

MERANCANG ASESMEN PAI DI ERA GENERATIVE AI: PERGESERAN DARI PENILAIAN PRODUK MENUJU PENILAIAN PROSES, REASONING, DAN AUTHENTICITY SISWA

Abdul Hakim, IAI Al Khoziny Buduran, Sidoarjo
e-Mail: abd.hakim@gmail.com

Abstrak

Kemunculan generative artificial intelligence (GenAI) seperti ChatGPT, Gemini, dan berbagai large language models (LLM) telah mengubah lanskap pembelajaran dan asesmen pendidikan. Dalam Pendidikan Agama Islam (PAI), perubahan tersebut menghadirkan persoalan yang lebih kompleks karena asesmen tidak hanya ditujukan untuk mengukur penguasaan pengetahuan, tetapi juga kemampuan bernalar, penghayatan nilai, sikap, dan kemampuan menerapkan ajaran Islam dalam konteks kehidupan nyata. Penelitian ini bertujuan menganalisis kebutuhan transformasi asesmen PAI pada era GenAI dan merumuskan model asesmen yang lebih berorientasi pada proses, reasoning, dan authenticity peserta didik. Penelitian menggunakan pendekatan kualitatif dengan metode studi kepustakaan. Data diperoleh melalui analisis artikel jurnal ilmiah, buku akademik, dokumen kebijakan pendidikan, dan literatur tentang asesmen autentik serta GenAI dalam pendidikan. Data dianalisis secara tematik melalui tahap identifikasi, kategorisasi, komparasi, interpretasi, dan sintesis. Hasil kajian menunjukkan bahwa asesmen berbasis produk tunggal semakin rentan kehilangan validitas sebagai bukti kemampuan individual ketika proses pengerjaannya dapat dialihkan kepada GenAI. Oleh karena itu, asesmen PAI perlu menggabungkan bukti proses, argumentasi atau reasoning, refleksi personal, performa autentik, dialog lisan, dan transparansi penggunaan AI. Artikel ini menawarkan model Process-Reasoning-Authenticity (PRA) yang menempatkan produk sebagai salah satu bukti, bukan satu-satunya dasar penilaian. Model tersebut dapat menjadi kerangka praktis bagi guru PAI untuk merancang asesmen yang tetap valid, adil, reflektif, dan berorientasi pada pembentukan karakter.

Kata kunci: asesmen PAI; generative AI; reasoning; authentic assessment; academic integrity

Abstract

The emergence of generative artificial intelligence (GenAI), including ChatGPT, Gemini, and other large language models (LLMs), has transformed the landscape of teaching, learning, and assessment. In Islamic Religious Education (PAI), this transformation presents a more complex challenge because assessment is expected not only to measure students' knowledge but also their reasoning, internalization of Islamic values, attitudes, and ability to apply religious principles in real-life contexts. This study aims to analyze the need to transform PAI assessment in the GenAI era and to formulate an assessment model oriented toward process, reasoning, and student authenticity. A qualitative library research

approach was employed by examining scholarly journal articles, academic books, educational policy documents, and literature on authentic assessment and GenAI in education. Data were analyzed thematically through identification, categorization, comparison, interpretation, and synthesis. The findings indicate that single-product assessment is increasingly vulnerable to losing its validity as evidence of individual competence when the production process can be delegated to GenAI. Therefore, PAI assessment needs to integrate evidence of learning processes, reasoning, personal reflection, authentic performance, oral dialogue, and transparency regarding AI use. This study proposes the Process–Reasoning–Authenticity (PRA) model, positioning the final product as one form of evidence rather than the sole basis for assessment. The model provides a practical framework for PAI teachers to design assessment that remains valid, fair, reflective, and oriented toward character development.

Keywords: *PAI assessment; generative AI; reasoning; authentic assessment; academic integrity*

Pendahuluan

Perkembangan *generative artificial intelligence* (GenAI) dalam beberapa tahun terakhir telah membawa perubahan signifikan terhadap ekosistem pendidikan, khususnya dalam cara peserta didik memperoleh informasi, menyusun gagasan, memproduksi teks, dan menyelesaikan tugas akademik. Kehadiran aplikasi berbasis *large language model* (LLM), seperti ChatGPT dan berbagai platform sejenis, memungkinkan peserta didik memperoleh bantuan dalam merumuskan ide, menyusun argumentasi, meringkas informasi, menerjemahkan teks, hingga menghasilkan produk akademik dalam waktu yang relatif singkat. Perubahan tersebut memberikan peluang bagi pembelajaran yang lebih personal, adaptif, dan efisien, tetapi pada saat yang sama memunculkan persoalan

mendasar mengenai validitas bukti yang digunakan dalam asesmen. UNESCO menegaskan bahwa perkembangan GenAI dalam pendidikan perlu direspons melalui pendekatan yang berpusat pada manusia dengan tetap mempertahankan agensi peserta didik, transparansi, keamanan, dan tanggung jawab etis dalam penggunaan teknologi (Miao & Holmes, 2023).

Persoalan tersebut menjadi semakin kompleks ketika asesmen masih berorientasi pada produk akhir sebagai bukti utama pencapaian belajar. Esai, makalah, rangkuman, jawaban uraian, presentasi, dan berbagai bentuk tugas tertulis selama ini banyak digunakan untuk mengukur penguasaan pengetahuan dan kemampuan peserta didik. Dalam desain asesmen konvensional, produk akhir relatif mudah diperlakukan sebagai representasi kemampuan

individual siswa. Namun, asumsi tersebut menjadi semakin problematis ketika produk tersebut dapat dihasilkan, diperbaiki, atau dikembangkan dengan bantuan *GenAI*. Cotton et al. (2024) menunjukkan bahwa kemunculan *ChatGPT* menimbulkan persoalan baru terkait integritas akademik karena teknologi tersebut dapat menghasilkan teks yang menyerupai karya manusia. Dengan demikian, persoalan utama asesmen di era *GenAI* bukan sekadar apakah peserta didik menggunakan AI, melainkan apakah produk yang dinilai masih memberikan bukti yang memadai mengenai kompetensi yang sebenarnya hendak diukur.

Dalam perspektif pengukuran pendidikan, persoalan tersebut berkaitan erat dengan validitas asesmen. Asesmen dikatakan bermakna apabila bukti yang diperoleh benar-benar dapat digunakan untuk mendukung interpretasi mengenai kemampuan peserta didik. Hao et al. (2024) menjelaskan bahwa perkembangan *large language models* dan *GenAI* menimbulkan implikasi terhadap validitas, reliabilitas, fairness, transparansi, dan keamanan asesmen. Ketika guru hanya menilai produk akhir, terdapat kemungkinan bahwa skor yang diperoleh lebih merepresentasikan kemampuan teknologi dalam menghasilkan produk daripada kemampuan peserta didik dalam memahami, menganalisis, dan mengonstruksi pengetahuan. Dengan

kata lain, perubahan teknologi telah mengubah hubungan antara produk asesmen dan konstruk yang hendak diukur.

Kondisi tersebut mengharuskan pendidik melakukan pergeseran paradigma dari *product-centred assessment* menuju asesmen yang menghasilkan bukti pembelajaran secara lebih komprehensif. Bukti pembelajaran tidak cukup diperoleh dari jawaban akhir, tetapi perlu mencakup bagaimana peserta didik mengidentifikasi masalah, mencari dan mengevaluasi informasi, mengembangkan argumentasi, melakukan revisi, menerima umpan balik, serta merefleksikan keputusan yang dibuat. Lin dan Chiu (2024), melalui kajian terhadap perkembangan asesmen dalam konteks pendidikan tinggi, menunjukkan bahwa *GenAI* mendorong kebutuhan untuk mendesain kembali asesmen agar lebih menekankan *higher-order thinking*, regulasi diri, tanggung jawab, dan integritas akademik. Perubahan tersebut menunjukkan bahwa asesmen tidak lagi cukup dipahami sebagai instrumen untuk memberikan nilai, tetapi sebagai mekanisme untuk memperoleh bukti mengenai proses dan kualitas belajar peserta didik.

Persoalan transformasi asesmen tersebut memiliki relevansi khusus dalam Pendidikan Agama Islam (PAI). Tujuan PAI tidak terbatas pada penguasaan pengetahuan keagamaan, tetapi mencakup pembentukan sikap,

karakter, kemampuan bernalar secara etis, serta kemampuan menerapkan nilai-nilai Islam dalam kehidupan sehari-hari. Oleh karena itu, keberhasilan pembelajaran PAI tidak dapat sepenuhnya direpresentasikan melalui kemampuan peserta didik menghasilkan jawaban yang benar dalam tes atau tulisan yang secara linguistik baik. Peserta didik juga perlu menunjukkan kemampuan memahami makna ajaran, menghubungkan dalil dengan persoalan kehidupan, mempertimbangkan berbagai alternatif, mengambil keputusan berdasarkan prinsip Islam, dan merefleksikan konsekuensi dari keputusan tersebut.

Dalam konteks tersebut, asesmen PAI menghadapi tantangan ganda. Di satu sisi, guru perlu memastikan penguasaan pengetahuan keislaman peserta didik. Di sisi lain, guru perlu memperoleh bukti bahwa pengetahuan tersebut benar-benar dipahami dan dapat digunakan secara bertanggung jawab. Tantangan tersebut semakin besar ketika GenAI mampu memberikan jawaban yang tampak sistematis mengenai persoalan keagamaan. Teknologi dapat menghasilkan penjelasan mengenai ayat, hadis, fikih, akhlak, atau sejarah Islam, tetapi kualitas dan kebenaran informasi yang dihasilkan tetap perlu diverifikasi. Penggunaan GenAI dalam konteks pendidikan agama karena itu tidak dapat dipisahkan dari persoalan literasi digital, literasi informasi,

kemampuan berpikir kritis, dan kemampuan melakukan verifikasi terhadap sumber keagamaan.

Kajian mengenai AI dalam pendidikan menunjukkan bahwa penggunaan teknologi generatif tidak semestinya hanya dipandang sebagai ancaman terhadap integritas akademik. Tlili et al. (2023) menunjukkan bahwa ChatGPT dapat memiliki potensi pedagogis dalam pembelajaran sekaligus menimbulkan tantangan terkait akurasi, bias, privasi, dan integritas akademik. Perspektif yang sama dikembangkan oleh Miao dan Holmes (2023) yang menekankan pentingnya penggunaan GenAI secara manusiawi, aman, adil, dan bertanggung jawab. Dengan demikian, respons pendidikan terhadap GenAI tidak cukup berupa pelarangan penggunaan teknologi, tetapi perlu diarahkan pada desain pembelajaran dan asesmen yang mampu memastikan bahwa teknologi tetap berada di bawah kendali dan tanggung jawab peserta didik.

