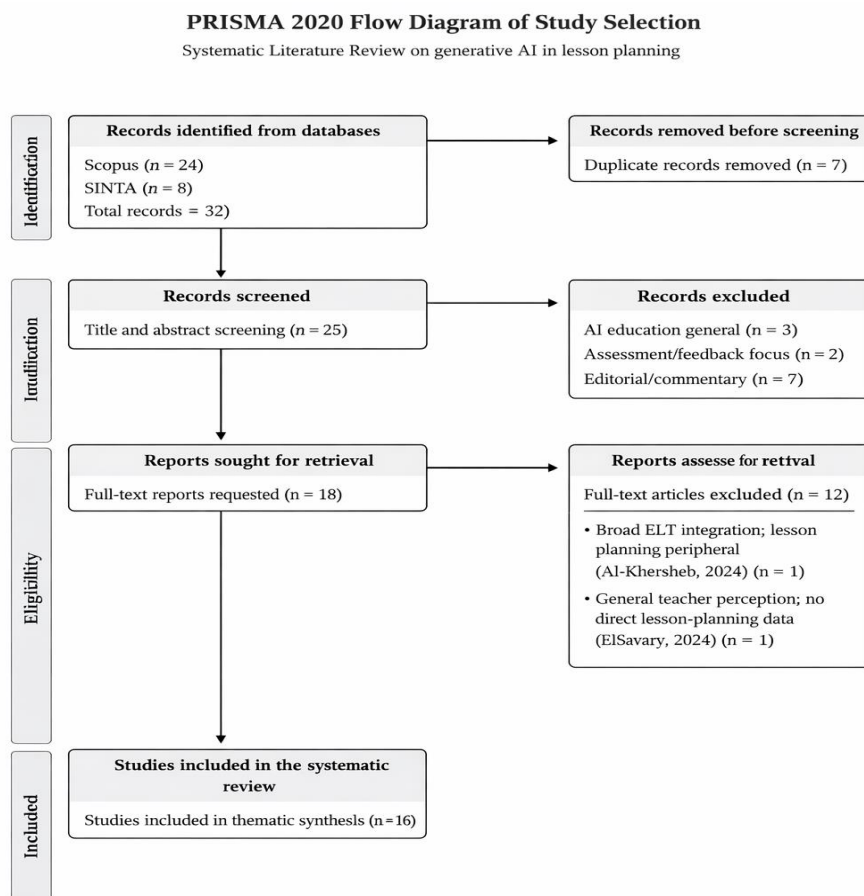
Artikel dimasukkan bila berupa jurnal berbahasa Indonesia atau Inggris, tersedia *full text*, terindeks Scopus atau terdaftar di SINTA, dan membahas AI generatif secara langsung untuk penyusunan *lesson plan*, modul ajar, RPP, atau *instructional planning*. Artikel opini, editorial, duplikat, dan studi AI pendidikan umum tanpa kaitan langsung dengan perencanaan pembelajaran dikeluarkan.

Tabel 2. Kriteria Inklusi dan Eksklusi

Aspek	Kriteria inklusi	Kriteria eksklusi
Jenis dokumen	Artikel jurnal ilmiah dengan naskah penuh.	Editorial, opini, berita, atau naskah dengan bobot ilmiah perifer.
Indeks sumber	Terindeks Scopus atau terdaftar di SINTA.	Publikasi di luar dua basis seleksi utama.
Fokus substansi	AI generatif berhubungan langsung dengan <i>lesson plan</i> , modul ajar, RPP, <i>instructional planning</i> , <i>teaching plan</i> , atau <i>teaching scenario</i> .	Fokus utama berada pada AI pendidikan umum, asesmen, atau pengembangan materi dengan kedekatan rendah terhadap perencanaan pembelajaran.
Bahasa	Bahasa Indonesia atau Inggris.	Bahasa lain di luar dua bahasa kerja kajian.
Rentang tahun	2023–April 2026.	Terbitan di luar rentang akselerasi AI generatif yang dipilih.

Aspek	Kriteria inklusi	Kriteria eksklusi
Akses teks	Teks penuh tersedia dan dapat dibaca utuh.	Akses teks terbatas sehingga analisis tematik tidak dapat dijalankan secara memadai.
Kandungan tematik	Memuat data, temuan, atau refleksi yang menjelaskan efektivitas, strategi <i>prompt</i> , atau risiko.	Kedekatan tema sangat tipis sehingga kontribusi sintesis menjadi minimal.

Seleksi mengikuti PRISMA 2020 melalui deduplikasi, screening judul-abstrak, dan penilaian *full-text*. Dari 32 rekaman awal, 7 duplikat dihapus, 25 artikel *discreening*, 18 artikel dinilai *full-text*, dan 16 studi diinklusi. Seluruh proses dilakukan oleh satu *reviewer* menggunakan lembar seleksi terstruktur; karena itu, *inter-rater reliability* tidak dihitung dan kondisi ini diakui sebagai keterbatasan metodologis.



Gambar 1. Diagram alur PRISMA 2020 hasil seleksi artikel

Data diekstrak melalui lembar terstruktur yang memuat identitas studi, konteks, platform AI, bentuk penggunaan, temuan efektivitas, strategi *prompt*, risiko, dan rekomendasi. Al-khresheh³ dan ElSayary⁴ tetap digunakan sebagai rujukan latar belakang, tetapi tidak masuk korpus sintesis.

Tabel 3. Komponen Ekstraksi Data dan Fungsi Analitis

Komponen ekstraksi	Fungsi analitis
Penulis dan tahun	Menandai kronologi perkembangan riset dan sumber sitasi utama.
Konteks atau negara	Membaca variasi konteks internasional dan nasional.
Sumber indeks	Membedakan karakter temuan Scopus dan SINTA.
Tujuan penelitian	Menentukan orientasi utama setiap studi.
Desain penelitian	Membaca bobot bukti empiris dan karakter metodologis studi.
Subjek atau konteks penggunaan	Mengidentifikasi guru, calon guru, atau konteks pelatihan.
Platform AI yang dikaji	Menandai perbedaan platform, seperti ChatGPT, EduAid, Magic School, atau alat sejenis.
Bentuk penggunaan AI	Membedakan penggunaan untuk draf awal, revisi, evaluasi, atau desain terstruktur.
Temuan efektivitas	Menangkap bentuk nilai praktis AI generatif untuk <i>lesson planning</i> .
Strategi <i>prompt</i>	Membaca jenis <i>prompt</i> yang paling produktif.
Risiko atau tantangan	Mengidentifikasi pola risiko yang muncul lintas studi.
Rekomendasi penelitian	Menemukan arah pengembangan riset selanjutnya.

Kualitas studi ditelaah dengan MMAT untuk artikel empiris dan appraisal deskriptif untuk artikel konseptual. Sintesis dilakukan secara tematik kualitatif melalui *coding*, pengelompokan tema, perbandingan lintas studi, pembacaan reporting bias, dan penetapan kategori kepastian bukti: sedang-berkembang (dukungan awal dan konteks terbatas), sedang (tema konsisten lintas beberapa desain), serta kuat atau sedang-kuat (didukung lintas konteks dengan desain lebih kokoh).

