



EL LENGUAJE MATEMÁTICO ANTE LA INTELIGENCIA ARTIFICIAL GENERATIVA: ANÁLISIS CONCEPTUAL DE SU INCIDENCIA EN LAS RESPUESTAS DE ESTUDIANTES, DOCENTES Y UNA TIPOLOGÍA DE PROMPTS

MATHEMATICAL LANGUAGE IN THE FACE OF GENERATIVE ARTIFICIAL INTELLIGENCE: A CONCEPTUAL ANALYSIS OF ITS INCIDENCE ON STUDENTS' AND TEACHERS' RESPONSES AND A TYPOLOGY OF PROMPTS

Reiman Yitsak Acuña Chacón¹



ORCID iD: <https://orcid.org/0000-0002-0094-3282>

RESUMEN

La incorporación de la inteligencia artificial (IA) generativa en la educación matemática se discute, sobre todo, en clave de integridad académica y de gobernanza institucional. Este artículo, de naturaleza teórico-metodológica y no empírica, desplaza la mirada hacia el lenguaje matemático mismo, entendido, desde Duval, Pimm y Alcalá, como el medio semiótico, discursivo y notacional en el que se constituyen los objetos matemáticos. Sostiene que los modelos generativos no solo resuelven problemas, sino que reorganizan la actividad semiótica de quienes los usan. El estudio se sustenta en una revisión narrativa y crítica de literatura reciente, complementada con una lectura documental de marcos institucionales. A partir de este abordaje, se examina la incidencia de la inteligencia artificial en las respuestas producidas por estudiantes y docentes, mediante la articulación de las nociones de fluidez sintética, génesis instrumental y contrato didáctico. Sobre esa base se deriva una tipología de seis categorías de prompts en las que el manejo del lenguaje matemático es la decisión de diseño central, orientada a preservar los procesos de representar y comunicar; cada categoría se ilustra con ejemplos y con episodios, de carácter exploratorio e ilustrativo, de una experiencia de aula desarrollada en un curso universitario de ecuaciones diferenciales. El aporte es un marco inicial para orientar la práctica de aula y la formación docente, susceptible de validación empírica posterior. Se concluye que el riesgo central no es que la IA produzca matemática incorrecta, sino matemática impecable y uniforme, que disuelve la fricción entre registros de la que depende el aprendizaje.

¹ Instituto Tecnológico de Costa Rica, Cartago, Costa Rica. Correo electrónico: reiacuna@itcr.ac.cr



Palabras clave: lenguaje matemático, registros de representación semiótica, inteligencia artificial generativa, prompts, educación matemática.

ABSTRACT

The integration of generative artificial intelligence (AI) into mathematics education is mostly discussed in terms of academic integrity and institutional governance. This theoretical-methodological and non-empirical paper shifts the focus toward mathematical language itself, understood, following Duval, Pimm and Alcalá, as the semiotic, discursive and notational medium in which mathematical objects are constituted, and argues that generative models not only solve problems but reorganize the semiotic activity of those who use them. The study is based on a narrative and critical review of recent literature, complemented by a documentary analysis of institutional frameworks. From this approach, it examines the influence of artificial intelligence on the responses produced by students and teachers by articulating the notions of synthetic fluency, instrumental genesis, and didactical contract. On this basis it derives a typology of six categories of prompts in which the handling of mathematical language is the central design decision, aimed at preserving the processes of representing and communicating; each category is illustrated with examples and with episodes, of an exploratory and illustrative nature, from a classroom experience in a university differential-equations course. The contribution is an initial framework to guide classroom practice and teacher education, open to subsequent empirical validation. It concludes that the central risk is not that AI produces incorrect mathematics, but impeccable and uniform mathematics that dissolves the friction between registers on which learning depends.

Keywords: mathematical language; semiotic registers of representation; generative artificial intelligence; prompts; mathematics education.

1. INTRODUCCIÓN

La irrupción de las herramientas de inteligencia artificial (IA) generativa ha transformado, en muy pocos años, las preguntas que docentes e investigadores se formulan sobre la enseñanza de la matemática. Las revisiones disponibles coinciden en que estas tecnologías abren oportunidades de personalización, retroalimentación y acceso, a la vez que plantean desafíos relativos a la integridad académica, los sesgos y la dependencia cognitiva (Kasneci et al., 2023; Qian, 2025; Walkington, 2025; Zawacki-Richter et al., 2019). En paralelo, numerosos sistemas educativos han comenzado a regular su uso mediante políticas públicas y marcos institucionales (MICITT, 2024).