Dalam konteks Indonesia, arah kebijakan pendidikan juga semakin menekankan pentingnya asesmen sebagai bagian integral dari proses pembelajaran. Prinsip asesmen dalam implementasi kurikulum menempatkan asesmen sebagai proses memperoleh informasi mengenai pencapaian belajar yang digunakan untuk memberikan umpan balik dan memperbaiki pembelajaran. Asesmen juga perlu memperhatikan prinsip

validitas, keadilan, objektivitas, fleksibilitas, dan kebermaknaan. Perspektif tersebut memberikan landasan bahwa asesmen seharusnya tidak hanya berfungsi untuk menentukan nilai akhir, tetapi juga memberikan gambaran mengenai perkembangan kompetensi peserta didik. Dalam konteks era GenAI, prinsip tersebut menjadi semakin relevan karena guru perlu memperoleh bukti pembelajaran yang tidak mudah direduksi menjadi sekadar produk yang dihasilkan dengan bantuan teknologi.

Salah satu pendekatan yang dapat digunakan untuk merespons persoalan tersebut adalah *authentic assessment*. Penilaian autentik menempatkan peserta didik pada tugas yang memungkinkan mereka menggunakan pengetahuan dan keterampilan dalam situasi yang memiliki relevansi dengan kehidupan nyata. Dalam pembelajaran, asesmen autentik dapat digunakan untuk menilai kemampuan berpikir tingkat tinggi melalui tugas yang menuntut penerapan, analisis, evaluasi, dan penciptaan. Muthohharoh et al. (2020), misalnya, menunjukkan bahwa penerapan *authentic assessment* dapat digunakan untuk mengidentifikasi kemampuan berpikir tingkat tinggi peserta didik. Namun, implementasi asesmen autentik juga membutuhkan kemampuan guru dalam merancang indikator, instrumen, rubrik, serta prosedur penilaian yang sesuai dengan kompetensi yang hendak diukur.

Dalam era GenAI, konsep asesmen autentik perlu diperluas. Autentisitas tidak cukup dipahami sebagai keadaan ketika peserta didik mengerjakan tugas tanpa bantuan teknologi. Definisi semacam itu berpotensi menyederhanakan persoalan dan tidak sesuai dengan realitas lingkungan belajar digital. Autentisitas lebih tepat dipahami sebagai kemampuan peserta didik untuk menunjukkan keterlibatan intelektual, pertimbangan, refleksi, dan tanggung jawab terhadap proses serta hasil belajar. Dengan demikian, peserta didik masih dapat menggunakan GenAI dalam kondisi tertentu, tetapi penggunaan tersebut harus transparan dan tidak menggantikan proses berpikir yang menjadi konstruk utama asesmen.

Perspektif tersebut mengarahkan perhatian pada pentingnya *reasoning* sebagai salah satu bukti utama dalam asesmen PAI. Reasoning tidak sekadar berarti memberikan jawaban yang benar, tetapi kemampuan menjelaskan bagaimana suatu kesimpulan diperoleh, bukti apa yang digunakan, mengapa suatu sumber dianggap relevan, bagaimana alternatif dipertimbangkan, serta bagaimana keputusan dikaitkan dengan prinsip tertentu. Dalam PAI, kemampuan tersebut sangat penting karena peserta didik sering berhadapan dengan persoalan yang membutuhkan interpretasi dan pertimbangan etis,

seperti penggunaan media sosial, transaksi digital, pergaulan, toleransi, lingkungan, konsumsi, dan berbagai persoalan moral kontemporer.

Dengan memasukkan reasoning ke dalam asesmen, guru tidak hanya menilai apa yang diketahui siswa, tetapi juga bagaimana siswa menggunakan pengetahuannya untuk berpikir dan mengambil keputusan. Perubahan ini memiliki implikasi penting terhadap penggunaan GenAI. Apabila peserta didik hanya diminta menghasilkan teks, GenAI dapat mengambil alih sebagian besar proses produksi jawaban. Sebaliknya, apabila peserta didik harus mempertanggungjawabkan argumentasinya melalui jurnal proses, refleksi, diskusi, presentasi, atau dialog lisan, maka guru memperoleh bukti tambahan mengenai pemahaman individual siswa.

Sejumlah penelitian terdahulu telah membahas hubungan antara GenAI dan asesmen. Fischer et al. (2024) menunjukkan bahwa penggunaan GenAI dalam tugas akademik dapat menimbulkan perbedaan antara kualitas produk yang dihasilkan dan kemampuan aktual peserta didik sebagaimana dipersepsikan melalui proses asesmen. Lin dan Chiu (2024) juga menunjukkan bahwa transformasi asesmen akibat GenAI membutuhkan pendekatan yang lebih menekankan keterampilan tingkat tinggi, pembelajaran regulatif, dan integritas. Sementara itu, Hao et al. (2024) menekankan perlunya

mempertimbangkan kembali validitas dan prinsip pengukuran ketika LLM digunakan dalam konteks asesmen. Literatur tersebut memberikan dasar kuat bahwa desain asesmen perlu beradaptasi terhadap perubahan teknologi.

Di Indonesia, kajian mengenai pemanfaatan kecerdasan buatan dalam pembelajaran PAI mulai berkembang. Penelitian mengenai pengembangan penilaian pembelajaran PAI berbasis kecerdasan buatan menunjukkan adanya peluang untuk meningkatkan efisiensi evaluasi, personalisasi pembelajaran, dan pemberian umpan balik, tetapi juga menghadirkan tantangan berkaitan dengan kesiapan guru, kualitas teknologi, dan integritas asesmen (Sofyan & Salito, 2024). Penelitian lain mengenai optimalisasi AI dalam Pendidikan Agama Islam juga menunjukkan bahwa teknologi dapat mendukung proses evaluasi, tetapi implementasinya perlu disertai penguatan kompetensi digital dan pedagogis guru (Rahmawati et al., 2025). Temuan tersebut menunjukkan bahwa diskursus AI dalam PAI di Indonesia telah berkembang, tetapi sebagian kajian masih lebih banyak berfokus pada pemanfaatan AI sebagai alat evaluasi daripada pada rekonstruksi desain asesmen akibat perubahan karakteristik lingkungan belajar.

Di sinilah terdapat *research gap* yang menjadi dasar penelitian ini. Pertama, sebagian penelitian mengenai

AI dalam pendidikan berfokus pada efisiensi penggunaan AI sebagai alat bantu pembelajaran atau penilaian, sedangkan persoalan mengenai bagaimana guru memperoleh bukti autentik tentang kompetensi siswa di tengah kemampuan GenAI menghasilkan produk akademik masih membutuhkan elaborasi lebih lanjut. Kedua, kajian mengenai asesmen autentik telah berkembang sebelum munculnya GenAI secara masif, tetapi belum seluruhnya menjawab bagaimana konsep autentisitas perlu ditafsirkan ulang ketika proses produksi pengetahuan dapat dibantu oleh mesin. Ketiga, dalam konteks PAI, masih terbuka ruang untuk mengintegrasikan proses belajar, *reasoning*, dan *authenticity* menjadi satu kerangka desain asesmen yang secara khusus merespons karakteristik GenAI.

Berdasarkan kesenjangan tersebut, artikel ini menawarkan kebaruan konseptual melalui model *Process-Reasoning-Authenticity (PRA)*. Model ini memandang bahwa kualitas asesmen PAI di era GenAI tidak seharusnya bertumpu pada produk akhir semata, tetapi pada kombinasi beberapa bukti: proses pengerjaan, kemampuan bernalar, autentisitas penerapan, kualitas produk, serta refleksi dan transparansi penggunaan AI. Dalam kerangka ini, produk akhir tetap memiliki posisi penting, tetapi tidak menjadi satu-satunya dasar untuk membuat keputusan asesmen.

Model PRA dibangun atas asumsi bahwa semakin kompleks teknologi yang tersedia bagi peserta didik, semakin penting bagi asesmen untuk mengumpulkan *evidence of learning* yang beragam. Bukti tersebut dapat diperoleh melalui portofolio proses, draft dan revisi, jurnal reflektif, analisis kasus, presentasi, dialog lisan, proyek autentik, serta pernyataan penggunaan AI. Pendekatan tersebut memungkinkan guru menggeser orientasi dari sekadar mencari bukti "apakah siswa menggunakan AI" menuju pertanyaan yang lebih pedagogis, yaitu "apakah siswa memahami, menalar, merefleksikan, menerapkan, dan mampu mempertanggungjawabkan hasil belajarnya?"

Berdasarkan uraian tersebut, penelitian ini berangkat dari tiga pertanyaan utama. Pertama, bagaimana perkembangan GenAI memengaruhi validitas asesmen PAI yang masih berorientasi pada produk akhir? Kedua, bagaimana konsep proses, reasoning, dan authenticity dapat digunakan sebagai sumber bukti pembelajaran yang lebih kuat? Ketiga, bagaimana model asesmen PAI berbasis *Process-Reasoning-Authenticity (PRA)* dapat dirancang sebagai respons pedagogis terhadap era GenAI?

Penelitian ini bertujuan untuk menganalisis kebutuhan transformasi asesmen PAI di era GenAI, mengidentifikasi karakteristik asesmen yang mampu menghasilkan bukti

pembelajaran yang lebih autentik, serta merumuskan model konseptual *Process-Reasoning-Authenticity (PRA)* yang dapat digunakan sebagai kerangka praktis dalam perancangan asesmen PAI. Secara teoretis, artikel ini diharapkan memperluas diskursus mengenai asesmen autentik dan validitas bukti pembelajaran dalam lingkungan yang semakin dipengaruhi teknologi generatif. Secara praktis, model yang ditawarkan diharapkan dapat membantu guru PAI merancang asesmen yang tidak hanya menilai produk akhir, tetapi juga mampu mengungkap proses berpikir, argumentasi, refleksi, penerapan nilai, dan tanggung jawab peserta didik dalam menggunakan teknologi AI.