C. Hasil dan Pembahasan

1. Profil Studi Terinklusi

Korpus akhir mencakup 16 studi: 13 Scopus dan 3 SINTA, dengan akselerasi publikasi pada 2024–2025. Sebagian besar studi bergerak pada pendidikan guru, calon guru, atau pelatihan guru; ChatGPT menjadi platform yang paling dominan, disusul Gemini dan variasi GenAI lain. Konteks studi memperlihatkan bahwa AI digunakan terutama untuk drafting awal *lesson plan*, perbaikan struktur, penyusunan modul ajar, dan latihan *prompt design*

³Al-khresheh, "Bridging technology and pedagogy from a global lens" (2024).

⁴ElSayary, "An investigation of teachers' perceptions of using ChatGPT as a supporting tool for teaching and learning in the digital era" (2024).

pada pendidikan guru. Ragam ini membuat sintesis dapat menangkap manfaat praktis sekaligus batas kontekstual implementasinya. Secara kualitas, korpus memperlihatkan kombinasi tiga studi kuat, sepuluh studi memadai, dan tiga studi eksploratif. Profil ini cukup untuk mendukung sintesis tematik, meskipun kekuatan bukti tidak merata pada seluruh domain.

Tabel 4. Karakteristik Studi Terinklusi (Bagian 1)

Kode	Studi	Indeks/Konteks	Desain	Fokus inti
S1	van den Berg & du Plessis (2023)	Scopus; teacher education	Analisis dokumen	Kontribusi AI generatif pada <i>lesson planning</i> , critical thinking, dan openness.
S2	Baytak (2024)	Scopus; kelas 7 berbagai mata pelajaran	Analisis isi	Struktur <i>lesson plan</i> ChatGPT dan Gemini serta kedalaman aktivitas belajar.
S3	Karaman & Göksu (2024)	Scopus; matematika sekolah dasar	Eksperimen	Efek <i>lesson plan</i> berbantuan ChatGPT terhadap capaian belajar matematika.
S4	Moundridou et al. (2024)	Scopus; inquiry-based <i>lesson planning</i>	Analisis dokumen	Pemetaan peran GenAI pada fase perancangan <i>lesson plan</i> berbasis inkuiri.
S5	Hashem et al. (2024)	Scopus; guru sekolah menengah	Studi penjelas	Tailored queries, perencanaan personal, dan pengurangan beban kerja guru.
S6	Gurl et al. (2025)	Scopus; calon guru matematika	Kualitatif	ChatGPT sebagai <i>lesson planning assistant</i> dan bahan refleksi kritis.
S7	Kalenda et al. (2025)	Scopus; metode STEM, TESOL, social studies	Mixed methods	Persepsi calon guru setelah guided analysis terhadap <i>lesson plan</i> ChatGPT.
S8	Luo & Tahir (2025)	Scopus; STEAM arts education	Eksperimen	Manfaat, tantangan, dan <i>prompt framework</i> untuk <i>lesson planning</i> .

Tabel 5. Karakteristik Studi Terinklusi (Bagian 2)

Kode	Studi	Indeks/Konteks	Desain	Fokus inti
S9	Peikos & Stavrou (2025)	Scopus; sains sekolah dasar	Eksploratif	PCK, layer <i>prompt</i> s, dan reference text untuk kualitas <i>lesson plan</i> sains.
S10	Dornburg & Davin (2025)	Scopus; pembelajaran bahasa asing	Analisis komparatif	Variabilitas kualitas <i>lesson plan</i> dan historical bias.
S11	ElSayary et al. (2025)	Scopus; 17 guru sekolah menengah	Studi kasus kualitatif	<i>Prompt engineering</i> untuk meningkatkan relevansi <i>lesson plan</i> .
S12	Biberman-Shalev (2025)	Scopus; curriculum planning course	Fenomenologi	Translasi teori ke praktik melalui ChatGPT-4.
S13	Şimşek (2025)	Scopus; 5E <i>lesson planning</i> matematika	Analisis interaksi	Interaksi guru dan calon guru dengan ChatGPT pada desain 5E.
S14	Istiarsyah et al. (2024)	SINTA; guru SMP Aceh	Quasi-eksperimen	Kompetensi guru dan kualitas <i>lesson modules</i> berbantuan ChatGPT.
S15	Siregar et al. (2025)	SINTA; guru SMA	Deskriptif-pelatihan	Perangkat pembelajaran, kepercayaan diri, dan efisiensi penyusunan.
S16	Sukmadana & Halim (2026)	SINTA; calon guru EFL Indonesia	Mixed methods	Persepsi kegunaan, struktur <i>lesson plan</i> , dan kebutuhan revisi.

2. Efektivitas AI Generatif dalam Penyusunan Rencana Pembelajaran

Jawaban terhadap RQ1 memperlihatkan bahwa efektivitas AI generatif hadir paling jelas pada efisiensi waktu. Hashem, Ali, El Zein, Fidalgo, dan Abu Khurma⁵ menunjukkan bahwa *tailored queries* mendorong percepatan kerja guru pada aktivitas perencanaan yang selama ini menyerap energi tinggi. Istiarsyah, Maisura, Nurhabibah, Kumita, dan Kamarullah⁶ serta Siregar, Kurniawan, Uyun, dan Gea⁷ memperlihatkan pola serupa pada konteks Indonesia, yaitu guru merasakan percepatan ketika AI digunakan untuk menyusun modul ajar, perangkat pembelajaran, dan struktur RPP. Pada level pendidikan guru, Kalenda, Rath, Abugasea Heidt, dan Wright⁸ serta Sukmadana dan Halim⁹ juga merekam persepsi positif terhadap kecepatan memperoleh *draf lesson plan*, ide aktivitas, dan pengorganisasian komponen inti.

⁵Hashem et al., "AI to the rescue" (2024).

⁶Istiarsyah et al., "Transforming teacher competence through ChatGPT" (2024).

⁷Siregar et al., "Penggunaan ChatGPT untuk menyusun perangkat pembelajaran tingkat SMA" (2025).

⁸Kalenda et al., "Pre-service teacher perceptions of ChatGPT for lesson plan generation" (2025).

⁹Sukmadana and Halim, "Perceptions of pre-service English teachers on using ChatGPT for lesson planning" (2026).

Temuan lintas studi ini menegaskan bahwa AI generatif memberi nilai awal yang tinggi sebagai penghemat waktu kerja perencanaan. Dari sudut cognitive load theory, pola ini dapat dipahami sebagai penurunan beban kognitif ekstrinsik pada tahap drafting awal, sehingga perhatian guru dapat dialihkan ke keputusan yang lebih substantif seperti diferensiasi, validasi isi, dan penyesuaian pedagogik keluaran.¹⁰

Dimensi efektivitas kedua bergerak pada kualitas struktur awal rencana pembelajaran. Baytak¹¹ memperlihatkan bahwa *lesson plan* yang dihasilkan ChatGPT dan Gemini memiliki organisasi yang menyerupai rancangan guru pada unsur tujuan, aktivitas, dan asesmen. Moundridou, Matzakos, dan Doukakis¹² menunjukkan bahwa GenAI dapat dipetakan secara produktif pada fase-fase inquiry-based *lesson planning*, sehingga alat ini membantu guru menyusun alur desain secara lebih runtut. Gurl, Markinson, dan Artzt¹³ menemukan bahwa calon guru matematika memanfaatkan ChatGPT sebagai titik mula yang memudahkan pembentukan struktur, seraya menjalankan telaah kritis pada substansi keluaran. Dari sini terlihat bahwa AI generatif mempunyai kekuatan kuat sebagai pembentuk kerangka awal yang koheren.

