

Adoption of AI-assisted coding tools in programming courses: evidence and guidelines for e-learning in engineering

Dayron Rumbaut-Rangel¹; Franklin Parrales-Bravo^{1,2}; Lorenzo Cevallos-Torres^{1,2}; Angel Yasmil Echeverria Guzman³; Gary Reyes^{1,2}; Manuel Fabricio Reyes Wagnio^{1,2}; Maikel Yelandi Leyva Vázquez^{1,2}

¹Artificial Intelligence Research Group, Universidad Bolivariana del Ecuador, Km 5 ½ vía Yaguachi, Durán 092405, Ecuador.

²Grupo de Investigación en Inteligencia Artificial, Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, Guayaquil 090514, Ecuador.

³Universidad Bolivariana del Ecuador, Km 5 ½ vía Yaguachi, Durán 092405, Ecuador.

Abstract– The integration of artificial intelligence (AI)-assisted coding tools into programming education offers opportunities to enhance autonomy and immediate feedback in e-learning environments, but it also introduces risks associated with insufficient verification, instrumental dependence, and academic integrity. The overall objective of this study is to design a pedagogical guide for the integration of AI-assisted coding tools into e-learning/hybrid programming courses, based on empirical evidence and geared toward responsible, verifiable use that is compatible with authentic assessment. A non-experimental, cross-sectional, exploratory-descriptive quantitative study was conducted at the Bolivarian University of Ecuador with a sample of 111 students. An ad hoc questionnaire with 28 items (Likert 1–5) was administered, organized into four dimensions: use and frequency, perceived usefulness, trust/control, and ethics/risks. The instrument showed excellent internal consistency (overall $\alpha = 0.9736$). The results show high perceived usefulness ($M = 3.71$) and moderate ethical awareness ($M = 3.57$), along with heterogeneous adoption (use and frequency: $M = 3.29$). Significant positive correlations were observed between dimensions and differences between inconclusive degrees in most comparisons. Based on the diagnosis, an operational pedagogical guide for e-learning/hybrid learning was designed (rules of use, traceability log, mandatory verification, and authentic assessment), validated by expert judgment ($n = 12$, one round) with an overall average Aiken's V of 0.87. It is concluded that effective adoption requires integrated guidelines that articulate productivity, deep learning, and academic integrity.

Keywords– AI-assisted coding, language models (LLM), engineering education, authentic assessment

Adopción de herramientas de codificación asistida por IA en cursos de programación: evidencia y lineamientos para e-learning en ingeniería

Dayron Rumbaut-Rangel¹; Franklin Parrales-Bravo^{1,2}; Lorenzo Cevallos-Torres^{1,2}; Angel Yasmil Echeverria Guzman³; Gary Reyes^{1,2}; Manuel Fabricio Reyes Wagnio^{1,2}; Maikel Yelandi Leyva Vázquez^{1,2}

¹Artificial Intelligence Research Group, Universidad Bolivariana del Ecuador, Km 5 ½ vía Yaguachi, Durán 092405, Ecuador.

²Grupo de Investigación en Inteligencia Artificial, Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, Guayaquil 090514, Ecuador.

³Universidad Bolivariana del Ecuador, Km 5 ½ vía Yaguachi, Durán 092405, Ecuador.

Resumen— La integración de herramientas de codificación asistida por inteligencia artificial (IA) en la enseñanza de la programación ofrece oportunidades para potenciar la autonomía y la retroalimentación inmediata en entornos e-learning, pero también introduce riesgos asociados a verificación insuficiente, dependencia instrumental e integridad académica. El objetivo general de este estudio es diseñar una guía pedagógica de integración de herramientas de codificación asistida por IA para cursos de programación en modalidad e-learning/híbrida, sustentada en evidencia empírica y orientada a un uso responsable, verificable y compatible con la evaluación auténtica. Se desarrolló un estudio cuantitativo no experimental, transversal y exploratorio-descriptivo en la Universidad Bolivariana del Ecuador con una muestra de 111 estudiantes. Se aplicó un cuestionario ad hoc de 28 ítems (Likert 1–5) organizado en cuatro dimensiones: uso y frecuencia, utilidad percibida, confianza/control y ética/riesgos. El instrumento mostró excelente consistencia interna (α global = 0.9736). Los resultados evidencian alta utilidad percibida ($M = 3.71$) y conciencia ética moderada ($M = 3.57$), junto con una adopción heterogénea (uso y frecuencia: $M = 3.29$). Se observaron correlaciones positivas significativas entre dimensiones y diferencias entre titulaciones no concluyentes en la mayoría de comparaciones. Con base en el diagnóstico, se diseñó una guía pedagógica operativa para e-learning/híbrido (reglas de uso, bitácora de trazabilidad, verificación obligatoria y evaluación auténtica), validada mediante juicio de expertos ($n = 12$, una ronda) con Aiken's V promedio general de 0.87. Se concluye que la adopción efectiva requiere lineamientos integrados que articulen productividad, aprendizaje profundo e integridad académica.

Palabras clave— codificación asistida por IA, modelos de lenguaje (LLM), educación en ingeniería, evaluación auténtica

Herramientas basadas en modelos de lenguaje de gran tamaño, como asistentes de codificación, chatbots y sistemas de autocompletado inteligente, permiten sugerir, explicar, depurar, refactorizar y generar fragmentos de código a partir de instrucciones en lenguaje natural [1]. Este escenario consolida una dinámica de cocreación humano–IA que modifica la actividad tradicional de programar, desplazándola desde una práctica centrada exclusivamente en la escritura manual de código hacia procesos mediados por interacción, iteración, revisión y validación [2]. En modalidades e-learning e híbridas, esta transformación resulta especialmente relevante, dado que la retroalimentación inmediata, la autonomía del estudiante y la trazabilidad del proceso de aprendizaje adquieren un papel central [3].

Desde una perspectiva pedagógica, las herramientas de codificación asistida por IA pueden funcionar como recursos de andamiaje para apoyar la resolución de problemas, la comprensión de conceptos, la depuración de errores y el aprendizaje autónomo [4]. Estudios recientes señalan que estos sistemas pueden contribuir a reducir la carga cognitiva asociada a tareas rutinarias, facilitar la continuidad del trabajo académico y ofrecer retroalimentación personalizada en el proceso de aprendizaje [5], [6]. En el caso específico de la programación, se ha reportado que los estudiantes valoran positivamente la posibilidad de recibir explicaciones inmediatas, ejemplos alternativos y orientación durante la resolución de problemas, lo cual puede favorecer la autonomía y la productividad académica [7].

I. INTRODUCCIÓN

La irrupción de la inteligencia artificial generativa en los entornos de desarrollo de software está transformando las prácticas de enseñanza y aprendizaje de la programación.