Metode Penelitian

Penelitian ini menggunakan pendekatan kualitatif dengan metode studi kepustakaan. Pemilihan metode ini didasarkan pada tujuan penelitian yang berfokus pada analisis konseptual mengenai transformasi asesmen Pendidikan Agama Islam (PAI) di tengah perkembangan *generative artificial intelligence* (GenAI), khususnya pergeseran dari penilaian produk menuju penilaian proses, *reasoning*, dan *authenticity*. Studi kepustakaan memungkinkan peneliti melakukan sintesis kritis terhadap berbagai hasil penelitian, teori asesmen, kajian tentang kecerdasan buatan dalam pendidikan, serta dokumen kebijakan yang relevan

dengan konteks pendidikan Indonesia. Pendekatan kualitatif digunakan karena penelitian tidak diarahkan untuk menguji hipotesis secara statistik, melainkan untuk memahami perubahan konseptual dan pedagogis yang muncul akibat perkembangan GenAI serta merumuskan model asesmen yang relevan dengan kebutuhan pembelajaran PAI.

Sumber Data

Sumber data penelitian berupa literatur ilmiah dan dokumen akademik yang memiliki relevansi langsung dengan fokus penelitian. Literatur yang dianalisis meliputi artikel jurnal nasional dan internasional, buku ilmiah, laporan penelitian, dokumen kebijakan pendidikan, serta publikasi lembaga internasional yang membahas GenAI, asesmen pendidikan, *authentic assessment*, *assessment literacy*, *reasoning*, *higher-order thinking skills* (HOTS), integritas akademik, dan Pendidikan Agama Islam.

Literatur diprioritaskan pada publikasi dalam rentang 2016–2026 agar kajian mencerminkan perkembangan mutakhir dalam asesmen dan teknologi pendidikan. Meskipun demikian, beberapa sumber yang lebih lama tetap digunakan apabila memiliki posisi fundamental sebagai landasan teori atau konsep utama yang masih relevan. Sumber dari konteks Indonesia digunakan secara khusus untuk memperkuat keterkaitan analisis dengan kebijakan, karakteristik

pembelajaran, dan praktik asesmen di sekolah dan madrasah.

Pemilihan sumber dilakukan berdasarkan empat pertimbangan utama, yaitu: (1) relevansi dengan fokus penelitian; (2) kredibilitas akademik penerbit atau jurnal; (3) kontribusi konseptual atau empiris terhadap pembahasan; dan (4) keterkinian sumber, terutama untuk literatur mengenai GenAI yang mengalami perkembangan sangat cepat. Literatur yang hanya bersifat opini tanpa dasar akademik, sumber anonim, blog nonakademik, serta informasi yang tidak dapat diverifikasi tidak dijadikan sebagai dasar utama argumentasi.

Teknik Pengumpulan Data

Pengumpulan data dilakukan melalui penelusuran dan dokumentasi literatur secara sistematis. Penelusuran diarahkan pada beberapa kelompok kata kunci, yaitu *generative AI and education, generative AI and assessment, AI and assessment validity, authentic assessment, reasoning assessment, assessment literacy, academic integrity and ChatGPT, Islamic religious education and artificial intelligence, PAI assessment*, dan *AI in Islamic education*. Kata kunci tersebut dikombinasikan untuk memperoleh literatur yang tidak hanya membahas teknologi AI, tetapi juga menghubungkannya dengan persoalan pedagogis dan asesmen.

Literatur yang ditemukan kemudian diseleksi berdasarkan judul,

abstrak, kata kunci, dan relevansi substansi. Sumber yang memiliki keterkaitan langsung dengan persoalan penelitian dibaca secara lebih mendalam untuk mengidentifikasi konsep, temuan, argumentasi, dan rekomendasi yang dapat digunakan dalam penyusunan sintesis. Setiap sumber yang relevan didokumentasikan berdasarkan informasi bibliografis, fokus kajian, metode penelitian, temuan utama, keterbatasan, dan relevansinya terhadap pengembangan model asesmen PAI.

Proses pengumpulan data dalam penelitian kepustakaan tidak berhenti pada kegiatan mengumpulkan referensi, tetapi dilanjutkan dengan proses ekstraksi informasi. Informasi yang diekstraksi dikelompokkan ke dalam beberapa kategori, yaitu: (1) perubahan karakteristik asesmen akibat GenAI; (2) persoalan validitas dan integritas akademik; (3) asesmen autentik; (4) proses pembelajaran sebagai bukti asesmen; (5) *reasoning* dan kemampuan berpikir tingkat tinggi; (6) *authenticity* dan refleksi; (7) penggunaan AI secara transparan dan bertanggung jawab; serta (8) konteks implementasi asesmen PAI di Indonesia.

Teknik Analisis Data

Analisis data dilakukan secara analisis tematik (*thematic analysis*). Analisis ini digunakan untuk menemukan pola, tema, hubungan

antarkonsep, serta kesenjangan penelitian dari berbagai literatur yang dikaji. Analisis dilakukan secara bertahap melalui proses familiarisasi terhadap sumber, identifikasi informasi penting, pemberian kode, pengelompokan kode ke dalam tema, interpretasi hubungan antartema, dan penyusunan sintesis konseptual.

Tahap pertama adalah identifikasi isu, yaitu mengidentifikasi persoalan utama yang muncul dalam literatur mengenai penggunaan GenAI dan perubahan asesmen. Pada tahap ini perhatian diarahkan pada pertanyaan mengenai apakah produk akhir masih dapat menjadi bukti yang memadai terhadap kompetensi individual peserta didik ketika proses produksi tugas dapat dibantu oleh AI.

Tahap kedua adalah kategorisasi tematik. Temuan dari berbagai literatur dikelompokkan ke dalam tema-tema yang memiliki kesamaan konseptual. Tema utama yang diperoleh meliputi *product-centred assessment*, *process evidence*, *reasoning*, *authentic assessment*, *academic integrity*, *AI literacy*, dan *human agency*. Kategorisasi tersebut digunakan untuk melihat hubungan antara perkembangan teknologi dan kebutuhan perubahan desain asesmen.

Tahap ketiga adalah analisis komparatif. Temuan penelitian terdahulu dibandingkan untuk mengidentifikasi persamaan, perbedaan, keterbatasan, serta kecenderungan argumentasi. Analisis

komparatif diperlukan agar penelitian tidak sekadar menyusun ringkasan literatur, tetapi menghasilkan pembacaan kritis mengenai perkembangan asesmen pada era GenAI.

Tahap keempat adalah sintesis konseptual. Hasil analisis berbagai tema kemudian diintegrasikan untuk merumuskan hubungan antara proses pembelajaran, *reasoning*, dan *authenticity*. Sintesis tersebut menjadi dasar pengembangan model *Process-Reasoning-Authenticity (PRA)* sebagai kerangka konseptual asesmen PAI pada era GenAI.

Secara analitis, model yang dikembangkan berangkat dari proposisi bahwa kualitas bukti asesmen dapat diperkuat melalui kombinasi:

Bukti Pembelajaran = *Process + Reasoning + Authenticity + Product + Reflection*

Produk akhir tetap dipertahankan sebagai bagian dari bukti asesmen, tetapi tidak ditempatkan sebagai satu-satunya indikator kompetensi. Proses digunakan untuk menunjukkan perkembangan dan keterlibatan peserta didik, *reasoning* digunakan untuk melihat kualitas penalaran dan pengambilan keputusan, sedangkan *authenticity* digunakan untuk melihat kemampuan menerapkan pengetahuan dalam konteks yang bermakna.

Keabsahan Data

Keabsahan data dalam penelitian kepustakaan dijaga melalui *triangulasi* sumber. Triangulasi dilakukan dengan membandingkan informasi yang diperoleh dari artikel jurnal, buku ilmiah, dokumen kebijakan, dan publikasi lembaga yang memiliki kredibilitas akademik. Suatu argumentasi tidak digunakan sebagai dasar utama apabila hanya ditemukan dalam satu sumber yang tidak memiliki dukungan dari literatur lain.

Selain triangulasi sumber, dilakukan pula pemeriksaan konsistensi konseptual dengan membandingkan definisi dan penggunaan istilah seperti *generative AI*, *authentic assessment*, *reasoning*, *academic integrity*, dan *assessment validity*. Langkah tersebut diperlukan karena istilah-istilah tersebut digunakan dalam berbagai disiplin dengan penekanan konseptual yang berbeda.

Kredibilitas sintesis juga dijaga dengan membedakan secara tegas antara temuan penelitian terdahulu, interpretasi peneliti, dan proposisi konseptual yang dikembangkan dalam artikel. Dengan cara tersebut, model PRA tidak diposisikan sebagai temuan empiris, melainkan sebagai hasil sintesis konseptual yang perlu diuji lebih lanjut melalui penelitian lapangan.

Prosedur Pengembangan Model Konseptual

Pengembangan model PRA dilakukan melalui empat tahapan. Pertama, mengidentifikasi keterbatasan asesmen yang berorientasi pada produk akhir dalam konteks GenAI. Kedua, mengidentifikasi bentuk bukti alternatif yang dapat menunjukkan proses belajar dan kemampuan bernalar peserta didik. Ketiga, mengintegrasikan prinsip *authentic assessment* dengan kebutuhan transparansi dan penggunaan AI secara bertanggung jawab. Keempat, menyusun ketiga komponen tersebut menjadi model konseptual yang dapat diterapkan dalam perancangan asesmen PAI.

Model tersebut dirancang bukan untuk melarang penggunaan GenAI secara mutlak, melainkan untuk memastikan bahwa penggunaan teknologi tidak menghilangkan kompetensi yang hendak diukur. Dalam konteks tertentu, AI dapat diposisikan sebagai alat bantu yang penggunaannya perlu diungkapkan dan dievaluasi secara kritis. Sebaliknya, pada tugas yang memang dirancang untuk mengukur kemampuan individual tertentu, guru dapat menetapkan kondisi asesmen tanpa bantuan AI.

Dengan demikian, metode penelitian ini diarahkan untuk menghasilkan kerangka konseptual berbasis sintesis literatur, bukan klaim efektivitas empiris. Model Process-Reasoning-Authenticity (PRA) yang

dihasilkan selanjutnya menjadi dasar analisis dalam bagian hasil dan pembahasan untuk menjelaskan bagaimana asesmen PAI dapat direkonstruksi agar tetap valid, autentik, dan relevan di tengah perkembangan GenAI.

Hasil Penelitian

Berdasarkan hasil telaah dan analisis tematik terhadap literatur mengenai *generative artificial intelligence* (GenAI), asesmen pembelajaran, asesmen autentik, *higher-order thinking skills* (HOTS), integritas akademik, dan evaluasi Pendidikan Agama Islam (PAI), ditemukan bahwa kehadiran GenAI mendorong perubahan mendasar terhadap cara kompetensi peserta didik dibuktikan. Temuan literatur tidak menunjukkan bahwa produk pembelajaran menjadi tidak relevan, tetapi menunjukkan bahwa produk tidak lagi cukup apabila digunakan sebagai satu-satunya bukti untuk menilai kemampuan individual peserta didik. Lima tema utama ditemukan dalam sintesis literatur, yaitu melemahnya daya evidensial asesmen berbasis produk, meningkatnya pentingnya bukti proses, kebutuhan eksplisit terhadap *asesmen reasoning*, rekonstruksi konsep *authenticity*, serta perlunya transparansi penggunaan AI dalam asesmen.