El debate ha tendido a concentrarse en dos ejes: la integridad académica (¿produjo la respuesta el estudiante o la máquina?) y la gobernanza ética (¿qué reglas aseguran un uso responsable?). Ambos ejes son legítimos y necesarios; sin embargo, comparten el supuesto de que el problema se agota en quién produce la respuesta y bajo qué normas, de modo que bastaría un buen detector y una buena política para resolverlo.

Este trabajo sostiene que hay una cuestión más profunda y menos atendida. La matemática no es únicamente un conjunto de respuestas correctas, sino, ante todo, una actividad que se realiza en y a través de un lenguaje. Quien interviene ese lenguaje interviene la actividad. La pregunta decisiva no es solo quién produjo la respuesta, sino qué le ocurre al lenguaje matemático cuando estudiantes y docentes lo producen mediados por un mismo modelo.

La cuestión es de interés internacional, pues atañe a cualquier sistema que incorpore IA generativa en la enseñanza de la matemática. A manera de caso ilustrativo, este artículo recurre a la experiencia costarricense, cuya Reforma de la Educación Matemática incluyó, desde 2012, la representación y la comunicación entre sus cinco procesos fundamentales (MEP, 2012; Ruiz, 2013), donde marcos curriculares internacionales, como los estándares del National Council of Teachers of Mathematics (NCTM, 2000), también los colocan en el centro. Se trata, precisamente, de la actividad semiótica que la IA generativa pone en riesgo. El caso costarricense no es, así, un localismo, sino un caso bien documentado de un fenómeno global, que ofrece un vocabulario curricular preciso para nombrar un problema transferible a cualquier sistema que comparta esos procesos.

El objetivo es fundamentar, desde una perspectiva teórico-metodológica, la incidencia de los modelos generativos sobre las respuestas matemáticas y derivar una tipología de prompts orientada a preservar la actividad semiótica. No se trata de un estudio empírico, sino de una construcción conceptual que se ofrece como marco inicial y reconoce la necesidad de validación posterior. El manuscrito se organiza de la siguiente manera. Tras esta introducción, la sección 2 reúne los fundamentos teóricos, el análisis conceptual de la incidencia de la IA en las respuestas del estudiantado y del profesorado, y el caso costarricense. La sección 3 expone el abordaje metodológico. La sección 4 presenta y discute la tipología de prompts, la experiencia de aula que la ilustra, sus implicaciones y sus limitaciones. Finalmente, la sección 5 ofrece las conclusiones.

2. ELEMENTOS TEÓRICOS

La fundamentación se organiza en cuatro ejes teóricos que, leídos en conjunto, delimitan el problema que el trabajo aborda. En cada eje se construye un argumento antes que un catálogo de autores. Posteriormente, estos fundamentos se aplican al análisis conceptual de la incidencia de la IA en las respuestas del estudiantado y del profesorado y se sitúan en el caso costarricense.

2.1 El lenguaje matemático: registros, registro discursivo y simbolización

A diferencia de los objetos de otras ciencias, los objetos matemáticos no son directamente accesibles a la percepción, pues solo se accede a ellos a través de signos, es decir, de representaciones semióticas (Duval, 2006). En matemática, operar sobre el objeto y operar sobre su representación son inseparables, de modo que la actividad semiótica no es periférica al pensamiento, sino su sustancia.

Duval (2006) organiza esa actividad en torno a registros de representación como el lenguaje natural, la escritura simbólica y las representaciones gráficas, numéricas y tabulares. Asimismo, distingue los tratamientos, que permanecen dentro de un registro, de las conversiones, que implican el tránsito entre registros. La comprensión se evidencia, sobre todo, en la capacidad de convertir una representación en otra sin perder el objeto, y es en este proceso de conversión entre registros donde se localiza la mayoría de las dificultades de aprendizaje.

Esta mirada cognitiva se completa con su dimensión lingüística. Pimm (1987) mostró que la matemática escolar puede analizarse como un lenguaje, con un registro,

en sentido hallidayano, de vocabulario y sintaxis específicos, realizado en modalidades habladas y escritas que no se corresponden punto por punto. En su obra sobre los símbolos, Pimm (1995) anticipó, ante las primeras máquinas con capacidades matemáticas, la tensión entre fluidez y comprensión que hoy reaparece. En el ámbito hispanohablante, Alcalá (2002) desarrolló la perspectiva de la matemática como lenguaje, subrayando los procesos de simbolización y la apropiación progresiva de los códigos notacionales. Reunidas, estas miradas convergen: el lenguaje matemático es el medio constitutivo del pensamiento, con una cara representacional (Duval, 2006), una discursiva (Pimm, 1987) y una génesis escolar entendida como simbolización progresiva (Alcalá, 2002). De ahí se desprende una definición operativa según la cual una respuesta matemática no es un dato, sino una producción semiótica.