Sin embargo, la integración educativa de estas herramientas también plantea riesgos relevantes. El uso acrítico de asistentes de IA puede favorecer dependencia instrumental, aprendizaje superficial, aceptación de código sin

verificación y debilitamiento del razonamiento algorítmico [8]. Además, algunos estudios han cuestionado si estos sistemas pueden sustituir procesos formativos como la programación colaborativa o el acompañamiento humano, dado que la interacción con la IA no siempre garantiza comprensión profunda ni desarrollo de habilidades de depuración sistemática [9]. De igual manera, investigaciones recientes advierten una paradoja entre el incremento de la autoeficacia percibida y una mayor dependencia tecnológica, lo que exige diseñar estrategias pedagógicas que promuevan control crítico y autorregulación [10].

A estos desafíos se suman preocupaciones sobre integridad académica, transparencia y responsabilidad en el uso educativo de la IA. Revisiones recientes sobre IA generativa, chatbots y sistemas tutoriales en educación de programación señalan beneficios importantes, pero también riesgos asociados al uso excesivo, la delegación cognitiva y la dificultad para verificar la autoría real de los productos entregados por los estudiantes [11]. En este sentido, el debate actual no se limita a permitir o prohibir estas herramientas, sino a establecer criterios claros de uso responsable, citación del apoyo recibido, trazabilidad del proceso y evaluación auténtica [12]. En cursos e-learning e híbridos, estos criterios resultan aún más necesarios, debido a que gran parte de las evidencias de aprendizaje se produce mediante entregables digitales y actividades asincrónicas.

La dimensión técnica también debe ser considerada. Aunque herramientas como asistentes de código y modelos entrenados sobre lenguajes de programación pueden mejorar la productividad, generar soluciones preliminares o apoyar la depuración, sus salidas no siempre son correctas, óptimas ni seguras [13]. Investigaciones sobre seguridad del código generado por asistentes de IA han evidenciado la posibilidad de producir fragmentos vulnerables cuando las sugerencias son aceptadas sin revisión sistemática [14]. Por ello, la formación en programación no debe orientarse únicamente a usar IA para “hacer más rápido”, sino a desarrollar competencias de verificación, prueba, explicación, refactorización, seguridad básica y responsabilidad académica.

Pese al creciente interés internacional por la IA generativa en la educación en programación, todavía existen vacíos de evidencia sobre los patrones reales de uso, la utilidad percibida, los niveles de confianza, las prácticas de verificación y las percepciones éticas de estudiantes de ingeniería en contextos latinoamericanos. Esta brecha es relevante porque la adopción de asistentes de codificación no depende solo de la disponibilidad tecnológica, sino también de factores formativos, curriculares e institucionales. Por tanto, resulta necesario generar evidencia situada que permita diseñar orientaciones aplicables a la práctica docente y a la toma de decisiones académicas en programas de ingeniería.

En este marco, el objetivo general del presente estudio es diseñar una guía pedagógica para la integración de herramientas de codificación asistida por IA en cursos de programación en modalidad e-learning/híbrida, sustentada en evidencia empírica y orientada a un uso responsable, verificable y compatible con la integridad académica. Para ello, se plantean como objetivos específicos: (i) caracterizar el nivel de uso y los patrones de adopción de estas herramientas en estudiantes de Ingeniería en Sistemas Inteligentes e Ingeniería en Inteligencia Artificial y Ciberseguridad de la Universidad Bolivariana del Ecuador; (ii) analizar sus percepciones sobre utilidad, confianza/control y ética/riesgos; (iii) derivar lineamientos operativos que incorporen usos permitidos/no permitidos, trazabilidad mediante bitácora, verificación obligatoria y evaluación auténtica; y (iv) validar la guía mediante juicio de expertos para estimar su validez de contenido. Con ello, el estudio aporta evidencia empírica y una propuesta pedagógica aplicable para orientar la incorporación responsable de la IA generativa en la formación de ingenieros.

II. METODOLOGÍA

A. Enfoque y diseño metodológico

El estudio se desarrolló con un enfoque cuantitativo, bajo un diseño no experimental y transversal. Metodológicamente, se estructuró en dos fases complementarias:

1. Fase 1 (diagnóstico exploratorio-descriptivo): caracterización de patrones de uso y percepciones estudiantiles sobre herramientas de codificación asistida por IA, considerando utilidad percibida, confianza/control y dimensiones éticas.
2. Fase 2 (diseño y validación de un recurso pedagógico): elaboración de una guía de integración para e-learning/híbrido derivada del diagnóstico y validada mediante juicio de expertos ($n = 12$, una ronda), con estimación de validez de contenido mediante Aiken's V.

B. Población y muestra

La población estuvo conformada por estudiantes de la Universidad Bolivariana del Ecuador pertenecientes a las carreras de Ingeniería en Sistemas Inteligentes ($n = 79$) e Ingeniería en Inteligencia Artificial y Ciberseguridad ($n = 32$). La muestra total incluyó 111 estudiantes, seleccionados mediante muestreo no probabilístico por conveniencia, atendiendo criterios de accesibilidad, pertinencia curricular y voluntariedad de participación. Dado el carácter exploratorio del estudio, la muestra se consideró suficiente para identificar tendencias y diferencias entre titulaciones.

La muestra se consideró adecuada para un estudio exploratorio-descriptivo porque permitió cubrir estudiantes de dos titulaciones directamente vinculadas con programación, inteligencia artificial y ciberseguridad, garantizando diversidad curricular y pertinencia temática. No obstante, al

tratarse de un muestreo no probabilístico por conveniencia, los resultados no se orientan a generalización estadística, sino a la identificación de tendencias empíricas y necesidades pedagógicas para el diseño de la guía.

C. Instrumento de recolección de datos

Se diseñó un cuestionario estructurado ad hoc, compuesto por 28 ítems en escala Likert de cinco puntos (1 = totalmente en desacuerdo, 5 = totalmente de acuerdo), organizados en cuatro dimensiones:

(1) Uso y frecuencia, (2) Utilidad percibida, (3) Confianza y control/ exactitud, y (4) Ética y riesgos. Adicionalmente, se incluyeron variables de caracterización (carrera, nivel de estudios, experiencia previa en programación y lenguaje de programación habitual).

Previo a su aplicación, el instrumento fue sometido a una revisión de contenido y consistencia semántica (claridad, redacción y coherencia de ítems con las dimensiones definidas). La confiabilidad interna se evaluó posteriormente mediante alfa de Cronbach por dimensión y global, interpretando su magnitud conforme criterios ampliamente usados en medición educativa (Nunnally & Bernstein, 1994).

D. Procedimiento

Fase 1 (diagnóstico): la recolección se realizó de forma virtual mediante un formulario en Microsoft Forms, distribuido a través de canales institucionales y asignaturas vinculadas con programación. Se garantizó consentimiento informado, confidencialidad y anonimato; la participación fue voluntaria y no tuvo repercusiones evaluativas.