Pergeseran Daya Evidensial Asesmen Berbasis Produk

Temuan pertama menunjukkan adanya perubahan pada posisi produk sebagai bukti kompetensi peserta didik. Dalam pola asesmen konvensional, guru umumnya menggunakan esai, makalah, presentasi, jawaban tertulis, atau produk proyek sebagai representasi capaian belajar. Produk tersebut diasumsikan merefleksikan pengetahuan, keterampilan berpikir, dan kemampuan peserta didik dalam menyelesaikan tugas. Namun, perkembangan GenAI memungkinkan peserta didik memperoleh bantuan dalam menghasilkan teks, menyusun gagasan, memperbaiki bahasa, membuat ringkasan, bahkan mengembangkan respons terhadap pertanyaan akademik.

Literatur mengenai ChatGPT dan asesmen menunjukkan bahwa kemampuan *generative AI* menghasilkan respons yang menyerupai tulisan manusia menciptakan persoalan baru bagi validitas asesmen berbasis produk. Cotton et al. (2024) menunjukkan bahwa penggunaan *ChatGPT* menghadirkan tantangan terhadap integritas akademik karena batas antara bantuan teknologi dan pekerjaan peserta didik menjadi semakin sulit ditentukan. Fischer et al. (2024) juga menunjukkan adanya perbedaan antara klaim peserta didik mengenai penggunaan *GenAI* dan penilaian *assessor* terhadap

keterlibatan AI dalam menghasilkan suatu tugas.

Dengan demikian, hasil sintesis menunjukkan bahwa persoalan utama bukan sekadar apakah suatu produk dibuat menggunakan AI atau tidak. Persoalan yang lebih mendasar adalah apakah produk tersebut memberikan bukti yang memadai mengenai kompetensi yang hendak diukur. Hao et al. (2024) menempatkan perkembangan large language models sebagai faktor yang mendorong transformasi asesmen karena teknologi tersebut memengaruhi hubungan antara tugas, respons peserta didik, dan inferensi mengenai kemampuan yang mendasarinya. Temuan ini dapat diringkas sebagai berikut:

Kondisi Asesmen	Asesmen Berbasis Produk	Asesmen Berbasis Bukti
Fokus utama	Hasil akhir	Proses dan hasil
Bukti kemampuan	Produk yang dikumpulkan	Proses, reasoning, refleksi
Posisi AI	Sering diperlakukan sebagai ancaman	Dikelola melalui transparansi penggunaan
Kemampuan yang terlihat	Jawaban atau performa akhir	Cara berpikir dan pengambilan keputusan

Validasi karya	Berdasarkan produk	Melalui triangulasi berbagai bukti
Risiko utama	Sulit memastikan authorship	Beban asesmen lebih tinggi bagi guru
Peran peserta didik	Penghasil produk	Agen pembelajaran yang menjelaskan proses

Hasil tersebut menunjukkan bahwa transformasi asesmen tidak harus dilakukan dengan menghapus tugas tertulis atau produk digital. Yang diperlukan adalah memperluas sumber bukti sehingga produk menjadi salah satu komponen dari keseluruhan *evidence of learning*.

Penguatan Bukti Proses dalam Asesmen PAI

Tema kedua menunjukkan bahwa bukti proses memiliki posisi strategis dalam mempertahankan validitas asesmen di era *GenAI*. Proses pembelajaran dapat didokumentasikan melalui berbagai bentuk, seperti rancangan awal, catatan perkembangan, jurnal reflektif, revisi tugas, dokumentasi diskusi, catatan umpan balik, rekaman presentasi, atau penjelasan lisan mengenai perkembangan suatu karya.

Dalam perspektif assessment literacy, guru tidak hanya dituntut mampu memberikan skor, tetapi juga memahami bukti apa yang diperlukan untuk membuat keputusan yang valid mengenai capaian peserta didik (Aria et al., 2021; Zulaiha et al., 2020). Oleh sebab itu, penggunaan bukti proses memungkinkan guru memperoleh informasi yang tidak tampak dari produk akhir.

Literatur mengenai asesmen autentik juga menunjukkan bahwa penilaian yang mengintegrasikan aktivitas dan proses dapat memberikan gambaran yang lebih luas mengenai kemampuan peserta didik. Muthohharoh et al. (2020), misalnya, menunjukkan relevansi asesmen autentik dalam menilai kemampuan berpikir tingkat tinggi. Ananda et al. (2020) juga menekankan pentingnya penguatan kompetensi guru dalam merancang asesmen autentik berbasis HOTS.

Dalam konteks PAI, bukti proses menjadi semakin penting karena tujuan pembelajaran tidak terbatas pada kemampuan menghasilkan jawaban benar. Pembelajaran PAI juga berkaitan dengan kemampuan memahami nilai, mempertimbangkan konsekuensi moral, mengambil keputusan berdasarkan prinsip keislaman, dan menghubungkan pengetahuan agama dengan persoalan kehidupan. Kompetensi tersebut lebih tepat diamati melalui rangkaian aktivitas daripada hanya melalui produk akhir.

Hasil sintesis menunjukkan bahwa bukti proses dalam asesmen PAI dapat mencakup:

- a. perkembangan gagasan dari tahap awal hingga produk akhir;
- b. argumentasi peserta didik dalam memilih suatu solusi;
- c. catatan perubahan atau revisi setelah memperoleh umpan balik;
- d. jurnal refleksi terhadap proses belajar;
- e. diskusi dan tanya jawab;
- f. presentasi atau pembelaan terhadap hasil kerja;
- g. dokumentasi penggunaan sumber;
- h. deklarasi penggunaan GenAI apabila teknologi tersebut digunakan.

Dengan demikian, proses tidak hanya berfungsi sebagai dokumentasi administratif, tetapi sebagai *evidence of learning* yang memungkinkan guru melakukan inferensi lebih kuat mengenai kemampuan peserta didik.

Reasoning sebagai Bukti Kompetensi yang Perlu Dinilai secara Eksplisit

Tema ketiga berkaitan dengan reasoning atau kemampuan peserta didik menjelaskan proses berpikir yang menghasilkan suatu jawaban atau keputusan. Hasil sintesis menunjukkan bahwa kemampuan menghasilkan jawaban yang benar tidak selalu identik dengan kemampuan memahami alasan di balik jawaban tersebut. Kondisi ini menjadi semakin penting ketika GenAI mampu menghasilkan jawaban yang

relatif koheren tanpa menunjukkan proses kognitif peserta didik.

Hao et al. (2024) menunjukkan bahwa perkembangan GenAI mengharuskan asesmen mempertimbangkan kembali bagaimana bukti kompetensi dikumpulkan. Dalam konteks tersebut, tugas yang hanya meminta peserta didik menghasilkan jawaban akhir memiliki keterbatasan karena tidak selalu memungkinkan guru mengamati bagaimana jawaban tersebut diperoleh.

Oleh karena itu, asesmen reasoning perlu mengubah pertanyaan dari sekadar “apa jawabanmu?” menjadi “mengapa kamu memilih jawaban tersebut?”, “bukti apa yang mendukung keputusanmu?”, dan “bagaimana kamu mempertimbangkan alternatif yang berbeda?”.

Dalam PAI, *reasoning* dapat diwujudkan melalui kemampuan peserta didik untuk:

Dimensi Reasoning	Bukti yang Diharapkan
Identifikasi masalah	Peserta didik mampu merumuskan persoalan secara tepat
Penggunaan sumber	Peserta didik mampu memilih dan menjelaskan relevansi sumber
Interpretasi	Peserta didik mampu menjelaskan makna informasi atau dalil

Argumentasi	Peserta didik mampu memberikan alasan yang logis dan relevan
Pertimbangan alternatif	Peserta didik mampu membandingkan beberapa kemungkinan
Pengambilan keputusan	Peserta didik mampu menjelaskan dasar keputusan
Refleksi	Peserta didik mampu mengevaluasi kembali proses berpikirnya

Temuan ini menunjukkan bahwa *reasoning* dapat menjadi mekanisme penting untuk memperkuat validitas asesmen. Ketika peserta didik diminta menjelaskan alasan, pertimbangan, sumber, dan proses pengambilan keputusan, guru memperoleh bukti yang lebih kaya daripada sekadar produk akhir.

Rekonstruksi Konsep Authenticity di Era GenAI

Tema keempat menunjukkan bahwa konsep authenticity perlu ditafsirkan kembali. Dalam konteks tradisional, authenticity sering dipahami sebagai keaslian karya peserta didik dan bahkan secara praktis direduksi menjadi pertanyaan apakah tugas tersebut dibuat sendiri tanpa bantuan orang lain atau teknologi. Kehadiran GenAI membuat pemahaman tersebut menjadi kurang memadai.

Hasil telaah menunjukkan bahwa authenticity lebih produktif apabila dipahami sebagai keaslian keterlibatan peserta didik dalam proses belajar dan kemampuan mempertanggungjawabkan hasil kerjanya. Dengan pengertian tersebut, penggunaan AI tidak otomatis membuat suatu karya kehilangan *authenticity*. Yang menjadi perhatian adalah apakah peserta didik memahami isi produk, mampu menjelaskan proses pembentukannya, dapat mengevaluasi keluaran AI, dan bertanggung jawab terhadap keputusan akademik yang diambil.

UNESCO melalui Miao dan Holmes (2023) menekankan pentingnya penggunaan GenAI secara berpusat pada manusia, bertanggung jawab, dan memperhatikan aspek etis dalam pendidikan. Perspektif tersebut memberikan dasar untuk menggeser pendekatan asesmen dari sekadar pengawasan terhadap penggunaan AI menuju pembentukan kemampuan menggunakan AI secara kritis dan bertanggung jawab.

Dengan demikian, hasil sintesis mengidentifikasi beberapa indikator authenticity yang relevan:

1. *Personal engagement*, yaitu adanya keterlibatan nyata peserta didik dalam proses penyelesaian tugas.
2. *Ownership of reasoning*, yaitu peserta didik mampu menjelaskan dan mempertanggungjawabkan alasan yang digunakan.