Efektivitas berikutnya tampak pada kelengkapan komponen dasar. Karaman dan Göksu¹⁴ menunjukkan bahwa *lesson plan* berbantuan ChatGPT dapat mendukung proses pembelajaran matematika sekolah dasar dengan capaian belajar yang meningkat secara bermakna di dalam kelompok eksperimen. Biberman-Shalev¹⁵ menunjukkan bahwa ChatGPT-4 membantu calon guru menghubungkan teori dengan komponen kurikulum melalui proses penyederhanaan, aplikasi, dan visualisasi teori. Şimşek¹⁶ juga memperlihatkan bahwa interaksi guru dan calon guru dengan ChatGPT pada 5E *lesson planning* memperkaya unsur fase pembelajaran secara lebih sistematis. Dengan demikian, AI generatif memberi kontribusi pada kelengkapan tujuan, langkah kegiatan, dan asesmen, terutama ketika pengguna memasukkan kerangka pedagogik yang jelas.

Personalisasi muncul sebagai dimensi efektivitas yang bergerak paling dinamis. Luo dan Tahir¹⁷ menunjukkan bahwa kualitas *lesson plan* meningkat ketika *prompt* memuat konteks pembelajaran, tujuan, dan kebutuhan spesifik peserta didik. Peikos dan Stavrou¹⁸ mendemonstrasikan bahwa *prompt* yang memuat aspek *pedagogical content knowledge* menghasilkan *lesson plan* sains yang lebih relevan. Pada konteks Indonesia, Istiarsyah et al.¹⁹ dan Sukmadana dan Halim²⁰ menunjukkan bahwa AI membantu guru dan calon guru menyesuaikan struktur pembelajaran secara lebih cepat, sementara konteks kurikulum lokal, karakter bahasa, dan kebutuhan kelas tetap berkembang paling baik melalui penajaman guru.

¹⁰Sweller, "Cognitive load during problem solving" (1988).

¹¹Baytak, "The content analysis of the lesson plans created by ChatGPT and Google Gemini" (2024).

¹²Moundridou et al., "Generative AI tools as educators' assistants" (2024).

¹³Gurl et al., "Using ChatGPT as a lesson planning assistant with preservice secondary mathematics teachers" (2025).

¹⁴Karaman and Göksu, "Are lesson plans created by ChatGPT more effective? An experimental study" (2024).

¹⁵Biberman-Shalev, "Prompting theory into practice" (2025).

¹⁶Şimşek, "Integration of ChatGPT in mathematical story-focused 5E lesson planning" (2025).

¹⁷Luo and Tahir, "ChatGPT-assisted lesson planning for children's STEAM arts education" (2025).

¹⁸Peikos and Stavrou, "ChatGPT for science lesson planning" (2025).

¹⁹Istiarsyah et al., "Transforming teacher competence through ChatGPT" (2024).

²⁰Sukmadana and Halim, "Perceptions of pre-service English teachers on using ChatGPT for lesson planning" (2026).

Temuan ini memperlihatkan bahwa AI generatif menyediakan personalisasi potensial yang menjadi semakin kuat ketika guru memasukkan data kontekstual secara kaya.

Dimensi efektivitas lain yang konsisten ialah kemudahan revisi. ElSayary, Kuhail, dan Hojeij²¹ menggambarkan bahwa guru memperoleh hasil paling relevan ketika keluaran awal AI diperlakukan sebagai bahan revisi bertahap. Biberman-Shalev²² memperlihatkan bahwa dialog berulang dengan ChatGPT-4 menolong calon guru menata ulang relasi antara teori dan praktik. Pada studi Gurl et al.²³, draf yang disusun AI juga berfungsi sebagai objek refleksi yang memicu penilaian ulang terhadap tujuan, urutan kegiatan, dan orientasi pembelajaran. Dengan kata lain, nilai AI generatif berkembang lebih tinggi ketika guru memosisikannya sebagai mitra revisi yang memudahkan penulisan ulang, penyempurnaan rubrik, dan penataan kembali alur kegiatan.

Secara sintesis, efektivitas AI generatif dalam penyusunan rencana pembelajaran paling kuat hadir dalam bentuk efisiensi waktu, kualitas struktur awal, kelengkapan komponen dasar, dan kemudahan revisi. Personalisasi turut berkembang secara positif ketika *prompt* memuat konteks pedagogik secara kaya. RQ1 dengan demikian dapat dijawab bahwa efektivitas yang paling konsisten muncul pada efisiensi waktu, kualitas struktur, kelengkapan komponen, dan kemudahan revisi, sementara personalisasi mencapai mutu tertinggi melalui co-design antara AI dan guru.

3. Strategi *Prompt* yang Paling Produktif untuk *Lesson planning*

Jawaban terhadap RQ2 menunjukkan bahwa strategi *prompt* yang paling produktif bergerak dari *prompt* yang kaya konteks menuju *prompt* yang terstruktur secara pedagogik. *Prompt* umum tetap berguna sebagai pembuka eksplorasi ide, terutama untuk menghasilkan draf pertama, variasi aktivitas, atau daftar komponen awal. Hashem et al.²⁴ serta Kalenda et al.²⁵ memperlihatkan bahwa *prompt* umum memberi manfaat tinggi pada fase brainstorming. Nilai tersebut menjadi semakin kuat ketika guru melanjutkan percakapan ke *prompt* yang lebih spesifik.

Prompt kontekstual muncul sebagai strategi dengan dukungan lintas studi yang paling konsisten. Luo dan Tahir²⁶ merumuskan *prompt* framework yang menempatkan konteks pembelajaran sebagai fondasi awal, meliputi tema, tujuan, kelompok usia, alokasi waktu, dan target keluaran. Siregar et al.²⁷ serta Istiarsyah et al.²⁸ memberi bukti praktis bahwa guru Indonesia memperoleh hasil lebih sesuai ketika konteks kelas, kurikulum, dan kebutuhan modul ajar dituliskan secara rinci. *Prompt* kontekstual bekerja efektif karena AI generatif merespons petunjuk yang konkret, sehingga keluaran bergerak lebih dekat ke realitas kelas. *Prompt* berbasis peran juga memperlihatkan daya guna yang tinggi. Pengguna dapat memosisikan AI sebagai guru sains sekolah dasar, guru bahasa Inggris calon praktisi,

²¹ElSayary et al., "Examining the role of prompt engineering in utilizing generative AI tools for lesson planning" (2025).

²²Biberman-Shalev, "Prompting theory into practice" (2025).

²³Gurl et al., "Using ChatGPT as a lesson planning assistant with preservice secondary mathematics teachers" (2025).

²⁴Hashem et al., "AI to the rescue" (2024).