Fase 2 (guía pedagógica): a partir de los resultados del diagnóstico se redactó una guía de integración pedagógica para el uso de IA en cursos de programación en modalidad e-learning/híbrida. Posteriormente, la guía y el formulario de evaluación fueron remitidos digitalmente a un panel de expertos para su valoración cuantitativa y cualitativa en una única ronda.

E. Análisis de datos

El procesamiento y análisis estadístico se efectuó en Python, utilizando bibliotecas como pandas y scipy. Se aplicaron:

Estadística descriptiva (frecuencias, medias, desviaciones estándar) para resumir cada dimensión del cuestionario.

Confiabilidad interna mediante alfa de Cronbach por dimensión y global [15].

Pruebas de normalidad (Shapiro–Wilk) para orientar la selección de pruebas comparativas.

Comparación bivariada entre titulaciones por dimensión: ante distribución no normal en varias dimensiones, se empleó Mann–Whitney U como prueba no paramétrica, reportando estadístico U y significancia (p).

Este análisis permitió identificar patrones generales (p. ej., mayor valoración de la utilidad percibida) y diferencias relevantes entre titulaciones (p. ej., diferencias en utilidad percibida y ética/riesgos).

F. Diseño de la guía pedagógica derivada del diagnóstico

La guía pedagógica se construyó explícitamente a partir de los hallazgos cuantitativos del diagnóstico, empleando un criterio de “traducción pedagógica” desde las dimensiones evaluadas hacia componentes operativos para cursos e-learning/híbridos.

En particular:

- La mayor media en Utilidad percibida fundamentó la inclusión de casos de uso permitidos (p. ej., apoyo conceptual, depuración guiada, refactorización y generación de pruebas), orientados a potenciar el aprendizaje y la eficiencia sin sustituir el razonamiento.
- La presencia de una media moderada en Ética y riesgos, junto con diferencias significativas entre titulaciones en esta dimensión, sustentó la incorporación de principios de integridad, delimitación de usos no permitidos, y un protocolo de transparencia/trazabilidad mediante bitácora.
- La media relativamente menor en Uso y frecuencia se consideró como oportunidad para proponer un módulo de inducción (orientación inicial) y recomendaciones de implementación gradual, con criterios de acompañamiento docente.
- La dimensión Confianza y control/ exactitud motivó la incorporación de prácticas de verificación, pruebas, revisión y explicación del código como requisito de aprendizaje auténtico.

Con ello, la guía se estructuró en: propósito y alcance; reglas de uso responsable; casos de uso permitidos/no permitidos; protocolo de uso con bitácora (antes–durante–después); estrategia didáctica para e-learning; recomendaciones de evaluación auténtica; y consideraciones de privacidad y seguridad.

G. Juicio de expertos y validez de contenido de la guía

La guía fue evaluada por 12 expertos (una ronda), seleccionados por su experiencia en docencia de programación/ingeniería, e-learning/diseño instruccional, integridad académica y/o uso educativo de IA. Cada sección/ítem de la guía se valoró en escala 1–5 bajo cinco criterios: claridad, pertinencia, coherencia, aplicabilidad y suficiencia, incorporando además observaciones cualitativas.

La validez de contenido se estimó mediante Aiken’s V por ítem y criterio, utilizando:

$$V = \frac{\Sigma(r-l_0)}{n(c-l_0)} \quad (1)$$

donde r es la puntuación asignada por el experto, $l_0=1$ el mínimo de la escala, $c=5$ el máximo y $n=12$ el número de expertos. Como referencia operativa, se consideraron valores $V \geq 0.70$ como aceptables y valores inferiores como indicadores de revisión editorial o de ajuste del contenido.

H. Consideraciones éticas

La investigación se condujo bajo principios de ética en investigación educativa: respeto a la autonomía, privacidad y confidencialidad. La participación fue voluntaria, anónima y con consentimiento informado. No se recolectaron datos personales sensibles ni se vinculó información individual con los resultados. El estudio se ajustó a los lineamientos institucionales y normativas aplicables a investigaciones en educación superior.

III. RESULTADOS

A. Confiabilidad del instrumento

Con el fin de evaluar la consistencia interna del cuestionario aplicado, se estimó el coeficiente alfa de Cronbach para cada una de las cuatro dimensiones: Uso y frecuencia, Utilidad percibida, Confianza y exactitud, y Ética y riesgos. Este análisis permite determinar en qué medida los ítems de cada dimensión se comportan de manera homogénea al medir un mismo constructo. Asimismo, se calculó el alfa global considerando los 28 ítems del instrumento.

TABLA I
COEFICIENTES ALFA DE CRONBACH POR DIMENSIÓN

ESCALA	ÍTEM	ALFA DE CRONBACH
Uso y frecuencia	7	0.9164750078733073
Utilidad percibida	7	0.9528576793708144
Confianza y exactitud	7	0.9240177291965208
Ética y riesgos	7	0.9322954461974002
Global (todos los ítems)	28	0.973617921674081

Todos los coeficientes superan 0.90, lo que indica una excelente consistencia interna de acuerdo con los criterios de referencia en medición psicométrica [15]. En conjunto, estos resultados aportan evidencia robusta de fiabilidad para el análisis descriptivo e inferencial presentado en las siguientes subsecciones, dentro del carácter exploratorio del estudio.

B. Caracterización de la muestra

Participaron 111 estudiantes de dos titulaciones: Ingeniería en Sistemas Inteligentes ($n = 79$; 71.2%) e Ingeniería en Inteligencia Artificial y Ciberseguridad ($n = 32$; 28.8%). En cuanto al sexo, se registró una mayor proporción de participantes masculinos ($n = 91$; 82%) en comparación con femeninos ($n = 20$; 18%), tal como se ilustra en la Figura 1.

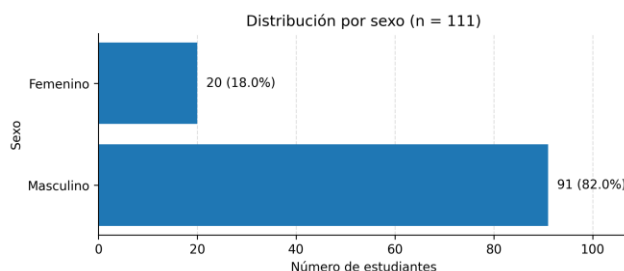


Fig. 1 Distribución por sexo.

Respecto a la experiencia previa en programación, predominaron niveles básicos (64.9%), seguidos de

intermedios (32.4%) y un porcentaje reducido avanzado (2.7%). En relación con el lenguaje de programación utilizado con mayor frecuencia, Python fue el más reportado (64.0%), seguido por JavaScript (10.8%), Java (9.0%) y C++ (5.4%). Esta caracterización contextualiza las percepciones y patrones de uso observados.