3. *Source awareness*, yaitu kemampuan mengidentifikasi dan mengevaluasi sumber informasi.
4. *AI transparency*, yaitu keterbukaan mengenai penggunaan GenAI dalam proses pengerjaan.
5. *Critical evaluation*, yaitu kemampuan mengkritisi keluaran AI dan tidak menerimanya secara otomatis.
6. *Contextual application*, yaitu kemampuan menghubungkan pengetahuan dengan konteks nyata.
7. *Reflection*, yaitu kemampuan mengevaluasi pengalaman belajar dan keputusan yang telah dibuat.

Dengan indikator tersebut, *authenticity* tidak lagi diposisikan sebagai kondisi biner “*menggunakan AI*” atau “*tidak menggunakan AI*”, tetapi sebagai kualitas keterlibatan, penguasaan, dan pertanggungjawaban peserta didik terhadap proses serta hasil belajarnya.

Transparansi Penggunaan GenAI sebagai Bagian dari Asesmen

Tema kelima menunjukkan bahwa transparansi penggunaan AI perlu menjadi bagian dari desain asesmen, terutama ketika penggunaan teknologi tersebut memang diperbolehkan. Literatur mengenai GenAI dalam pendidikan menunjukkan bahwa pendekatan yang hanya mengandalkan larangan menghadapi persoalan implementasi ketika teknologi AI telah menjadi bagian dari

lingkungan belajar peserta didik (Tlili et al., 2023; Cotton et al., 2024).

Oleh karena itu, hasil sintesis mengarah pada kebutuhan untuk membedakan antara *AI-assisted learning* dan *AI-substituted learning*. Pada kondisi pertama, AI digunakan sebagai alat bantu yang tetap berada di bawah kontrol intelektual peserta didik. Pada kondisi kedua, AI menggantikan sebagian besar proses berpikir yang seharusnya dilakukan peserta didik.

Perbedaan tersebut penting bagi asesmen PAI. Peserta didik dapat menggunakan GenAI, misalnya, untuk memperoleh alternatif penjelasan, menemukan pertanyaan pembanding, memperbaiki struktur tulisan, atau melakukan simulasi argumentasi. Namun, peserta didik tetap harus mampu melakukan verifikasi, memberikan alasan, mengoreksi kesalahan, dan menghubungkan hasil tersebut dengan prinsip atau materi PAI.

Dengan demikian, transparansi dapat dilakukan melalui pernyataan penggunaan AI yang sederhana, misalnya mencantumkan:

- a. apakah AI digunakan;
- b. aplikasi atau jenis AI yang digunakan;
- c. tujuan penggunaannya;
- d. bagian pekerjaan yang dibantu AI;
- e. bagaimana keluaran AI diverifikasi;
- f. perubahan apa yang dilakukan peserta didik terhadap keluaran tersebut.

Temuan ini menunjukkan bahwa transparansi bukan sekadar mekanisme pengawasan, tetapi juga dapat menjadi bagian dari literasi AI dan integritas akademik.

Formulasi Model Process-Reasoning-Authenticity (PRA)

Berdasarkan keseluruhan tema tersebut, penelitian ini menghasilkan formulasi konseptual *Process-Reasoning-Authenticity (PRA)* sebagai model asesmen PAI di era GenAI. Model ini tidak menghapus penilaian produk, tetapi menempatkan produk dalam sistem bukti yang lebih luas.

Model PRA terdiri atas tiga dimensi utama:

Dimensi	Fokus Penilaian	Contoh Bukti
Process	Bagaimana peserta didik mengembangkan tugas	draft, revisi, jurnal, diskusi, log aktivitas
	Mengapa peserta didik mengambil keputusan tertentu	argumentasi, penjelasan lisan, analisis, pertimbangan alternatif
Authenticity	Sejauh mana peserta didik memiliki keterlibatan dan tanggung jawab terhadap karya	refleksi, verifikasi sumber, deklarasi AI, pembelaan karya

Ketiga dimensi tersebut kemudian diintegrasikan dengan *Product* sebagai hasil akhir dan *Reflection/AI Transparency* sebagai bukti pertanggungjawaban.

Secara konseptual, struktur asesmen yang dihasilkan dapat dirumuskan sebagai:

Evidence of Learning = Process + Reasoning + Authenticity + Product + Reflection/AI Transparency

Formula tersebut bukan formula statistik, melainkan kerangka konseptual untuk membantu guru memahami bahwa keputusan asesmen sebaiknya tidak hanya didasarkan pada satu artefak pembelajaran.

Dalam model tersebut, produk tetap memiliki fungsi penting karena menunjukkan capaian akhir. Namun, kekuatan inferensi mengenai kompetensi peserta didik diperkuat melalui bukti proses, reasoning, authenticity, dan refleksi. Dengan demikian, penggunaan GenAI tidak secara otomatis merusak validitas asesmen; dampaknya sangat bergantung pada bagaimana tugas dan bukti pembelajaran dirancang.

Sintesis Hasil Penelitian

Secara keseluruhan, hasil penelitian menunjukkan bahwa GenAI tidak hanya menghadirkan persoalan baru bagi asesmen, tetapi juga mendorong redefinisi mengenai apa yang seharusnya dinilai. Lima temuan utama dapat dirumuskan sebagai

berikut: *Pertama*, asesmen yang hanya berorientasi pada produk semakin memiliki keterbatasan sebagai bukti tunggal kompetensi peserta didik karena produk dapat dihasilkan atau dimodifikasi dengan bantuan GenAI; *Kedua*, bukti proses menjadi semakin penting karena memungkinkan guru mengamati perkembangan gagasan, revisi, pengambilan keputusan, dan keterlibatan peserta didik; *Ketiga*, reasoning perlu dijadikan objek asesmen secara eksplisit agar guru dapat mengetahui alasan dan proses berpikir yang berada di balik suatu jawaban; *Keempat*, *authenticity* perlu direkonstruksi dari sekadar “*karya tanpa AI*” menjadi keterlibatan, penguasaan, transparansi, dan pertanggungjawaban peserta didik terhadap proses dan hasil belajar; *Kelima*, kombinasi *process*, *reasoning*, *authenticity*, *product*, dan *reflection/AI transparency* menghasilkan kerangka asesmen yang lebih kaya dibandingkan pendekatan yang hanya menilai produk akhir.

Temuan tersebut menjadi dasar bagi pembahasan mengenai bagaimana model PRA dapat diterapkan dalam pembelajaran PAI, termasuk bentuk tugas, instrumen, rubrik, dan strategi guru untuk menjaga validitas asesmen tanpa mengabaikan perkembangan teknologi generative AI.

Pembahasan

Transformasi Asesmen PAI di Tengah Perkembangan Generative AI

Perkembangan *generative artificial intelligence* (GenAI) membawa perubahan terhadap hubungan antara tugas, proses pengerjaan, produk, dan bukti kompetensi peserta didik. Dalam model asesmen konvensional, produk akhir sering kali diperlakukan sebagai representasi utama kemampuan peserta didik. Esai, makalah, laporan proyek, presentasi, dan berbagai bentuk tugas tertulis digunakan untuk menarik kesimpulan mengenai penguasaan pengetahuan dan keterampilan. Namun, ketika teknologi generatif mampu menghasilkan teks, menyusun argumentasi, memberikan alternatif jawaban, serta membantu melakukan penyuntingan, hubungan antara produk dan kompetensi individual menjadi semakin kompleks (Hao et al., 2024; Lin & Chiu, 2024).

Persoalan tersebut tidak berarti bahwa produk pembelajaran kehilangan nilai asesmennya. Produk tetap merupakan bukti penting karena memperlihatkan kemampuan peserta didik menghasilkan suatu keluaran. Persoalannya terletak pada kecukupan produk sebagai bukti tunggal. Ketika sebuah esai dapat dihasilkan dengan bantuan GenAI, guru tidak dapat secara langsung menyimpulkan bahwa kualitas produk identik dengan kemampuan berpikir peserta didik. Dengan demikian, tantangan utama asesmen di era GenAI sesungguhnya

adalah tantangan validitas inferensi, yaitu sejauh mana bukti yang dikumpulkan benar-benar dapat digunakan untuk menyimpulkan kompetensi yang hendak diukur (Hao et al., 2024).

Lin dan Chiu (2024) menunjukkan bahwa GenAI mendorong perubahan dari asesmen tradisional menuju pendekatan yang lebih menekankan kemampuan tingkat tinggi, proses berpikir, dan bentuk asesmen yang lebih adaptif. Temuan tersebut relevan dengan pembelajaran PAI karena kompetensi PAI tidak berhenti pada kemampuan mengingat konsep atau dalil, tetapi juga mencakup kemampuan memahami makna, menganalisis persoalan, mengambil keputusan berdasarkan nilai, dan menerapkan prinsip keislaman dalam kehidupan.

Dengan demikian, transformasi asesmen PAI seharusnya tidak dipahami sebagai upaya “melawan” teknologi. Guru justru perlu mendesain asesmen yang menjadikan kemampuan manusia yang sulit direduksi menjadi sekadar keluaran teks sebagai bagian penting dari bukti pembelajaran. Kemampuan tersebut antara lain penalaran, refleksi, pertimbangan moral, komunikasi, pengambilan keputusan, dan kemampuan menghubungkan pengetahuan dengan konteks kehidupan.

Dari Penilaian Produk menuju Penilaian Berbasis Bukti

Hasil penelitian menunjukkan bahwa pergeseran yang diperlukan bukanlah perpindahan sederhana dari *product assessment* menuju *process assessment*. Jika proses hanya dimaknai sebagai dokumentasi tahapan pengerjaan, transformasi tersebut belum cukup menjawab tantangan GenAI. Yang diperlukan adalah perubahan paradigma dari menilai produk menjadi menilai kumpulan bukti pembelajaran (*evidence of learning*).

Dalam kerangka tersebut, produk tetap dinilai, tetapi dikombinasikan dengan bukti lain yang dapat memperkuat interpretasi guru. Bukti tersebut dapat berupa rancangan awal, perkembangan gagasan, revisi, catatan refleksi, diskusi, presentasi, argumentasi lisan, serta penjelasan mengenai penggunaan sumber dan teknologi. Pendekatan semacam ini sejalan dengan prinsip asesmen autentik yang menempatkan kemampuan peserta didik dalam konteks tugas yang bermakna dan memungkinkan berbagai dimensi kompetensi diamati (Ananda et al., 2020; Muthohharoh et al., 2020). Perubahan tersebut dapat digambarkan sebagai berikut:

Model lama:

Tugas → Produk → Skor

Model berbasis bukti:

**Tugas → Proses → Reasoning →
Produk → Refleksi → Keputusan
Asesmen**

Model kedua menghasilkan informasi yang lebih beragam bagi guru. Produk menjawab pertanyaan tentang “*apa yang dihasilkan*”, sedangkan proses menjawab “*bagaimana karya dikembangkan*”. Reasoning menjawab “*mengapa keputusan tertentu diambil*”, sementara refleksi dan transparansi AI menjawab “*bagaimana peserta didik mempertanggungjawabkan proses dan hasil tersebut*”.