²⁵Kalenda et al., "Pre-service teacher perceptions of ChatGPT for lesson plan generation" (2025).

²⁶Luo and Tahir, "ChatGPT-assisted lesson planning for children's STEAM arts education" (2025).

²⁷Siregar et al., "Penggunaan ChatGPT untuk menyusun perangkat pembelajaran tingkat SMA" (2025).

²⁸Istiarsyah et al., "Transforming teacher competence through ChatGPT" (2024).

fasilitator inkuiri, atau penyusun modul ajar Kurikulum Merdeka. Peikos dan Stavrou²⁹ menunjukkan bahwa pelibatan aspek pedagogical *content knowledge* sesungguhnya menghadirkan *prompt* berbasis peran yang kaya akan identitas pedagogik. Şimşek³⁰ menampilkan pola serupa pada desain 5E, ketika interaksi guru dengan ChatGPT menjadi lebih produktif saat peran pedagogik dan model pembelajaran dinyatakan secara eksplisit. Strategi ini membantu AI menjaga fokus keluaran pada orientasi profesi dan kebutuhan kelas yang lebih spesifik.

Prompt berbasis rubrik dan kriteria memberi kontribusi penting pada kestabilan mutu keluaran. Dornburg dan Davin³¹ memperlihatkan bahwa *lesson plan* bahasa yang dihasilkan AI menunjukkan variasi kualitas yang cukup lebar, sehingga daftar kriteria dan indikator mutu menjadi penopang yang sangat berguna. Biberman-Shalev³² menunjukkan bahwa penggunaan bahasa kurikuler yang presisi membantu mahasiswa memindahkan teori ke praktik secara lebih tertata. Pada tataran operasional, rubrik, checklist, atau indikator mutu mengubah *prompt* dari instruksi umum menjadi instruksi evaluatif, sehingga AI memperoleh arah yang lebih jelas mengenai kualitas yang diharapkan.

Prompt iteratif tampil sebagai strategi yang paling berulang muncul di hampir seluruh korpus. ElSary et al.³³ menegaskan bahwa kualitas keluaran berkembang paling baik melalui refinement bertahap melalui serangkaian *prompt* yang saling menyempurnakan. Hashem et al.³⁴ menunjukkan fungsi tailored queries yang berkembang seiring dialog. Gurl et al.³⁵, Kalenda et al.³⁶, dan Sukmadana dan Halim³⁷ juga memperlihatkan bahwa evaluasi pengguna terhadap *lesson plan* AI menjadi semakin tajam setelah beberapa putaran revisi. Strategi iteratif pada akhirnya memperlihatkan bahwa interaksi dengan AI generatif bersifat dialogis dan progresif. Dalam perspektif *scaffolding*, pola ini menunjukkan bahwa AI dapat dikondisikan sebagai *more knowledgeable other* yang menyediakan dukungan sementara melalui *prompt* bertahap, sementara kontrol keputusan pedagogik final tetap berada pada guru.³⁸

Prompt berbasis contoh atau reference text memiliki nilai tinggi terutama ketika guru ingin menjaga format, kedalaman, dan gaya keluaran. Peikos dan Stavrou³⁹ menunjukkan bahwa reference text membantu ChatGPT menghasilkan *lesson plan* sains yang lebih kuat melalui lensa PCK. Biberman-Shalev⁴⁰ memperlihatkan bahwa contoh bahasa kurikuler dan contoh implementasi teori memperkuat translasi teori ke rancangan pembelajaran. Dalam konteks Indonesia, *prompt* berbasis contoh sangat potensial untuk menyelaraskan keluaran dengan format modul ajar atau perangkat pembelajaran yang berlaku di satuan pendidikan.

²⁹Peikos and Stavrou, "ChatGPT for science lesson planning" (2025).

³⁰Şimşek, "Integration of ChatGPT in mathematical story-focused 5E lesson planning" (2025).

³¹Dornburg and Davin, "ChatGPT in foreign language lesson plan creation" (2025).

³²Biberman-Shalev, "Prompting theory into practice" (2025).

³³ElSary et al., "Examining the role of prompt engineering in utilizing generative AI tools for lesson planning" (2025).

³⁴Hashem et al., "AI to the rescue" (2024).

³⁵Gurl et al., "Using ChatGPT as a lesson planning assistant with preservice secondary mathematics teachers" (2025).

³⁶Kalenda et al., "Pre-service teacher perceptions of ChatGPT for lesson plan generation" (2025).

³⁷Sukmadana and Halim, "Perceptions of pre-service English teachers on using ChatGPT for lesson planning" (2026).

³⁸Vygotsky, "Mind in society" (1978).

³⁹Peikos and Stavrou, "ChatGPT for science lesson planning" (2025).

⁴⁰Biberman-Shalev, "Prompting theory into practice" (2025).

Secara sintesis, RQ2 dapat dijawab bahwa strategi *prompt* lebih efektif ketika bersifat kontekstual, berbasis peran, berbasis rubrik, iteratif, dan berbasis contoh. *Prompt* umum berfungsi paling baik sebagai titik mula eksplorasi. Temuan ini mengarah pada satu implikasi penting: kompetensi *prompt* design bagi guru merupakan bagian dari kompetensi pedagogik-digital, karena di dalamnya guru menerjemahkan kebutuhan belajar ke dalam instruksi yang dapat dipahami AI secara tepat.

4. Risiko Penggunaan AI Generatif dalam Penyusunan Rencana Pembelajaran

Jawaban terhadap RQ3 memperlihatkan bahwa risiko paling dominan berada pada wilayah ketidaksesuaian pedagogik. Gurl et al.⁴¹ menunjukkan bahwa keluaran ChatGPT pada konteks matematika calon guru sering bergerak ke arah *teacher-centered* dan repetitif. Baytak⁴² memperlihatkan bahwa integrasi teknologi dan kedalaman konstruktivisme pada *lesson plan* hasil AI masih memerlukan penajaman. Peikos dan Stavrou⁴³ menambahkan bahwa *lesson plan* sains mencapai kualitas lebih baik ketika *prompt* diperkaya dengan elemen PCK. Temuan-temuan ini memberi gambaran kuat bahwa *lesson plan* yang tampak rapi secara tekstual tetap memerlukan pembacaan pedagogik yang cermat.

Halusinasi informasi menjadi risiko berikutnya yang sangat relevan, terutama pada mata pelajaran yang menuntut ketelitian konsep. Peikos dan Stavrou⁴⁴, Gurl et al.⁴⁵, serta Şimşek⁴⁶ memperlihatkan pentingnya verifikasi isi, konsep, dan contoh yang dihasilkan AI. Dalam konteks *lesson planning*, halusinasi dapat muncul sebagai definisi yang keliru, urutan aktivitas yang tidak logis, asesmen yang tidak sepadan dengan tujuan, atau materi yang tampak meyakinkan tetapi rapuh secara akademik. Karena itu, penggunaan AI generatif untuk perencanaan pembelajaran tetap menuntut validasi sumber, pemeriksaan kesesuaian konsep, dan pembacaan ulang berbasis pengetahuan pedagogik guru.