C. Análisis descriptivo por dimensiones

Se calcularon medidas de tendencia central y dispersión para cada dimensión del instrumento, complementadas con visualizaciones mediante diagramas de violín (Figura 2).

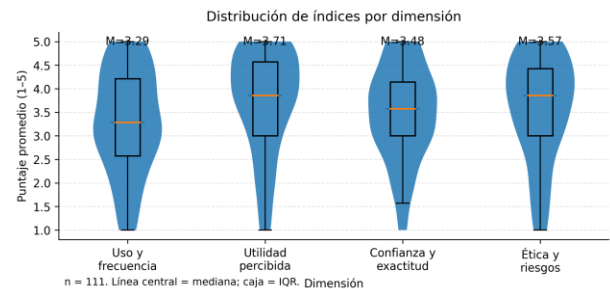


Fig. 2 Diagrama de violín índice de dimensiones.

La dimensión Utilidad percibida presenta la media más alta ($M = 3.71$), lo que evidencia una valoración favorable hacia el potencial de estas herramientas para apoyar el aprendizaje y el desempeño en tareas de programación. Le siguen Ética y riesgos ($M = 3.57$) y Confianza y exactitud ($M = 3.48$), lo cual sugiere una aceptación moderada del apoyo de la IA acompañada de conciencia sobre riesgos y limitaciones. Por su parte, Uso y frecuencia presenta la media más baja ($M = 3.29$), lo que indica que, aunque la IA es valorada positivamente, su adopción aún no está completamente generalizada.

La Figura 2 muestra distribuciones amplias (rango 1–5) en todas las dimensiones, con concentración relativamente mayor hacia valores altos en Utilidad percibida y Ética y riesgos, coherente con los valores promedio reportados.

D. Evidencia por ítems clave: adopción, control y ética

Para mejorar la interpretabilidad, se estimaron porcentajes de acuerdo alto (respuestas 4–5) en ítems representativos que resultan especialmente informativos para decisiones pedagógicas en e-learning:

Adopción y usos académicos:

“Utilizo herramientas de IA para codificar al menos una vez por semana”: 36.9% de acuerdo (4–5) y 36.0% en desacuerdo (1–2), sugiriendo una adopción heterogénea.

“Consulto la IA principalmente en tareas o proyectos académicos”: 40.5% de acuerdo, lo que indica presencia relevante del uso en actividades formativas.

“Utilizo la IA para aprender nuevos lenguajes o librerías”: 50.5% de acuerdo, evidenciando un rol frecuente como apoyo para autoaprendizaje.

Percepción de apoyo al aprendizaje:

“Las herramientas de IA me ayudan a comprender mejor los fundamentos”: 64.0% de acuerdo.

“Incrementan mi productividad”: 59.5% de acuerdo.

“Me permiten aprender de forma más autónoma”: 61.3% de acuerdo.

Estos indicadores refuerzan que la aceptación se sostiene principalmente en beneficios percibidos de comprensión, autonomía y eficiencia.

Control crítico y verificación:

“Verifico y entiendo cada línea de código antes de implementarla”: 56.8% de acuerdo, lo cual es positivo, aunque deja un margen relevante de estudiantes sin una práctica consistente de verificación.

“En ocasiones la IA genera errores o código redundante”: 54.1% de acuerdo, lo que sugiere conciencia crítica sobre limitaciones, pero también subraya la necesidad de protocolos explícitos de validación y pruebas.

Ética, integridad y necesidad de lineamientos:

“Considero necesario reconocer el apoyo de la IA en mis tareas”: 52.3% de acuerdo.

“Usar IA sin citar su intervención constituye una forma de plagio”: 45.0% de acuerdo y 27.0% en desacuerdo, evidenciando un vacío normativo/interpretativo que amerita orientación institucional.

“Las instituciones deberían establecer normas claras sobre su uso educativo”: 63.1% de acuerdo.

“La IA debe ser vista como herramienta complementaria, no sustitutiva”: 65.8% de acuerdo.

En conjunto, estos resultados muestran una legitimidad alta para establecer lineamientos y criterios compartidos de uso.

E. Asociación entre dimensiones

Con el fin de examinar la coherencia interna del patrón perceptual y la relación entre los constructos evaluados, se estimaron correlaciones no paramétricas de Spearman (ρ) entre las cuatro dimensiones del cuestionario. Los resultados se presentan en la Tabla 2.

TABLA II
COEFICIENTES DE CORRELACIÓN

Dimensión	Uso y frecuencia	Utilidad percibida	Confianza y exactitud	Ética y riesgos
Uso y frecuencia	1.000	0.714	0.727	0.654
Utilidad percibida	0.714	1.000	0.834	0.712
Confianza y exactitud	0.727	0.834	1.000	0.721
Ética y riesgos	0.654	0.712	0.721	1.000

En conjunto, las correlaciones positivas evidencian que un mayor nivel de uso se asocia con mayor utilidad percibida y, a su vez, con mayores niveles de confianza/control y mayor sensibilidad hacia aspectos éticos y de riesgo, lo que refuerza

la necesidad de lineamientos integrados para su incorporación en cursos de programación en modalidad e-learning.

Este patrón sugiere que, a medida que aumenta la adopción y la utilidad percibida, también tienden a incrementarse el sentido de control/confianza y la conciencia sobre implicaciones éticas y riesgos. Por tanto, la integración pedagógica debe abordarse como un conjunto articulado: uso + criterios de verificación + integridad académica.

F. Segmentación por nivel de adopción: usuarios frecuentes vs. no frecuentes

Dada la heterogeneidad observada en la dimensión Uso y frecuencia, se realizó una segmentación exploratoria definiendo como usuarios frecuentes a quienes presentaron un puntaje promedio ≥ 4 en dicha dimensión. Bajo este criterio, el 34.2% de la muestra se clasificó como usuario frecuente ($n = 38$) frente a un 65.8% no frecuente ($n = 73$).

Los usuarios frecuentes mostraron puntajes significativamente mayores en:

- Utilidad percibida ($p < 0.001$; efecto grande),
- Confianza y exactitud ($p < 0.001$; efecto grande), y
- Ética y riesgos ($p < 0.001$; efecto moderado-grande).

Este resultado es clave: la adopción intensiva se asocia no solo con mayor utilidad y confianza, sino también con mayor conciencia de riesgos, lo que refuerza la necesidad de formación guiada para que la adopción se traduzca en aprendizaje auténtico y prácticas responsables.

Adicionalmente, se observó una correlación positiva pequeña entre semestre y Uso y frecuencia ($\rho \approx 0.20$; $p = 0.038$), lo que sugiere una tendencia de incremento gradual de adopción conforme avanza la trayectoria académica, aunque el efecto es bajo.