Perubahan ini juga menuntut peningkatan *assessment literacy* guru. Guru perlu menentukan sejak awal kompetensi apa yang akan diukur, bukti apa yang diperlukan, bagaimana bukti tersebut dikumpulkan, dan bagaimana bukti yang berbeda digunakan untuk menghasilkan keputusan asesmen. Aria et al. (2021) menegaskan bahwa kompetensi asesmen guru berkaitan dengan kemampuan merancang dan menggunakan asesmen dalam praktik pembelajaran, bukan hanya kemampuan memberikan nilai. Zulaiha et al. (2020) juga menunjukkan pentingnya hubungan antara pemahaman guru mengenai asesmen dan praktik asesmen di kelas.

Oleh karena itu, era GenAI sekaligus memperlihatkan bahwa persoalan asesmen bukan hanya persoalan teknologi, tetapi juga persoalan kapasitas pedagogis guru.

Reasoning sebagai Pusat Baru Validitas Asesmen

Salah satu implikasi paling penting dari kehadiran GenAI adalah meningkatnya kebutuhan untuk menilai reasoning secara eksplisit. Jika teknologi mampu menghasilkan jawaban yang tampak benar, maka pertanyaan asesmen perlu bergeser dari sekadar meminta jawaban menuju permintaan untuk menjelaskan proses berpikir yang menghasilkan jawaban tersebut.

Reasoning dalam konteks ini tidak identik dengan *chain-of-thought* internal suatu sistem AI. Yang dimaksud adalah alasan yang dapat dijelaskan dan dipertanggungjawabkan oleh peserta didik, seperti dasar pemilihan sumber, interpretasi informasi, pertimbangan terhadap alternatif, alasan mengambil suatu keputusan, dan refleksi terhadap hasil kerja.

Hao et al. (2024) menunjukkan bahwa perkembangan *large language models* mengharuskan pendidik mempertimbangkan kembali bentuk tugas dan bukti yang digunakan dalam asesmen. Apabila tugas hanya meminta jawaban akhir, maka sebagian proses kognitif peserta didik menjadi tidak terlihat. Sebaliknya, tugas yang meminta argumentasi dan pembelaan terhadap jawaban memungkinkan guru memperoleh bukti tambahan mengenai penguasaan peserta didik.

Dalam PAI, *reasoning* memiliki posisi yang lebih kompleks karena

objek pembelajaran sering berkaitan dengan persoalan nilai dan keputusan moral. Misalnya, ketika peserta didik diminta menganalisis persoalan etika bermedia sosial, jawaban yang baik tidak cukup hanya menyebutkan bahwa suatu tindakan “boleh” atau “tidak boleh”. Peserta didik perlu menjelaskan prinsip yang digunakan, sumber yang relevan, konteks persoalan, alternatif tindakan, serta konsekuensi dari keputusan yang diambil.

Dengan demikian, bentuk pertanyaan asesmen dapat direkonstruksi dari: “*Jelaskan pengertian toleransi dalam Islam.*” menjadi tugas yang lebih menuntut *reasoning*: “*Analisis satu kasus interaksi digital yang berpotensi menimbulkan konflik. Jelaskan prinsip Islam yang relevan, pertimbangkan sedikitnya dua alternatif respons, tentukan respons yang Anda pilih, dan jelaskan alasan serta sumber yang mendasari keputusan tersebut.*”

Pertanyaan kedua menghasilkan bukti yang lebih kaya karena peserta didik tidak hanya dituntut menghasilkan jawaban, tetapi juga mempertanggungjawabkan proses pengambilan keputusan. Pendekatan ini sejalan dengan orientasi asesmen HOTS yang menekankan analisis, evaluasi, dan penciptaan (Ananda et al., 2020).

Rekonstruksi Authenticity: Bukan Sekadar “Tanpa AI”

Konsep *authenticity* menjadi salah satu aspek yang paling perlu direkonstruksi dalam asesmen PAI di era GenAI. Pemaknaan *authenticity* secara sempit sebagai karya yang sepenuhnya dibuat tanpa bantuan teknologi berpotensi menghasilkan pendekatan yang sulit diterapkan dalam lingkungan digital. GenAI telah menjadi bagian dari ekosistem informasi yang diakses peserta didik, sehingga persoalan pedagogis yang lebih produktif adalah bagaimana memastikan bahwa peserta didik tetap menjadi agen utama dalam proses belajar.

UNESCO menekankan bahwa penggunaan GenAI dalam pendidikan harus ditempatkan dalam kerangka yang berpusat pada manusia, memperhatikan etika, keamanan, kesetaraan, serta tanggung jawab pengguna (Miao & Holmes, 2023). Perspektif tersebut mengarah pada pemahaman bahwa literasi AI tidak hanya berarti mampu menggunakan AI, tetapi juga mampu memahami keterbatasannya dan mengevaluasi hasil yang diberikan.

Oleh sebab itu, *authenticity* dalam artikel ini dipahami sebagai kualitas keterlibatan dan pertanggungjawaban peserta didik terhadap proses serta hasil pembelajaran. Peserta didik dapat menggunakan AI secara terbatas apabila kebijakan tugas

mengizinkannya, tetapi penggunaan tersebut harus transparan dan tidak menggantikan kompetensi yang sedang dinilai.

Misalnya, dalam tugas penulisan artikel PAI, penggunaan GenAI untuk mencari alternatif struktur tulisan dapat diperbolehkan. Namun, peserta didik harus tetap mampu memeriksa kebenaran informasi, mengevaluasi sumber, mengoreksi kekeliruan, dan menjelaskan alasan memilih argumen tertentu. Dalam kondisi tersebut, aspek *authenticity* tidak lagi ditentukan hanya oleh ada atau tidak adanya AI, tetapi oleh *agency*, *critical evaluation*, *ownership*, dan *accountability* peserta didik.

Pendekatan tersebut juga dapat mengurangi dikotomi “*AI versus non-AI*”. Yang lebih relevan adalah membedakan antara penggunaan AI yang mendukung pembelajaran dan penggunaan AI yang menggantikan aktivitas kognitif yang seharusnya dilakukan peserta didik. Tlili et al. (2023) menunjukkan bahwa *ChatGPT* menghadirkan peluang sekaligus tantangan dalam pendidikan, sehingga respons pedagogis perlu mempertimbangkan bagaimana teknologi tersebut digunakan, bukan hanya keberadaannya.

Model Process–Reasoning–Authenticity sebagai Rekonstruksi Asesmen PAI

Berdasarkan sintesis teoritis, penelitian ini mengusulkan model *Process–Reasoning–Authenticity (PRA)*

sebagai kerangka untuk merancang asesmen PAI di era *GenAI*. Model tersebut dibangun atas asumsi bahwa validitas keputusan asesmen akan lebih kuat apabila guru menggunakan berbagai bukti yang saling melengkapi.

Process mengacu pada bukti perkembangan pekerjaan peserta didik. *Reasoning* mengacu pada kemampuan menjelaskan dasar pemikiran dan pengambilan keputusan. *Authenticity* mengacu pada keterlibatan, kepemilikan, transparansi, dan pertanggungjawaban peserta didik terhadap karya. Ketiganya tetap dikaitkan dengan *Product* dan *Reflection/AI Transparency*.

Dimensi	Pertanyaan Utama	Bukti Asesmen	Instrumen yang Dapat Digunakan
Process	Bagaimana karya dikembangkan?	Draft, revisi, catatan proses	Portfolio/process log
Reasoning	Mengapa keputusan tersebut diambil?	Argumentasi, analisis reasoning, , , oral defense	Rubrik
Authenticity	Apakah peserta didik memahami dan memiliki karya?	Refleksi, verifikasi, sumber	Reflektif journal

Product	Apa hasil yang dihasilkan?	Esai, proyek, Rubrik presen produk tasi
Reflection/AI Transparency	Bagaimana proses dan AI penggunaan AI dipertanggungjawabkan?	Form deklarasi deklarasi tion, dan refleksi refleksi

Model tersebut menunjukkan bahwa asesmen tidak perlu meninggalkan produk. Sebaliknya, produk ditempatkan sebagai salah satu komponen dari ekosistem bukti. Konsekuensinya, tugas yang baik di era *GenAI* bukan sekadar tugas yang sulit dikerjakan oleh AI, melainkan tugas yang memerlukan keterlibatan intelektual peserta didik dan menghasilkan bukti keterlibatan tersebut.

Implikasi PRA terhadap Desain Tugas PAI

Model PRA dapat diterapkan pada berbagai materi PAI. Perubahan paling penting terletak pada desain tugas. Tugas yang hanya meminta reproduksi informasi relatif mudah dialihkan kepada *GenAI*. Sebaliknya, tugas berbasis kasus, konteks lokal, pengalaman reflektif, argumentasi, dan pembelaan lisan memungkinkan guru memperoleh bukti yang lebih autentik mengenai kompetensi peserta didik.

Contohnya, pada materi akhlak bermedia sosial, tugas dapat dirancang

dengan meminta peserta didik mengidentifikasi permasalahan nyata atau simulatif, menganalisisnya berdasarkan prinsip Islam, menyusun beberapa alternatif penyelesaian, memilih salah satu solusi, menghasilkan produk kampanye digital, dan kemudian mempertahankan keputusan tersebut melalui presentasi singkat.

Struktur tugas tersebut memungkinkan guru memperoleh beberapa jenis bukti sekaligus:

Kasus → Analisis → Reasoning → Produk → Presentasi → Refleksi

Pada tahap analisis, guru menilai kemampuan mengidentifikasi masalah. Pada tahap reasoning, guru menilai argumentasi. Pada tahap produk, guru menilai kualitas hasil. Pada tahap presentasi, guru dapat memverifikasi penguasaan peserta didik terhadap karya. Sementara itu, refleksi dapat digunakan untuk mengetahui apa yang dipelajari, kesulitan yang dihadapi, serta bagaimana teknologi digunakan selama proses pengerjaan.

Dengan demikian, desain asesmen dapat sekaligus berfungsi sebagai strategi pembelajaran. Peserta didik tidak hanya “dinilai”, tetapi diarahkan untuk menyadari bagaimana mereka belajar, bagaimana mengambil keputusan, dan bagaimana bertanggung jawab terhadap penggunaan teknologi.