Bias hadir baik sebagai bias isi, bias perspektif, maupun bias historis. Dornburg dan Davin⁴⁷ menunjukkan bahwa AI dapat mereproduksi historical bias pada *lesson plan* bahasa asing, yakni kecenderungan untuk menghidupkan kembali praktik pedagogik yang sudah lama memperoleh kritik. Luo dan Tahir⁴⁸, van den Berg dan du Plessis⁴⁹, serta Moundridou et al.⁵⁰ juga menandai pentingnya kecocokan konteks budaya, kebutuhan lokal, dan tujuan kurikuler. Pada konteks nasional, temuan SINTA memperlihatkan pentingnya penajaman kurikulum, karakter bahasa, dan kondisi kelas Indonesia agar keluaran AI tidak bergerak pada template yang seragam. Risiko bias ini menjelaskan bahwa kualitas keluaran AI tidak hanya ditentukan oleh kecanggihan model, tetapi juga oleh kualitas *prompt*, konteks yang diberikan, dan kecermatan guru saat menilai hasilnya.

⁴¹Gurl et al., "Using ChatGPT as a lesson planning assistant with preservice secondary mathematics teachers" (2025).

⁴²Baytak, "The content analysis of the lesson plans created by ChatGPT and Google Gemini" (2024).

⁴³Peikos and Stavrou, "ChatGPT for science lesson planning" (2025).

⁴⁴Peikos and Stavrou, "ChatGPT for science lesson planning" (2025).

⁴⁵Gurl et al., "Using ChatGPT as a lesson planning assistant with preservice secondary mathematics teachers" (2025).

⁴⁶Şimşek, "Integration of ChatGPT in mathematical story-focused 5E lesson planning" (2025).

⁴⁷Dornburg and Davin, "ChatGPT in foreign language lesson plan creation" (2025).

⁴⁸Luo and Tahir, "ChatGPT-assisted lesson planning for children's STEAM arts education" (2025).

⁴⁹van den Berg and du Plessis, "ChatGPT and generative AI" (2023).

⁵⁰Moundridou et al., "Generative AI tools as educators' assistants" (2024).

Plagiarisme, ketergantungan guru, serta isu privasi dan etika juga muncul sebagai risiko penting. Siregar et al.⁵¹ dan Istiarsyah et al.⁵² menunjukkan antusiasme guru terhadap percepatan kerja yang dibawa ChatGPT, sehingga ruang pelatihan literasi AI menjadi penting agar kecepatan berjalan seiring dengan orisinalitas desain dan kualitas penalaran pedagogik. Sukmadana dan Halim⁵³ menunjukkan bahwa calon guru menikmati manfaat AI untuk ide dan struktur, sambil tetap menyadari perlunya revisi besar pada hasil awal. ElSayary et al.⁵⁴ memberi penekanan pada AI literacy dan continuous professional *development*, sedangkan ElSayary⁵⁵ menempatkan etika dan evaluasi guru sebagai bagian integral dari penggunaan AI yang sehat.

Jika tiga RQ dibaca secara terpadu, maka risiko paling dominan dapat dirumuskan sebagai ketidaksesuaian pedagogik yang sering bergerak bersama halusinasi informasi dan bias. Plagiarisme, ketergantungan guru, serta privasi dan etika mengikuti sebagai lapisan risiko yang menuntut tata kelola profesional. RQ3 karena itu dapat dijawab dengan menempatkan ketidaksesuaian pedagogik pada posisi pertama, diikuti halusinasi informasi dan bias, kemudian plagiarisme, ketergantungan guru, serta privasi dan etika.

Ketidaksesuaian pedagogik yang muncul sebagai risiko dominan dapat dijelaskan lebih jauh melalui kerangka TPACK. AI generatif relatif kuat dalam menghasilkan *content-rich draft* dan variasi representasi umum, tetapi kualitas *lesson plan* baru mencapai kecocokan pedagogik ketika guru memadukan pengetahuan konten, pedagogi, teknologi, tujuan kurikuler, dan karakter peserta didik secara terintegrasi.⁵⁶ Karena itu, keluaran AI yang tampak rapi secara tekstual tidak otomatis memadai secara pedagogis; tanpa pengarahan *prompt* yang tepat dan tanpa validasi guru, AI cenderung kuat pada level permukaan struktur, namun belum selalu selaras dengan strategi pembelajaran, urutan pengalaman belajar, dan asesmen yang kontekstual. Implikasi praktisnya, pelatihan guru perlu dikembangkan sebagai workshop *prompt* design berbasis TPACK yang mengintegrasikan keputusan konten, strategi pedagogik, dan penggunaan teknologi secara serempak, sehingga guru menilai keluaran AI bukan hanya dari kerapian format, tetapi juga dari relevansinya terhadap tujuan belajar, karakter peserta didik, dan rancangan asesmen.

Tabel 6. Sintesis Tematik Efektivitas, Strategi *Prompt* , dan Risiko

Tema analitis	Pola temuan utama	Kecenderungan Scopus	Kecenderungan SINTA
Efektivitas	Efisiensi waktu, struktur awal, kelengkapan komponen, revisi, dan personalisasi berbasis konteks.	Lebih kuat pada eksperimen, analisis kualitas, dan teacher education.	Lebih kuat pada manfaat praktis guru, modul ajar, dan perangkat pembelajaran.

⁵¹Siregar et al., "Penggunaan ChatGPT untuk menyusun perangkat pembelajaran tingkat SMA" (2025).

⁵²Istiarsyah et al., "Transforming teacher competence through ChatGPT" (2024).

⁵³Sukmadana and Halim, "Perceptions of pre-service English teachers on using ChatGPT for lesson planning" (2026).

⁵⁴ElSayary et al., "Examining the role of prompt engineering in utilizing generative AI tools for lesson planning" (2025).

⁵⁵ElSayary, "An investigation of teachers' perceptions of using ChatGPT as a supporting tool for teaching and learning in the digital era" (2024).

⁵⁶Mishra and Koehler, "Technological pedagogical content knowledge" (2006).

Tema analitis	Pola temuan utama	Kecenderungan Scopus	Kecenderungan SINTA
Strategi <i>prompt</i>	<i>Prompt</i> kontekstual, berbasis peran, rubrik, iteratif, dan berbasis contoh memberi keluaran paling produktif.	Menawarkan framework dan evaluasi <i>prompt</i> lebih rinci.	Menunjukkan kebutuhan pelatihan <i>prompt</i> design dan adaptasi konteks lokal.
Risiko	Ketidakesesuaian pedagogik, halusinasi, bias, plagiarisme, ketergantungan guru, dan privasi-etika.	Menyoroti bias, kualitas keluaran, dan evaluasi kritis secara eksplisit.	Menekankan keselarasan kurikulum, konteks kebahasaan, dan pendampingan guru.
Kebaruan kontekstual	Pembacaan gabungan Scopus-SINTA mempertemukan kedalaman metodologis dan relevansi lokal.	Memperkuat kerangka <i>prompt engineering</i> dan evaluasi mutu.	Memperkuat implementasi pada modul ajar, Kurikulum Merdeka, dan kebutuhan guru Indonesia.