G. Análisis comparativo entre titulaciones

Con el objetivo de identificar diferencias por carrera, y dado que las pruebas de normalidad (Shapiro-Wilk) indicaron desviaciones significativas respecto a la normalidad en las dimensiones ($p < 0.01$), se aplicó la prueba no paramétrica Mann-Whitney U para comparar titulaciones.

Los resultados muestran:

Uso y frecuencia: diferencia marginal ($U = 1560.5$; $p = 0.054$), con tendencia a mayor puntaje en Sistemas Inteligentes ($M = 3.43$) respecto a IA y Ciberseguridad ($M = 2.94$), con tamaño de efecto pequeño.

Utilidad percibida: no se observaron diferencias significativas ($U = 1419.5$; $p = 0.312$).

Confianza y exactitud: no se observaron diferencias significativas ($U = 1324.0$; $p = 0.698$).

Ética y riesgos: no se observaron diferencias significativas ($U = 1368.0$; $p = 0.500$).

En términos generales, las percepciones sobre utilidad, confianza/control y ética/riesgos se muestran relativamente

homogéneas entre titulaciones en esta muestra, mientras que el nivel de uso presenta variación leve.

H. Implicaciones pedagógicas del diagnóstico

En conjunto, los hallazgos muestran un escenario consistente: (i) alta utilidad percibida y aceptación de la IA como herramienta de apoyo; (ii) adopción heterogénea (uso semanal no generalizado); (iii) presencia de prácticas de verificación, aunque no uniformes; y (iv) demanda explícita de normas institucionales y criterios compartidos sobre citación, integridad y uso responsable. Además, la segmentación evidencia que el mayor uso se acompaña de mayor utilidad, control y conciencia ética, lo que sugiere que el desafío principal no es restringir la adopción, sino orientarla pedagógicamente con reglas claras, trazabilidad y evaluación auténtica, especialmente en entornos e-learning/híbridos.

Con base en esta evidencia, en la siguiente sección se presenta la guía pedagógica de uso e integración de herramientas de codificación asistida por IA, elaborada a partir del diagnóstico, así como su validación mediante juicio de expertos (n = 12, una ronda) para estimar validez de contenido y fortalecer su aplicabilidad como recurso instruccional.

IV. PROPUESTA PEDAGÓGICA

Guía para la integración responsable de herramientas de codificación asistida por IA en e-learning de ingeniería.

A. Propósito y fundamento de la propuesta

A partir del diagnóstico, se identificó un escenario caracterizado por: (i) una utilidad percibida alta de las herramientas de codificación asistida por IA, (ii) un uso heterogéneo (no generalizado), (iii) prácticas de verificación presentes pero no uniformes, y (iv) una demanda clara de normas institucionales y criterios compartidos sobre integridad y transparencia. Estos hallazgos justifican el diseño de una propuesta pedagógica orientada a e-learning que no se limite a permitir o prohibir el uso de IA, sino que lo operacionalice mediante reglas, procedimientos y evidencias de aprendizaje.

En consecuencia, se plantea una guía de uso que traduce los resultados en lineamientos aplicables en cursos de programación en modalidad e-learning/híbrida, articulando tres metas:

- Aprendizaje profundo (comprensión y razonamiento algorítmico),
- Uso responsable y trazable (integridad académica, transparencia), y
- Control técnico (verificación, pruebas, seguridad básica del código).

B. Principios de diseño instruccional para e-learning

La guía se estructura sobre principios que responden directamente a los riesgos y oportunidades identificados:

Complementariedad pedagógica: la IA se integra como apoyo al aprendizaje, no como sustituto del trabajo cognitivo del estudiante.

Transparencia y trazabilidad: todo uso relevante de IA debe quedar documentado mediante una bitácora mínima (prompts, salidas, decisiones y modificaciones).

Verificación obligatoria: ninguna sugerencia de IA se acepta sin comprobación (pruebas, revisión, ejecución y justificación).

Autoría y responsabilidad: el estudiante conserva responsabilidad por el producto final (código, decisiones, explicaciones).

Seguridad y calidad: se promueve la revisión de vulnerabilidades comunes, calidad de código y buenas prácticas de documentación, acorde con el contexto formativo.

Evaluación auténtica: la evaluación se centra en evidencias de proceso, explicación y defensa breve del código, reduciendo la posibilidad de delegación acrítica.

C. Estructura de la guía pedagógica

La guía se organiza en secciones operativas para su adopción docente:

C.1. Alcance y definiciones operativas

Delimita qué se entiende por “herramientas de codificación asistida por IA”, tipos de apoyo (explicación, generación parcial, depuración, refactorización) y el alcance de la guía (tareas, proyectos, evaluaciones).

C.2. Reglas de uso responsable y ética académica

Incluye normas explícitas sobre:

- Transparencia: obligación de declarar cuándo y cómo se usó IA.
- Citación o reconocimiento del apoyo de IA: criterios mínimos de reporte en entregas académicas.
- Privacidad y datos: prohibición de ingresar datos sensibles, credenciales o información confidencial.
- Conductas no permitidas: uso de IA para generar entregables completos sin comprensión, suplantación de autoría o evasión de objetivos de aprendizaje.

C.3. Casos de uso permitidos y no permitidos (con ejemplos)

Se establecen categorías con ejemplos prácticos (prompting seguro y pedagógico):

Permitidos (orientados a aprendizaje): explicación conceptual, sugerencias de estrategia, depuración guiada, refactorización con justificación, generación de pruebas, documentación y análisis de complejidad (cuando aplique).

No permitidos (riesgo de superficialidad/deshonestidad): entrega de soluciones completas sin evidencia de proceso, respuestas directas a evaluaciones cerradas, generación de código sin verificación ni explicación, y uso de IA para ocultar autoría.

C.4. Protocolo Antes–Durante–Después (bitácora obligatoria)

Se propone un procedimiento simple, compatible con e-learning:

Antes: definir objetivo de aprendizaje, restricciones de la tarea y qué tipo de ayuda se solicitará a la IA (no “solución final”).

Durante: registrar prompts/salidas clave y aplicar verificación (ejecución, revisión, pruebas).

Después: documentar qué se aceptó, qué se rechazó y por qué; incluir breve reflexión (qué aprendí, qué corrigió la IA, qué validé yo).

La bitácora funciona como evidencia de autoría y como herramienta metacognitiva para fortalecer control y aprendizaje.

D. Estrategia didáctica para implementación en e-learning

Para facilitar adopción sin sobrecarga docente, la propuesta plantea una secuencia modular:

D.1. Módulo 0: Inducción al uso responsable (microlección + actividad diagnóstica)

Microcontenido: límites, riesgos, reglas y ejemplos de prompting.

Actividad breve: analizar un fragmento de código sugerido por IA, detectar errores y justificar correcciones.

Resultado: el estudiante demuestra comprensión de reglas, verificación y reporte.