Transparansi AI dan Integritas Akademik

Penggunaan GenAI juga membawa persoalan integritas akademik. Namun, pendekatan yang hanya mengandalkan deteksi AI memiliki keterbatasan. Detektor AI tidak seharusnya menjadi satu-satunya dasar untuk menentukan apakah suatu karya autentik. Keputusan asesmen perlu didasarkan pada keseluruhan bukti pembelajaran.

Cotton et al. (2024) menunjukkan bahwa perkembangan ChatGPT menghadirkan tantangan bagi integritas akademik dan mendorong institusi pendidikan untuk memikirkan kembali kebijakan serta strategi asesmennya. Dalam konteks ini, transparansi penggunaan AI dapat menjadi salah satu mekanisme pedagogis untuk membangun integritas.

Guru dapat meminta peserta didik membuat pernyataan singkat seperti: *“Dalam pengerjaan tugas ini, saya menggunakan GenAI untuk membantu menghasilkan alternatif pertanyaan dan memperbaiki struktur tulisan. Informasi yang dihasilkan kemudian saya periksa menggunakan sumber akademik dan saya revisi berdasarkan materi pembelajaran.”*

Pernyataan tersebut tidak dimaksudkan sebagai pengganti penilaian, tetapi sebagai salah satu bukti proses. Lebih penting lagi, peserta didik belajar bahwa penggunaan

teknologi akademik harus disertai tanggung jawab.

Fischer et al. (2024) memperlihatkan pentingnya melihat penggunaan GenAI secara lebih kompleks daripada sekadar membedakan karya “AI” dan “non-AI”. Perbedaan antara penggunaan teknologi, persepsi peserta didik, dan penilaian assessor menunjukkan bahwa asesmen perlu menyediakan mekanisme verifikasi yang lebih kaya.

Dalam PAI, prinsip transparansi tersebut memiliki relevansi etis yang kuat. Pembelajaran agama tidak hanya bertujuan menghasilkan pengetahuan, tetapi juga membentuk tanggung jawab moral. Karena itu, kejujuran dalam melaporkan penggunaan sumber dan teknologi dapat diposisikan sebagai bagian dari proses pendidikan karakter, bukan sekadar aturan administratif.

Peran Guru dalam Implementasi Asesmen PRA

Transformasi asesmen tidak akan berjalan tanpa perubahan peran guru. Guru tidak lagi cukup berperan sebagai pemeriksa produk akhir, tetapi perlu berfungsi sebagai perancang bukti pembelajaran, fasilitator reasoning, dan evaluator proses belajar.

Perubahan tersebut menuntut kemampuan guru dalam tiga aspek. Pertama, kemampuan merancang tugas yang menghasilkan bukti beragam. Kedua, kemampuan mengembangkan

rubrik yang membedakan kualitas produk, proses, reasoning, dan refleksi. Ketiga, kemampuan melakukan dialog atau verifikasi ketika terdapat keraguan terhadap penguasaan peserta didik.

Assessment literacy menjadi faktor penting dalam proses tersebut. Aria et al. (2021), Latif (2021), dan Zulaiha et al. (2020) menunjukkan bahwa kemampuan guru dalam memahami prinsip, tujuan, dan praktik asesmen berhubungan dengan kualitas pelaksanaan asesmen. Dalam konteks GenAI, assessment literacy perlu diperluas menjadi kemampuan memahami hubungan antara teknologi, tugas, bukti, dan validitas keputusan asesmen.

Guru juga perlu menghindari beban dokumentasi yang berlebihan. Tidak semua tugas harus menggunakan seluruh komponen PRA. Untuk tugas sederhana, guru dapat menggunakan produk dan reasoning. Untuk proyek besar, guru dapat menambahkan process portfolio, presentasi, dan refleksi. Dengan demikian, model PRA bersifat fleksibel dan dapat disesuaikan dengan tujuan pembelajaran.

Relevansi Model PRA bagi Pembelajaran PAI di Indonesia

Konteks pendidikan Indonesia memberikan alasan tambahan untuk memperkuat asesmen yang berorientasi pada proses, kompetensi, dan pengalaman belajar. PAI tidak ideal apabila direduksi menjadi pengukuran

hafalan semata. Pembelajaran agama perlu memberi ruang bagi peserta didik untuk memahami nilai, menginterpretasikan persoalan, berdialog dengan konteks kehidupan, dan mengembangkan sikap yang bertanggung jawab.

Penelitian mengenai asesmen autentik menunjukkan bahwa pendekatan yang menilai kemampuan melalui tugas bermakna dapat mendukung pengukuran kemampuan berpikir tingkat tinggi (Muthohharoh et al., 2020). Sementara itu, penelitian mengenai pemanfaatan AI dalam pembelajaran dan evaluasi PAI menunjukkan bahwa teknologi memberikan peluang efisiensi sekaligus menghadirkan tantangan baru terkait validitas dan kesiapan pendidik (Sofyan & Salito, 2024; Rahmawati et al., 2025).

Dengan demikian, penggunaan GenAI dalam PAI sebaiknya tidak diarahkan hanya pada otomatisasi pembuatan soal, koreksi, atau pemberian umpan balik. Fokus yang lebih strategis adalah bagaimana teknologi tersebut mendorong **rekonstruksi desain asesmen**. Guru perlu bertanya bukan hanya *"bagaimana AI dapat membantu saya menilai?"*, tetapi juga *"bukti apa yang sekarang harus saya minta dari peserta didik agar penilaian tetap valid?"*

Pertanyaan kedua merupakan titik penting dari transformasi asesmen. Teknologi berubah, sehingga desain bukti pembelajaran juga perlu

berubah. Akan tetapi, tujuan fundamental asesmen tetap sama, yaitu memperoleh informasi yang cukup dan relevan untuk membuat keputusan mengenai perkembangan belajar peserta didik.

Keunggulan dan Keterbatasan Model PRA

Model PRA menawarkan beberapa kontribusi konseptual. Pertama, model ini mengintegrasikan dimensi proses, reasoning, dan authenticity dalam satu kerangka asesmen. Kedua, model ini tidak memandang AI semata-mata sebagai ancaman terhadap integritas akademik, tetapi sebagai bagian dari lingkungan belajar yang harus dikelola secara transparan. Ketiga, model ini mempertahankan produk sebagai bukti penting sekaligus mengurangi ketergantungan terhadap produk sebagai satu-satunya indikator kompetensi.

Namun, penerapan model PRA juga memiliki keterbatasan. Penilaian proses dan reasoning membutuhkan waktu yang lebih besar dibandingkan pemeriksaan produk tunggal. Guru juga memerlukan kemampuan dalam merancang rubrik dan mengelola berbagai jenis bukti. Selain itu, konsep authenticity tetap memerlukan indikator operasional yang lebih teruji melalui penelitian empiris.

Karena model PRA dalam penelitian ini merupakan hasil sintesis konseptual dari literatur, bobot

penilaian dan instrumen yang digunakan tidak dapat dianggap sebagai standar universal. Implementasi di lapangan perlu mempertimbangkan jenjang pendidikan, karakteristik materi PAI, jumlah peserta didik, kompetensi guru, akses teknologi, serta kebijakan institusi.

Dengan demikian, kontribusi utama model PRA bukan menyediakan satu formula penilaian yang kaku, melainkan menawarkan kerangka berpikir untuk mendesain asesmen yang lebih kaya bukti. Validasi empiris melalui penelitian tindakan kelas, eksperimen, studi kasus, atau penelitian pengembangan masih diperlukan untuk menguji efektivitas model tersebut.

Sintesis Pembahasan

Pembahasan menunjukkan bahwa tantangan utama GenAI terhadap asesmen PAI bukan sekadar persoalan kecurangan akademik, melainkan persoalan yang lebih fundamental mengenai apa yang dianggap sebagai bukti kompetensi peserta didik. Jika produk akhir digunakan sebagai bukti tunggal, kemampuan *GenAI* menghasilkan produk akademik dapat memperlemah validitas inferensi. Sebaliknya, jika asesmen mengintegrasikan proses, reasoning, authenticity, produk, dan refleksi, guru memperoleh gambaran kompetensi yang lebih komprehensif.

Atas dasar itu, pergeseran asesmen PAI di era GenAI dapat dirumuskan sebagai perubahan: dari *"menilai apa yang dihasilkan peserta didik"* menuju *"menilai bagaimana peserta didik berpikir, bekerja, menggunakan teknologi, menghasilkan karya, dan mempertanggungjawabkan hasilnya."*

Pergeseran tersebut menempatkan peserta didik sebagai agen pembelajaran dan guru sebagai perancang sistem bukti pembelajaran. Dalam kerangka ini, GenAI tidak harus diposisikan sebagai musuh asesmen, tetapi sebagai faktor yang memaksa pendidikan memperjelas kompetensi manusia yang memang hendak dinilai.

Model *Process-Reasoning-Authenticity (PRA)* dengan demikian dapat dipandang sebagai salah satu kerangka konseptual untuk melakukan rekonstruksi asesmen PAI. Model tersebut tetap mempertahankan produk, tetapi menambahkan bukti proses, reasoning, authenticity, dan transparansi penggunaan AI agar keputusan asesmen tidak bergantung pada satu artefak pembelajaran.

Kesimpulan

Perkembangan *generative artificial intelligence* (GenAI) telah mengubah kondisi epistemik asesmen pembelajaran, khususnya karena produk akademik yang sebelumnya dipandang sebagai representasi kemampuan peserta didik kini dapat dihasilkan atau dimodifikasi dengan

bantuan teknologi generatif. Berdasarkan hasil kajian literatur, persoalan utama yang muncul bukan terletak pada hilangnya relevansi produk pembelajaran, melainkan pada keterbatasan produk apabila digunakan sebagai satu-satunya bukti untuk menyimpulkan kompetensi peserta didik. Oleh karena itu, asesmen PAI memerlukan pergeseran dari orientasi yang dominan pada hasil akhir menuju sistem asesmen yang menggunakan bukti pembelajaran secara lebih komprehensif.