5. Perbandingan Kecenderungan Scopus dan SINTA serta Kebaruan Kontekstual

Perbandingan antara artikel Scopus dan artikel SINTA memperlihatkan dua corak yang saling menguatkan. Artikel Scopus cenderung lebih kaya pada sisi kerangka konseptual, desain penelitian, *prompt* framework, PCK, dan evaluasi kualitas keluaran. Artikel-artikel ini memperluas pemahaman tentang bagaimana *prompt* dirancang, bagaimana *lesson plan* dievaluasi, dan bagaimana interaksi teacher–AI membentuk mutu desain pembelajaran. Sementara itu, artikel SINTA memperlihatkan kedekatan yang lebih kuat dengan kebutuhan operasional guru Indonesia, khususnya penyusunan modul ajar, perangkat pembelajaran, kecocokan kurikulum, dan kebutuhan pelatihan *prompt* design.

Pertemuan dua horizon tersebut menghasilkan kebaruan kontekstual yang penting. Scopus memberi dasar metodologis untuk menilai kualitas AI generatif dalam *lesson planning*, sedangkan SINTA memberi gambaran implementasi pada realitas Indonesia. Melalui kombinasi ini, diskusi tentang *prompt engineering*, evaluasi mutu, dan risiko dapat langsung dihubungkan dengan kebutuhan guru di dalam ekosistem Kurikulum Merdeka, budaya belajar lokal, dan praktik penyusunan modul ajar di sekolah.

Kebaruan kontekstual kajian ini terletak pada pembacaan dua lapis yang jarang dipertemukan dalam satu sintesis, yaitu horizon teoritis-global dan horizon implementatif-lokal. Melalui pembacaan silang tersebut, artikel ini tidak hanya memetakan apa yang efektif secara umum, tetapi juga menjelaskan bagaimana temuan internasional dapat diterjemahkan menjadi praktik *lesson planning* yang relevan bagi guru Indonesia. Dengan demikian, kontribusi kajian bergerak dari sekadar ringkasan literatur menuju model pembacaan yang menghubungkan *prompt* design, mutu pedagogik, dan kebutuhan kurikulum lokal secara lebih operasional.

6. Implikasi Penelitian, Keterbatasan, dan Agenda Lanjutan

Implikasi praktis kajian ini mengarah pada penguatan literasi AI guru. Guru memperoleh manfaat paling tinggi dari AI generatif ketika mereka memiliki kemampuan merancang *prompt*, menilai kualitas keluaran, memeriksa akurasi isi, dan menyelaraskan hasil dengan kebutuhan peserta didik. Lembaga pendidikan dapat menggunakan temuan ini untuk menyusun pedoman penggunaan AI generatif pada aktivitas perencanaan pembelajaran yang mencakup batas aman data, validasi isi, transparansi penggunaan, dan tanggung jawab revisi. Pembuat kebijakan memperoleh dasar untuk merancang pelatihan yang bergerak melampaui pengenalan platform menuju *prompt design*, evaluasi keluaran, etika, dan akuntabilitas pedagogik.

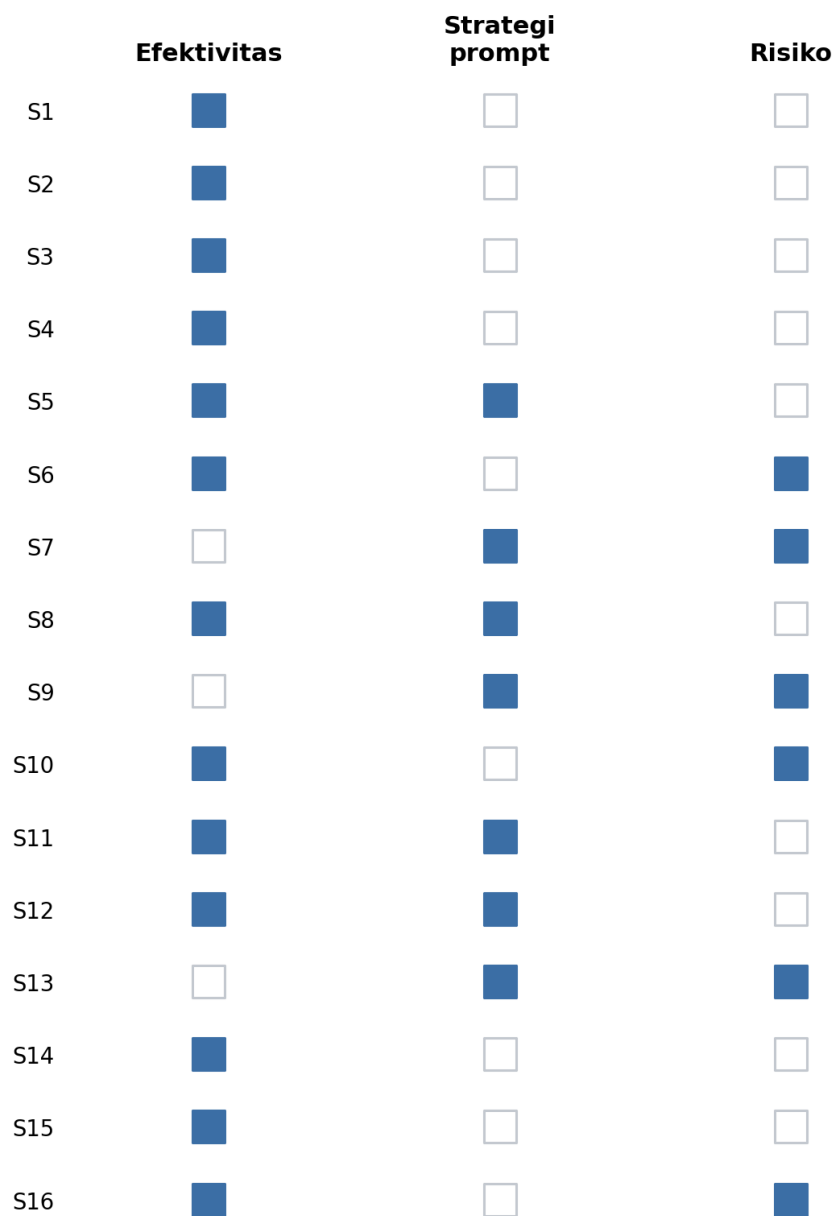
Keterbatasan kajian ini berhubungan dengan ruang basis data yang difokuskan pada Scopus dan SINTA, rentang waktu 2023-April 2026, serta heterogenitas desain penelitian yang membuat bentuk bukti bergerak pada intensitas yang beragam. Variasi istilah seperti *lesson planning*, *curriculum planning*, *teaching modules*, dan *AI-assisted planning* juga memberi tantangan pada sensitivitas pencarian. Selain itu, proses seleksi, appraisal, dan ekstraksi data dilakukan oleh *reviewer* tunggal karena keterbatasan sumber daya penelitian, sehingga *inter-rater reliability* tidak dapat disajikan. Untuk menjaga konsistensi, penulis menggunakan kriteria yang ditetapkan sejak awal, lembar seleksi terstruktur, dan pembacaan ulang pada kasus batas relevansi; meskipun demikian, ruang subjektivitas keputusan tetap perlu dibaca sebagai keterbatasan metodologis. Karakter AI generatif yang berkembang cepat turut membuat lanskap literatur bergerak dinamis, sehingga pembacaan sintesis perlu terus diperbarui sejalan dengan pertumbuhan model, platform, dan praktik baru.