D.2. Actividades guiadas durante el curso (aprendizaje basado en tareas)

Tarea tipo A (comprensión): explicar línea por línea un código sugerido, proponer mejoras y justificar.

Tarea tipo B (depuración): la IA sugiere hipótesis; el estudiante aplica pruebas y reporta evidencias.

Tarea tipo C (refactorización): mejorar legibilidad/calidad y comparar versiones con métricas simples.

Tarea tipo D (pruebas): diseñar casos de prueba, ejecutar y reportar cobertura/errores hallados.

Estas actividades están diseñadas para promover aprendizaje profundo, no solo productividad.

E. Estrategia de evaluación auténtica y aseguramiento de integridad

La guía recomienda que las evaluaciones en e-learning incorporen evidencias de proceso, reduciendo la dependencia de un único entregable final:

Evidencias obligatorias: código + bitácora + pruebas + breve reflexión.

Microdefensa (sincrónica o asincrónica): explicación corta (2–5 minutos) de decisiones, errores encontrados y cómo se verificó la solución.

Rúbrica de autenticidad: criterios como trazabilidad, coherencia entre bitácora y producto final, capacidad de explicar, calidad de pruebas y control del código.

Este enfoque mantiene la flexibilidad del e-learning y fortalece la validez de la evaluación del aprendizaje.

F. Roles y responsabilidades para la adopción

Para viabilizar la propuesta, se delimitan responsabilidades:

Docente: define qué usos son permitidos por actividad, exige bitácora, y evalúa proceso + producto.

Estudiante: mantiene transparencia, verifica el código y justifica decisiones.

Coordinación académica (opcional): establece lineamientos comunes y ejemplos institucionales para coherencia entre asignaturas.

G. Producto de la propuesta y transición a la validación

Como resultado, se dispone de una guía pedagógica operativa para cursos de programación en e-learning/híbrido, incluyendo: (i) reglas claras de uso, (ii) protocolo de trazabilidad, (iii) actividades tipo, y (iv) estrategia de evaluación auténtica.

En la siguiente sección se presenta la validación de contenido de la guía mediante juicio de expertos (n = 12, una ronda), con el fin de estimar su claridad, pertinencia, coherencia, aplicabilidad y suficiencia, y consolidarla como un recurso instruccional listo para adopción.

H. Diseño de la validación

Con el objetivo de estimar la validez de contenido de la guía pedagógica para la integración de herramientas de codificación asistida por IA en cursos de programación (modalidad e-learning/híbrida), se aplicó un juicio de expertos en una (1) ronda, con un panel de 12 expertos (n = 12). La evaluación se realizó mediante un instrumento estructurado que valora cada componente de la guía a partir de cinco criterios: claridad, pertinencia, coherencia, aplicabilidad y suficiencia, utilizando una escala ordinal de 1 a 5 (1 = muy deficiente, 2 = deficiente, 3 = aceptable, 4 = bueno, 5 = excelente).

La guía se descompuso en 18 elementos evaluables (secciones y componentes operativos), incluyendo apartados como propósito y alcance, definiciones, reglas de uso responsable, usos permitidos/no permitidos, protocolo de trazabilidad, estrategia didáctica, evaluación auténtica, rúbrica, integridad/privacidad y consistencia global.

I. Resultados globales de validez por criterio

En términos globales, la guía obtuvo un Aiken's V promedio general de 0.87, con un rango de 0.80 a 0.93 considerando el promedio por elemento. La Tabla 3 resume los valores promedio, mínimo y máximo de Aiken's V por criterio.

TABLA III
COEFICIENTES DE AIKEN'S

Criterio	V promedio	V mín.	V máx.
Claridad	0.86	0.72	0.97
Pertinencia	0.88	0.78	0.97
Coherencia	0.86	0.78	0.97

Aplicabilidad	0.87	0.72	0.97
Suficiencia	0.89	0.75	0.97

Estos resultados evidencian que, en promedio, los expertos valoraron la guía como buena en los cinco criterios, destacando especialmente suficiencia ($\bar{V} = 0.89$) y pertinencia ($\bar{V} = 0.88$), lo cual respalda que el contenido cubre adecuadamente los componentes esenciales y es relevante para su implementación en e-learning.

J. Resultados por elemento de la guía

La Tabla 3 presenta los valores de Aiken's V por elemento y criterio, junto con el V promedio por elemento. Todos los elementos obtuvieron $V \geq 0.80$, por lo que superan el umbral de aceptación definido y se clasifican como "BUENO".

En cuanto a los elementos con mejor valoración, destacan los usos permitidos orientados a refactorización ($V \text{ prom.} = 0.93$) y los componentes de depuración y definiciones operativas ($V \text{ prom.} \approx 0.90$). Por otra parte, aunque todos los elementos alcanzan niveles aceptables, se identificaron valores relativamente menores en: (i) usos permitidos de carácter conceptual ($V \text{ prom.} = 0.80$), lo que sugiere fortalecer ejemplos o delimitar mejor este tipo de apoyo para evitar ambigüedad; (ii) la suficiencia de la sección de transparencia/bitácora (Suficiencia = 0.75), lo que apunta a enriquecer campos mínimos o incluir un ejemplo completo diligenciado; (iii) la claridad del componente de integridad, privacidad y seguridad (Claridad = 0.72), recomendándose una redacción más directa y con ejemplos de "datos no permitidos"; y (iv) la aplicabilidad del apartado de usos no permitidos/faltas (Aplicabilidad = 0.72), que podría beneficiarse de una formulación modular adaptable a reglamentos institucionales.

V. DISCUSIÓN

Los hallazgos de este estudio deben interpretarse a la luz de la literatura reciente sobre IA generativa y asistentes de codificación en educación superior, la cual muestra un escenario ambivalente: por un lado, estas herramientas son valoradas positivamente por su capacidad para ofrecer retroalimentación inmediata, apoyar la resolución de problemas y favorecer la autonomía; por otro, su uso no guiado puede generar dependencia instrumental, aprendizaje superficial y debilitamiento de la autoría académica. En este sentido, la alta utilidad percibida identificada en la muestra analizada coincide con estudios recientes que reportan una valoración favorable de la IA generativa como apoyo para tareas académicas y de programación [5], [7]. Sin embargo, dicha utilidad no debe asumirse automáticamente como evidencia de mejora del aprendizaje, pues investigaciones recientes advierten que la percepción de eficiencia puede coexistir con riesgos de dependencia tecnológica y reducción del esfuerzo cognitivo independiente [10].