Kajian ini menunjukkan bahwa proses, *reasoning*, dan *authenticity* merupakan tiga dimensi penting yang perlu diperkuat dalam asesmen PAI di era GenAI. Bukti proses memungkinkan guru melihat perkembangan gagasan, revisi, pengambilan keputusan, dan keterlibatan peserta didik. Penilaian *reasoning* memungkinkan guru mengidentifikasi kemampuan peserta didik dalam memberikan alasan, mengevaluasi informasi, mempertimbangkan alternatif, dan mempertanggungjawabkan keputusan. Sementara itu, *authenticity* perlu dipahami kembali bukan semata-mata sebagai karya yang bebas dari bantuan teknologi, tetapi sebagai keterlibatan, kepemilikan, transparansi, evaluasi kritis, dan tanggung jawab peserta didik terhadap proses serta hasil pembelajaran.

Atas dasar temuan tersebut, penelitian ini merumuskan model *Process-Reasoning-Authenticity* (PRA)

sebagai kerangka konseptual asesmen PAI di era GenAI. Model PRA menempatkan produk tetap sebagai salah satu bukti pembelajaran, tetapi mengintegrasikannya dengan bukti proses, *reasoning*, *authenticity*, serta refleksi dan transparansi penggunaan AI. Kerangka tersebut memberikan alternatif terhadap pendekatan asesmen yang hanya berfokus pada deteksi penggunaan AI atau pemeriksaan produk akhir. Dalam model PRA, perhatian utama dialihkan pada kualitas bukti yang memungkinkan guru memastikan bahwa peserta didik memahami, mengembangkan, dan mampu mempertanggungjawabkan hasil belajarnya.

Kontribusi utama kajian ini terletak pada penegasan bahwa respons terhadap GenAI dalam asesmen PAI tidak cukup dilakukan melalui pelarangan atau peningkatan pengawasan, tetapi memerlukan rekonstruksi desain asesmen. Tugas pembelajaran perlu dirancang sedemikian rupa sehingga mendorong analisis, argumentasi, refleksi, pengambilan keputusan, penerapan nilai Islam dalam konteks nyata, serta kemampuan menjelaskan proses berpikir. Dengan demikian, GenAI dapat ditempatkan sebagai bagian dari lingkungan belajar yang penggunaannya dikelola secara etis dan transparan, sementara kompetensi intelektual dan moral peserta didik tetap menjadi pusat asesmen.

Saran

Implementasi model PRA memerlukan penguatan kompetensi asesmen guru. Guru PAI perlu meningkatkan kemampuan dalam merancang tugas berbasis proses, mengembangkan rubrik *reasoning*, menilai refleksi, serta menetapkan aturan penggunaan GenAI yang jelas sesuai tujuan pembelajaran. Kebijakan asesmen juga perlu membedakan penggunaan AI sebagai alat bantu belajar dengan penggunaan AI yang menggantikan kompetensi yang sedang dinilai.

Bagi satuan pendidikan dan pengelola program studi, diperlukan pedoman yang lebih adaptif mengenai penggunaan GenAI dalam asesmen. Pedoman tersebut tidak hanya mengatur larangan dan sanksi, tetapi juga menjelaskan kapan AI diperbolehkan, bagaimana penggunaannya harus dilaporkan, jenis bukti yang perlu dikumpulkan, dan bagaimana guru melakukan verifikasi terhadap penguasaan peserta didik.

Bagi peneliti berikutnya, model PRA perlu diuji secara empiris pada berbagai jenjang dan konteks pembelajaran PAI. Penelitian selanjutnya dapat mengembangkan instrumen dan rubrik PRA, menguji validitas dan reliabilitasnya, serta membandingkan efektivitas asesmen berbasis produk dengan asesmen berbasis bukti yang mengintegrasikan proses, *reasoning*, dan *authenticity*. Penelitian empiris juga diperlukan

untuk mengetahui bagaimana model tersebut memengaruhi kemampuan berpikir kritis, integritas akademik, literasi AI, dan keterlibatan belajar peserta didik.

Dengan demikian, rekonstruksi asesmen PAI di era GenAI pada akhirnya bukan sekadar perubahan teknik pemberian nilai, tetapi perubahan cara pendidikan memahami bukti belajar. Asesmen yang relevan di era kecerdasan generatif adalah asesmen yang mampu melihat bukan hanya apa yang dihasilkan peserta didik, tetapi juga bagaimana ia berpikir, bagaimana ia belajar, bagaimana ia menggunakan teknologi, dan sejauh mana ia mampu mempertanggungjawabkan pengetahuan serta keputusan yang dihasilkannya.

Daftar Pustaka

- Ananda, L. J., Fahrurrozi, & Febrian, D. (2020). Penguatan kompetensi guru dalam penilaian autentik berbasis higher order thinking skills. *School Education Journal PGSD FIP Unimed*, 9(3), 263–270.
<https://doi.org/10.24114/sejpgsd.v9i3.16155>
- Aria, D., Sukyadi, D., & Kurniawan, E. (2021). Teacher assessment literacy: Indonesian EFL secondary teachers' self-perceived on classroom-based assessment practice. *English Review: Journal of English*

- Education*, 10(1), 15–26.
<https://doi.org/10.25134/erjee.v10i1.5349>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.
<https://doi.org/10.1007/s11092-008-9068-5>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239.
<https://doi.org/10.1080/14703297.2023.2190148>
- Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? *Journal of Responsible Technology*, 13, 100060.
<https://doi.org/10.1016/j.jrt.2023.100060>
- Faizin, N., Alfian, M., Basid, A., Ramadhan, M. R., Panatik, S. A., & Kawakip, A. N. (2025). Muslim students' acceptance of artificial intelligence in Islamic religious education: An extended TAM approach. *Discover Education*, 4, 304.
<https://doi.org/10.1007/s44217-025-00767-1>
- Fischer, I., Sweeney, S., Lucas, M., & Gupta, N. (2024). Making sense of generative AI for assessments: Contrasting student claims and assessor evaluations. *The International Journal of Management Education*, 22(3), 101081.
<https://doi.org/10.1016/j.ijme.2024.101081>
- Hao, L., von Davier, M., & Harris, D. J. (2024). Transforming assessment: The impacts and implications of large language models and generative AI. *Educational Measurement: Issues and Practice*, 43(2), 5–15.
<https://doi.org/10.1111/emip.12602>
- Latif, M. W. (2021). Exploring tertiary EFL practitioners' knowledge base component of assessment literacy: Implications for teacher professional development. *Language Testing in Asia*, 11, 19.
<https://doi.org/10.1186/s40468-021-00130-9>
- Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21, 40.
<https://doi.org/10.1186/s41239-024-00468-z>
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO.
- Muthohharoh, S. R., Bharati, D. A. L., & Rozi, F. (2020). The

- implementation of authentic assessment to assess students' higher order thinking skills in writing at MAN 2 Tulungagung. *English Education Journal*, 10(3), 374–386.
<https://doi.org/10.15294/eej.v10i1.36590>
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218.
<https://doi.org/10.1080/03075070600572090>
- Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024). Higher education assessment practice in the era of generative AI tools. *Journal of Applied Learning and Teaching*, 7(1), 46–56.
<https://doi.org/10.37074/jalt.2024.7.1.28>
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129–144.
<https://doi.org/10.1016/j.edur.2013.01.002>
- Panadero, E., Jonsson, A., & Strijbos, J.-W. (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. In D. Laveault & L. Allal (Eds.), *Assessment for learning: Meeting the challenge of implementation* (pp. 311–326). Springer.
https://doi.org/10.1007/978-3-319-39211-0_18
- Rahmawati, A. R., Wulandari, R., Sadaruddin, M. A. S., Mustofa, T. A., & Chedimae, H. (2025). Optimization of artificial intelligence in Islamic Religious Education: Opportunities and challenges in learning evaluation. *Suhuf: International Journal of Islamic Studies*, 37(2), 274–287.
<https://doi.org/10.23917/suhuf.v37i2.11015>
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144.
<https://doi.org/10.1007/BF00117714>
- Sari, N. I. P., Thohir, M., & Al Fath, M. G. F. (2025). Between humans and religious values: Collaboration in AI-based Islamic Religious Education learning. *el-Tarbawi*, 18(2).
<https://doi.org/10.20885/tarbiawi.vol18.iss2.art5>
- Sofyan, A., & Salito. (2024). Pengembangan penilaian pembelajaran PAI berbasis kecerdasan buatan: Peluang dan tantangan di MTs Durul Jazil. *Al-Qalam: Jurnal Kajian Islam dan*

- Pendidikan*, 16(2).
<https://doi.org/10.47435/al-qalam.v16i2.3290>
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, 6(1), 31–40.
<https://doi.org/10.37074/jalt.2023.6.1.17>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10, 15.
- Weng, X., Xia, Q., Gu, M., Rajaram, K., & Chiu, T. K. F. (2024). Assessment and learning outcomes for generative AI in higher education: A scoping review on current research status and trends. *Australasian Journal of Educational Technology*, 40(6), 37–55.
<https://doi.org/10.14742/ajet.9540>
- Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. Jossey-Bass.
- Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21, 40.
<https://doi.org/10.1186/s41239-024-00468-z>
- Zulaiha, S., Mulyono, H., & Ambarsari, L. (2020). An investigation into EFL teachers' assessment literacy: Indonesian teachers' perceptions and classroom practice. *European Journal of Contemporary Education*, 9(1), 189–201.
<https://doi.org/10.13187/ejced.2020.1.189>
- Abdilah, Y. A. (2023). Tinjauan sistematis etika penggunaan ChatGPT di perguruan tinggi. *Integralistik*, 34(2).
<https://doi.org/10.15294/integralistik.v34i2.50278>
- Bin-Nashwan, S. A., Sadallah, M., & Bouteraa, M. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance. *Technology in Society*, 75, 102370.
<https://doi.org/10.1016/j.techsoc.2023.102370>
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, 6(1), 31–40.
<https://doi.org/10.37074/jalt.2023.6.1.17>

- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239.
<https://doi.org/10.1080/14703297.2023.2190148>
- Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024). Higher education assessment practice in the era of generative AI tools. *Journal of Applied Learning and Teaching*, 7(1), 46–56.
<https://doi.org/10.37074/jalt.2024.7.1.28>
- Weng, X., Xia, Q., Gu, M., Rajaram, K., & Chiu, T. K. F. (2024). Assessment and learning outcomes for generative AI in higher education: A scoping review on current research status and trends. *Australasian Journal of Educational Technology*, 40(6), 37–55.
<https://doi.org/10.14742/ajet.9540>
- Rahmawati, A. R., Wulandari, R., Sadaruddin, M. A. S., Mustofa, T. A., & Chedimae, H. (2025). Optimization of artificial intelligence in Islamic Religious Education: Opportunities and challenges in learning evaluation. *Suhuf: International Journal of Islamic Studies*, 37(2), 274–287.
<https://doi.org/10.23917/suhuf.v37i2.11015>