Agenda penelitian lanjutan dapat diarahkan pada studi empiris yang mengamati guru saat menggunakan AI generatif untuk menyusun rencana pembelajaran secara langsung, eksperimen komparatif antarjenis *prompt*, kajian per jenjang pendidikan dan per mata pelajaran, integrasi AI dengan kurikulum lokal dan konteks madrasah, serta studi etika AI generatif yang memeriksa privasi, akuntabilitas, dan agency guru dalam proses perencanaan pembelajaran. Arah ini akan memperkaya pemahaman tentang bagaimana AI generatif dapat mendukung kualitas pedagogik secara semakin matang.

7. Kualitas Bukti dan Robustitas Sintesis

Untuk menjaga keterlacakan hasil individual dan kualitas studi tanpa memperpanjang tabel, Gambar 2 menyajikan peta bukti studi individual menurut tiga domain sintesis, sedangkan Gambar 3 merangkum distribusi appraisal kualitas, potensi bias, dan kelengkapan pelaporan. Kedua visual ini menunjukkan bahwa korpus didominasi studi memadai dengan kontribusi paling kuat pada efektivitas drafting awal dan *prompt* kontekstual-literatif.

Peta bukti studi individual menurut domain sintesis

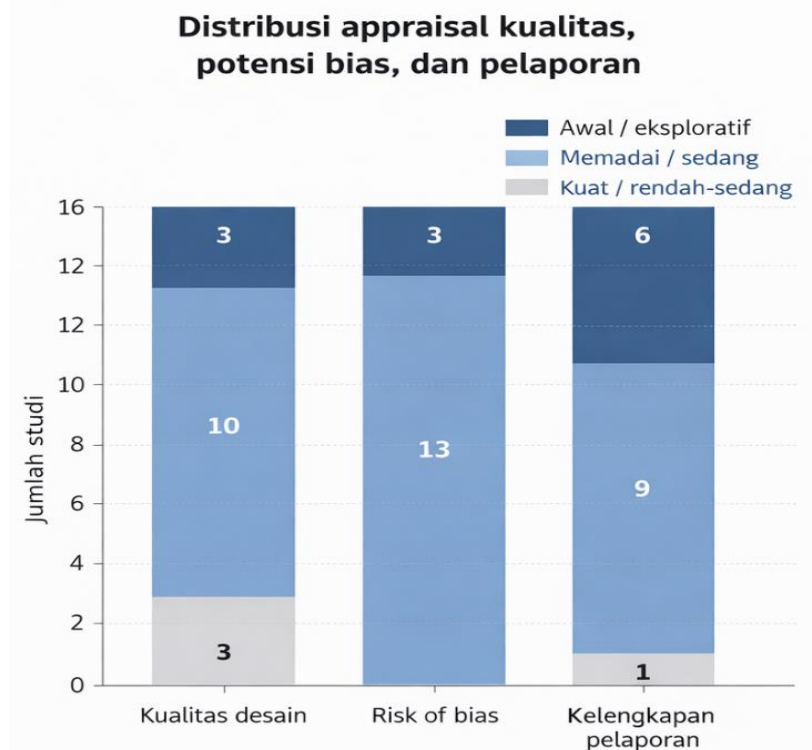


Kotak biru menandai kontribusi substantif studi pada domain terkait.

Gambar 2. Peta bukti studi individual menurut domain sintesis

Gambar 2 menunjukkan bahwa kontribusi bukti paling padat berada pada domain efektivitas, disusul strategi *prompt*, sedangkan domain risiko dihuni jumlah studi yang lebih terbatas. Pola ini menandakan bahwa literatur berkembang dari fokus awal pada kegunaan praktis AI generatif—terutama efisiensi drafting dan perbaikan struktur—lalu bergerak menuju perhatian yang lebih sistematis pada desain *prompt* dan konsekuensi pedagogiknya. Dengan

kata lain, peta bukti memperlihatkan bahwa pembahasan risiko mulai menguat, tetapi belum setebal bukti yang mendukung manfaat operasional AI dalam perencanaan pembelajaran. Gambar 3 memperlihatkan bahwa korpus didominasi studi dengan kualitas dan risiko bias pada tingkat memadai, sementara proporsi studi kuat sudah hadir tetapi belum menjadi mayoritas. Kelengkapan pelaporan tampak relatif lebih baik dibanding kekuatan desain, yang berarti banyak artikel telah cukup jelas dalam melaporkan proses dan temuan meskipun kedalaman rancangan penelitiannya masih beragam. Temuan ini memperkuat posisi sintesis sebagai dasar yang cukup andal untuk membaca kecenderungan utama, sekaligus menegaskan perlunya studi dengan desain yang lebih kokoh agar kepastian bukti meningkat pada tahap berikutnya.



Gambar 3. Distribusi appraisal kualitas, potensi bias, dan kelengkapan pelaporan

Tabel 7. Kepastian Bukti per Domain Utama

Domain utama	Kepastian	Interpretasi ringkas
Efektivitas drafting awal	Sedang-kuat	Paling konsisten pada efisiensi waktu, struktur awal, dan revisi.
<i>Prompt</i> kontekstual-iteratif	Kuat	Didukung lintas konteks dan paling stabil pada berbagai studi.
Ketidaksesuaian pedagogik	Sedang-kuat	Risiko paling dominan ketika keluaran diterima tanpa evaluasi guru.
Halusinasi, bias, dan etika	Sedang	Muncul berulang, tetapi intensitasnya bergantung pada konteks mata pelajaran.

Pembacaan kepekaan menunjukkan bahwa struktur tema tetap stabil ketika korpus penuh 16 studi dibandingkan dengan korpus inti 12 studi berkualitas kuat-memadai. Tema yang paling stabil ialah efisiensi waktu, struktur awal, *prompt* kontekstual-iteratif, dan risiko

ketidaksesuaian pedagogik, sedangkan personalisasi serta isu etika lebih sensitif terhadap variasi konteks.

D. Kesimpulan

Systematic Literature Review ini menunjukkan bahwa AI generatif menghadirkan nilai paling kuat dalam penyusunan rencana pembelajaran melalui efisiensi waktu, pembentukan struktur awal, kelengkapan komponen dasar, dan kemudahan revisi. Simpulan ini dibaca dengan mempertimbangkan batas ruang basis data, sifat reviewer tunggal pada proses seleksi dan ekstraksi, serta dinamika publikasi AI generatif yang berkembang sangat cepat. RQ1 terjawab dengan pola efektivitas yang paling stabil pada empat dimensi tersebut, sementara personalisasi berkembang paling optimal melalui kolaborasi intensif antara AI dan guru. RQ2 terjawab dengan dominasi *prompt* kontekstual, berbasis peran, berbasis rubrik, iteratif, dan berbasis contoh. RQ3 terjawab dengan penempatan ketidaksesuaian pedagogik, halusinasi informasi, dan bias sebagai tiga risiko yang paling menonjol, disusul plagiarisme, ketergantungan guru, serta privasi dan etika.