En particular, los resultados de Lepp y Kaimre [7] aportan un contraste relevante para la interpretación crítica de este estudio. Aunque dichos autores encontraron que los estudiantes utilizan chatbots de IA principalmente para tareas de programación y los perciben como recursos útiles, también reportaron una correlación negativa entre la frecuencia de uso y el rendimiento académico. Esta evidencia obliga a matizar la lectura optimista de la utilidad percibida: que los estudiantes consideren útil la IA no implica necesariamente que su uso produzca aprendizajes más profundos. Desde esta perspectiva, el hallazgo de una utilidad percibida elevada en el presente estudio debe entenderse como una oportunidad pedagógica, pero también como una señal de alerta. La incorporación de asistentes de codificación requiere diseñar actividades que obliguen al estudiante a explicar, verificar y justificar las soluciones, evitando que la herramienta sustituya procesos esenciales como el razonamiento algorítmico, la depuración sistemática y la comprensión del código.

La adopción heterogénea encontrada en este trabajo también se alinea con investigaciones recientes que señalan que el uso de asistentes de codificación con IA depende de factores cognitivos, motivacionales y de confianza. Por ejemplo, estudios basados en modelos de aceptación tecnológica han mostrado que la utilidad percibida, el interés por aprender, las metas de logro y la confianza influyen en la intención y el uso efectivo de estas herramientas en estudiantes de computación [16]. Este resultado es coherente con las correlaciones positivas observadas en el presente estudio entre uso/frecuencia, utilidad percibida, confianza/control y ética/riesgos. No obstante, mientras la literatura sobre aceptación tecnológica tiende a explicar la adopción desde variables individuales, los resultados aquí obtenidos sugieren que la adopción responsable también depende de condiciones institucionales y didácticas: reglas claras, acompañamiento docente, criterios de trazabilidad y evaluación auténtica.

Otro aspecto crítico se relaciona con la tensión entre confianza y control. En el presente estudio, más de la mitad de los estudiantes declara verificar y comprender el código antes de implementarlo; sin embargo, también se reconoce que las herramientas pueden generar errores o código redundante. Esta combinación muestra una conciencia inicial sobre las limitaciones de la IA, pero evidencia que las prácticas de verificación aún no son completamente uniformes. La literatura reciente coincide en que el problema no radica únicamente en el acceso a la IA, sino en la falta de estrategias pedagógicas para desarrollar una confianza calibrada: los estudiantes deben aprender cuándo apoyarse en la herramienta, cuándo desconfiar de sus salidas y cómo comprobar técnicamente la calidad, seguridad y pertinencia del código generado [11], [14]. Por ello, la propuesta de incorporar verificación obligatoria, pruebas, explicación del código y bitácora de uso responde directamente a una

necesidad identificada tanto en los datos del estudio como en el debate internacional reciente.

Desde la perspectiva ética, los resultados muestran una zona de ambigüedad especialmente relevante: aunque una parte importante de los estudiantes considera necesario reconocer el apoyo de la IA, no existe consenso pleno sobre si su uso no declarado constituye una forma de plagio. Este hallazgo coincide con preocupaciones recientes sobre integridad académica, transparencia y responsabilidad en el uso educativo de IA generativa [12]. En entornos e-learning e híbridos, esta problemática adquiere mayor complejidad, ya que las evidencias de aprendizaje suelen estar mediadas por plataformas digitales y entregables asincrónicos. Por tanto, la evaluación no puede centrarse exclusivamente en el producto final, sino que debe incorporar evidencias de proceso, tales como bitácoras de interacción, pruebas ejecutadas, reflexión metacognitiva y microdefensas. Esta orientación permite desplazar el foco desde la simple detección del uso de IA hacia la demostración de comprensión, autoría y control técnico.

En conjunto, la comparación con estudios recientes permite afirmar que el aporte principal de este trabajo no reside únicamente en confirmar que los estudiantes perciben útil la IA para programar, sino en traducir esa evidencia en una guía pedagógica operativa para orientar su uso responsable. Frente a investigaciones que advierten sobre dependencia, errores, sobreconfianza o integridad académica [7], [10], [11], [12], este estudio propone una respuesta instruccional concreta basada en cuatro mecanismos: delimitación de usos permitidos y no permitidos, trazabilidad mediante bitácora, verificación obligatoria del código y evaluación auténtica. De este modo, la guía no plantea una postura prohibitiva, sino una integración regulada y formativa, donde la IA funciona como andamiaje para el aprendizaje y no como sustituto del razonamiento del estudiante.

Asimismo, el contexto latinoamericano del estudio constituye un elemento diferenciador. Aunque la literatura internacional sobre IA en programación ha crecido rápidamente, todavía son necesarios estudios situados que analicen cómo estas herramientas son adoptadas en instituciones de educación superior de América Latina, particularmente en carreras de ingeniería vinculadas con inteligencia artificial, sistemas inteligentes y ciberseguridad. En este sentido, los resultados aportan evidencia empírica desde Ecuador y dialogan con estudios recientes desarrollados en otros contextos regionales, lo que contribuye a ampliar la comprensión de la adopción de asistentes de codificación más allá de escenarios anglosajones o altamente tecnologizados. Sin embargo, debido al carácter transversal, exploratorio y basado en autorreporte, los hallazgos deben interpretarse con cautela y ser complementados en el futuro con estudios longitudinales, mediciones de desempeño, análisis de calidad

del código y evaluación del impacto real de la guía en cursos de programación.

CONCLUSIONES

El presente estudio caracterizó la adopción y percepción de herramientas de codificación asistida por IA en estudiantes de ingeniería, y transformó ese diagnóstico en una propuesta pedagógica aplicable a cursos de programación en modalidad e-learning/híbrida. En conjunto, los resultados muestran que los estudiantes mantienen una valoración favorable sobre el aporte de estas herramientas, especialmente en términos de utilidad para apoyar el aprendizaje, la productividad y la autonomía. Sin embargo, la dimensión de uso y frecuencia se ubicó en un nivel relativamente menor, lo que evidencia una adopción todavía heterogénea y sugiere que el beneficio educativo no depende solo del acceso a la tecnología, sino de contar con orientaciones que faciliten un uso formativo y consistente.

La evidencia por ítems revela un hallazgo relevante para la docencia en e-learning: aunque una proporción importante de estudiantes declara verificar y comprender el código antes de implementarlo, persisten señales de riesgo asociadas a la variabilidad en prácticas de control, a la presencia de errores o redundancias en las salidas de IA, y a la ambigüedad ética sobre transparencia y citación del apoyo de IA. De manera consistente, los estudiantes expresan una expectativa clara de contar con normas institucionales y criterios comunes, lo cual legitima la construcción de lineamientos explícitos que reduzcan interpretaciones divergentes y favorezcan integridad académica.

El análisis de relaciones entre dimensiones refuerza que la incorporación de IA no debe abordarse de forma fragmentada: mayores niveles de uso se asocian con mayor utilidad percibida y, a su vez, con mayor sentido de control/confianza y mayor sensibilidad ante riesgos. Adicionalmente, la segmentación exploratoria sugiere que quienes utilizan la IA con mayor frecuencia tienden a reportar mayor utilidad y control, sin que ello implique menor conciencia ética; por el contrario, se observa una tendencia a mayor sensibilidad frente a riesgos, lo que respalda la conveniencia de un enfoque de adopción acompañada más que uno meramente restrictivo.