Secara teoretis, kajian ini menyatukan efektivitas, strategi *prompt*, dan risiko dalam satu kerangka yang utuh. Secara praktis, kajian ini memberi dasar bagi guru, peneliti, lembaga pendidikan, dan pembuat kebijakan untuk mengembangkan penggunaan AI generatif yang lebih cermat, produktif, dan bertanggung jawab. Posisi akhir artikel ini menempatkan AI generatif sebagai mitra perancangan pedagogik yang memberi manfaat tertinggi melalui *prompt* yang terarah, refleksi kritis, dan kepemimpinan profesional guru dalam setiap keputusan pembelajaran.

Pernyataan Pendanaan, Konflik Kepentingan, dan Ketersediaan Data

Artikel ini ditulis tanpa dukungan pendanaan eksternal. Penulis bekerja secara independen dan menempatkan integritas sintesis sebagai dasar interpretasi. Konflik kepentingan personal, institusional, dan komersial tidak memengaruhi proses review. Data ekstraksi, matriks sintesis, template appraisal, dan catatan keputusan seleksi tersedia melalui penulis korespondensi dalam permintaan akademik yang wajar.

Bibliography

- Al-khresheh, M. H. (2024). Bridging technology and pedagogy from a global lens: Teachers' perspectives on integrating ChatGPT in English language teaching. *Computers and Education: Artificial Intelligence*, 6, 100218. <https://doi.org/10.1016/j.caeai.2024.100218>
- Hashem, R., Ali, N., El Zein, F., Fidalgo, P., & Khurma, O. A. (2024). AI to the rescue: Exploring the potential of ChatGPT as a teacher ally for workload relief and burnout prevention. *Research and Practice in Technology Enhanced Learning*, 19, 1–29. <https://doi.org/10.58459/rptel.2024.19016>
- Tran, P. T. H. (2025). Preparing preservice teachers for generative AI in *lesson planning*: Comparing AI mindset and tools-only training approaches. *Journal of Technology and Teacher Education*, 33(1), 45–71. <https://doi.org/10.1080/21532974.2025.2583516>
- Baytak, A. (2024). The content analysis of the *lesson plans* created by ChatGPT and Google Gemini. *Research in Social Sciences and Technology*, 9(1), 329–350. <https://doi.org/10.46303/ressat.2024.19>

- Biberman-Shalev, L. (2025). *Prompt* ing theory into practice: Utilizing ChatGPT-4 in a curriculum planning course. *Education Sciences*, 15(2), 196. <https://doi.org/10.3390/educsci15020196>
- Dornburg, A., & Davin, K. J. (2025). ChatGPT in foreign language *lesson plan* creation: Trends, variability, and historical biases. *ReCALL*, 37(3), 332–347. <https://doi.org/10.1017/S0958344024000272>
- ElSayary, A. (2024). An investigation of teachers' perceptions of using ChatGPT as a supporting tool for teaching and learning in the digital era. *Journal of Computer Assisted Learning*, 40(3), 931–945. <https://doi.org/10.1111/jcal.12926>
- ElSayary, A., Kuhail, M. A., & Hojeij, Z. (2025). Examining the role of *prompt* engineering in utilizing generative AI tools for *lesson planning*: Insights from teachers' experiences and perceptions. *Human Behavior and Emerging Technologies*, 2025(1), 9986139. <https://doi.org/10.1155/hbe2/9986139>
- Gurl, T. J., Markinson, M. P., & Artzt, A. F. (2025). Using ChatGPT as a *lesson planning* assistant with preservice secondary mathematics teachers. *Digital Experiences in Mathematics Education*, 11(1), 114–139. <https://doi.org/10.1007/s40751-024-00162-9>
- Hashem, R., Ali, N., El Zein, F., Fidalgo, P., & Abu Khurma, O. (2024). AI to the rescue: Exploring the potential of ChatGPT as a teacher ally for workload relief and burnout prevention. *Research and Practice in Technology Enhanced Learning*, 19, 23. <https://doi.org/10.58459/rptel.2024.19023>
- Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., ... Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information*, 34(4), 285–291. <https://doi.org/10.3233/EFI-180221>
- Istiaryah, I., Maisura, M., Nurhabibah, N., Kumita, K., & Kamarullah, K. (2024). Transforming teacher competence through ChatGPT: Designing innovative and effective lesson modules. *Ideguru: Jurnal Karya Ilmiah Guru*, 10(1), 505–513. <https://doi.org/10.51169/ideguru.v10i1.1586>
- Kalenda, P. J., Rath, L., Abugasea Heidt, M., & Wright, A. (2025). Pre-service teacher perceptions of ChatGPT for *lesson plan* generation. *Journal of Educational Technology Systems*, 53(3), 219–241. <https://doi.org/10.1177/00472395241301388>
- Karaman, M. R., & Göksu, İ. (2024). Are *lesson plans* created by ChatGPT more effective? An experimental study. *International Journal of Technology in Education*, 7(1), 107–127. <https://doi.org/10.46328/ijte.607>
- Luo, Z., & Tahir, R. (2025). ChatGPT-assisted *lesson planning* for children's STEAM arts education: An experimental study on benefits, challenges, methods, and a *prompt* framework. *International Journal of Artificial Intelligence in Education*, 35, 3696–3745. <https://doi.org/10.1007/s40593-025-00513-6>
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Moundridou, M., Matzakos, N., & Doukakis, S. (2024). Generative AI tools as educators' assistants: Designing and implementing inquiry-based *lesson plans*. *Computers and Education: Artificial Intelligence*, 7, 100277. <https://doi.org/10.1016/j.caeai.2024.100277>

- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Peikos, G., & Stavrou, D. (2025). ChatGPT for science *lesson planning*: An exploratory study based on pedagogical content knowledge. *Education Sciences*, 15(3), 338. <https://doi.org/10.3390/educsci15030338>
- Siregar, A. M., Kurniawan, A., Uyun, M. H., & Gea, D. P. (2025). Penggunaan ChatGPT untuk menyusun perangkat pembelajaran tingkat SMA. *Jurnal Bhakti Karya dan Inovatif*, 5(2), 156–164. <https://doi.org/10.37278/bhaktikaryadaninovatif.v5i2.1326>
- Sukmadana, S. R., & Halim, A. (2026). Perceptions of pre-service English teachers on using ChatGPT for *lesson planning*. *RIGGS: Journal of Artificial Intelligence and Digital Business*, 5(1), 2716–2725. <https://doi.org/10.31004/riggs.v5i1.5812>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. <https://doi.org/10.1111/j.1551-6708.1988.tb00823.x>
- Şimşek, N. (2025). Integration of ChatGPT in mathematical story-focused 5E *lesson planning*: Teachers and pre-service teachers' interactions with ChatGPT. *Education and Information Technologies*, 30(8), 11391–11462. <https://doi.org/10.1007/s10639-024-13258-x>
- van den Berg, G., & du Plessis, E. (2023). ChatGPT and generative AI: Possibilities for its contribution to *lesson planning*, critical thinking and openness in teacher education. *Education Sciences*, 13(10), 998. <https://doi.org/10.3390/educsci13100998>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Cambridge, MA: Harvard University Press.