Como respuesta a estos hallazgos, se diseñó una guía pedagógica operativa que establece reglas de uso responsable, protocolo de trazabilidad (bitácora), verificación obligatoria, ejemplos de casos permitidos/no permitidos y recomendaciones de evaluación auténtica para entornos virtuales. La guía fue sometida a juicio de expertos ($n = 12$, una ronda) y obtuvo evidencia de validez de contenido alta mediante Aiken's V, con valoraciones favorables en claridad, pertinencia, coherencia, aplicabilidad y suficiencia. Este resultado sustenta la propuesta como un recurso instructivo viable para orientar el uso de IA en cursos de programación, particularmente en modalidades e-learning donde la trazabilidad y la autenticidad de evidencias son críticas.

Entre las implicaciones principales, se concluye que la integración efectiva de codificación asistida por IA requiere: (i) normas explícitas de transparencia e integridad (incluyendo reporte del uso), (ii) prácticas sistemáticas de verificación y pruebas como parte del aprendizaje, y (iii) estrategias de evaluación auténtica centradas en el proceso (bitácora, pruebas, explicaciones y microdefensas). De este modo, la IA puede aprovecharse como andamiaje para el aprendizaje sin comprometer la comprensión profunda ni la responsabilidad académica.

Finalmente, este trabajo presenta limitaciones propias de su alcance: el muestreo por conveniencia y el enfoque transversal limitan la generalización, y las variables analizadas se basan en autorreporte. Como línea futura, se recomienda evaluar el impacto de la guía en resultados de aprendizaje (calidad del código, pruebas, desempeño en evaluaciones auténticas), realizar estudios longitudinales y comparar implementaciones por asignatura o metodología docente. Asimismo, resulta pertinente analizar el efecto de la guía en indicadores de integridad académica y en la capacidad de los estudiantes para justificar decisiones de diseño y seguridad en el código en escenarios reales de e-learning.

AGRADECIMIENTO/RECONOCIMIENTO

Los autores agradecen a la Universidad Bolivariana del Ecuador por el respaldo institucional brindado para el desarrollo de esta investigación. De manera especial, expresan su reconocimiento al Artificial Intelligence Research Group por su apoyo académico, aportes técnicos y acompañamiento durante el proceso de levantamiento y análisis de datos, así como en la consolidación de la propuesta pedagógica presentada.

REFERENCES

- [1] S. S. Oyeler and K. Aruleba, "A comparative study of student perceptions on generative AI in programming education across Sub-Saharan Africa," *Comput. Educ. Open*, vol. 8, p. 100245, Jun. 2025, doi: 10.1016/j.caeo.2025.100245.
- [2] W. Holmes, M. Bialik, and C. Fadel, *Artificial Intelligence in Education. Promise and Implications for Teaching and Learning*, 2019.
- [3] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education – where are the educators?," *Int. J. Educ. Technol. High. Educ.*, vol. 16, no. 1, p. 39, Dec. 2019, doi: 10.1186/s41239-019-0171-0.
- [4] M. Madanchian, G. Drazenovic, S. R. Ramzani, and H. Taherdoost, "Integrating AI Tools to Enhance Learning Outcomes in Modern Education Systems," *Procedia Comput. Sci.*, vol. 263, pp. 514–521, 2025, doi: 10.1016/j.procs.2025.07.062.
- [5] J. Iqbal, Z. F. Hashmi, M. Z. Asghar, and M. N. Abid, "Generative AI tool use enhances academic achievement in sustainable education through shared metacognition and cognitive offloading among preservice teachers," *Sci. Rep.*, vol. 15, no. 1, p. 16610, May 2025, doi: 10.1038/s41598-025-01676-x.
- [6] S. Ba, L. Yang, Z. Yan, C. K. Looi, and D. Gašević, "Unraveling the mechanisms and effectiveness of AI-assisted feedback in education: A systematic literature review," *Comput. Educ. Open*, vol. 9, p. 100284, Dec. 2025, doi: 10.1016/j.caeo.2025.100284.
- [7] M. Lepp and J. Kaimre, "Does generative AI help in learning programming: Students' perceptions, reported use and relation to performance," *Comput. Hum. Behav. Reports*, vol. 18, p. 100642, May 2025, doi: 10.1016/j.chbr.2025.100642.
- [8] N. B. P. Duy, T. N. M. Phuong, V. N. M. Chau, N. V. H. Nhi, V. T. M. Khuyen, and N. T. P. Giang, "AI-assisted learning: an empirical study on student application behavior," *Multidiscip. Sci. J.*, vol. 7, no. 6, p. 2025275, Nov. 2024, doi: 10.31893/multiscience.2025275.
- [9] S. Imai, "Is GitHub copilot a substitute for human pair-programming?," in *Proceedings of the ACM/IEEE 44th International Conference on Software Engineering: Companion Proceedings*, May 2022, pp. 319–321, doi: 10.1145/3510454.3522684.
- [10] L. Zhang and J. Xu, "The paradox of self-efficacy and technological dependence: Unraveling generative AI's impact on university students' task completion," *Internet High. Educ.*, vol. 65, p. 100978, Apr. 2025, doi: 10.1016/j.iheduc.2024.100978.
- [11] S. Elnaffar, F. Rashidi, and A. Z. Abualkashik, "Teaching with AI: A Systematic Review of Chatbots, Generative Tools, and Tutoring Systems in Programming Education," Oct. 2025, [Online]. Available: <http://arxiv.org/abs/2510.03884>.
- [12] A. Amigud and D. J. Pell, "Responsible and Ethical Use of AI in Education: Are We Forcing a Square Peg into a Round Hole?," *World*, vol. 6, no. 2, p. 81, Jun. 2025, doi: 10.3390/world6020081.
- [13] M. Chen *et al.*, "Evaluating Large Language Models Trained on Code," Jul. 2021, doi: <https://doi.org/10.48550/arXiv.2107.03374>.
- [14] H. Pearce, B. Ahmad, B. Tan, B. Dolan-Gavitt, and R. Karri, "Asleep at the Keyboard? Assessing the Security of GitHub Copilot's Code Contributions," Dec. 2021, [Online]. Available: <http://arxiv.org/abs/2108.09293>.
- [15] J. C. Nunnally and I. H. Bernstein, *Psychometric Theory*, no. n.º972. McGraw-Hill Companies, Incorporated, 1994.
- [16] F. Ali, A. Ahmed, M. A. Alipour, and H. Terashima-Marin, "Adoption of AI-coding assistants in programming education: exploring trust and learning motivation through an extended technology acceptance model," *Journal of Computers in Education*, 2025, doi: 10.1007/s40692-025-00375-w.