2.2 Fluidez sintética, génesis instrumental y comprensión

El efecto más estudiado de los modelos generativos sobre el trabajo estudiantil ha sido denominado fluidez sintética, entendida como la capacidad de producir resultados correctos y soluciones pulidas sin recorrer los procesos de razonamiento que construyen la comprensión (Wang et al., 2025). El fenómeno es congruente con hallazgos revisados por pares sobre la IA generativa, que tiende a homogeneizar el contenido y a reducir su diversidad colectiva (Doshi y Hauser, 2024) y que plantea riesgos de dependencia cognitiva, junto a oportunidades que la misma literatura reconoce (Kasneci et al., 2023). Este paralelismo entre fluidez y comprensión no es del todo nuevo (Pimm, 1995), pero la fluidez sintética lo lleva a un extremo en el que el producto puede desligarse por completo del proceso.

La noción de génesis instrumental ayuda a explicar por qué resulta relevante la forma en que la inteligencia artificial se integra en la actividad cognitiva del estudiante (Vérillon y Rabardel, 1995). Un artefacto no constituye por sí mismo un instrumento de aprendizaje, sino que llega a serlo cuando el aprendiz lo incorpora a su actividad cognitiva. En este proceso coexisten la instrumentación, mediante la cual el artefacto moldea el pensamiento del usuario, y la instrumentalización, por medio de la cual el usuario orienta el artefacto hacia sus propios fines.

El uso formativo de la inteligencia artificial exige que predomine la instrumentalización. Sin embargo, la evidencia preliminar describe un predominio de la instrumentación, en el que el modelo opera como una prótesis que ejecuta el trabajo en lugar del estudiante (Wang et al., 2025). En la misma línea, revisiones recientes advierten sobre el riesgo de sobredependencia y de externalización de habilidades cognitivas y metacognitivas que deberían permanecer en el estudiante (Qian, 2025).

En paralelo, la literatura internacional ha examinado la interacción entre seres humanos e inteligencia artificial y el aprendizaje mediado por esta tecnología. Estos fenómenos se han abordado mediante revisiones sistemáticas y estudios panorámicos (Holmes y Tuomi, 2022; Luckin et al., 2016; Zawacki-Richter et al., 2019), así como desde posiciones críticas (Selwyn, 2019). Asimismo, se han desarrollado bancos de evaluación del razonamiento matemático de los propios modelos, como GSM8K y MATH. Estos recursos documentan avances notables, pero también la persistencia de errores en la resolución de problemas de varios pasos (Cobbe et al., 2021; Hendrycks et al., 2021).

La distinción reactiva la de Skemp (1976) entre comprensión instrumental, la de seguir la regla sin razones, y comprensión relacional, el saber qué hacer y por qué. La fluidez sintética constituye un caso límite, pues el estudiante puede producir el resultado sin poseer siquiera la regla.

2.3 Contrato didáctico y la distinción entre ejercicio y problema

Brousseau (1997) describió el contrato didáctico como el conjunto de expectativas implícitas que regulan la relación entre docente, estudiante y saber. Asignar un problema supone, tácitamente, que resolverlo exige comprometerse con la matemática subyacente. Los modelos generativos pueden romper ese contrato al satisfacer la condición visible de la tarea, como una respuesta correcta, sin su finalidad educativa, en una forma sistematizada del efecto Topaze.

La ruptura del contrato no siempre resulta visible, pues la condición explícita de la tarea queda satisfecha. La evidencia experimental reciente permite observar sus huellas. Cuando estudiantes de matemática practicaron con un modelo generativo sin salvaguardas, su desempeño mejoró durante la práctica, pero cayó por debajo del grupo de control una vez retirada la herramienta (Bastani et al., 2025); en términos del contrato didáctico, la tarea se cumplió sin que ocurriera el compromiso con el saber que la justificaba. El fenómeno se ha descrito también como una forma de externalización o descarga cognitiva (cognitive offloading), en la que el aprendiz delega en un recurso externo procesos que la tarea pretendía ejercitar (Risko y Gilbert, 2016).

La distinción entre ejercicio y problema, central en el enfoque costarricense (Ruiz, 2013), resulta aquí decisiva. Un ejercicio aplica un procedimiento ya conocido; un problema obliga a construirlo y, por ello, moviliza los procesos de representar y comunicar. La evidencia disponible sugiere que, en ausencia de instrucciones que exijan lo contrario, las salidas de los modelos tienden a adoptar la forma del ejercicio, es decir, de un procedimiento indicado de antemano que el estudiante solo debe seguir (Bastani et al., 2025; Jia et al., 2024). Se trata de una tendencia observada y no de una propiedad inevitable; precisamente por ello, la tipología del apartado 4.1 sitúa en el prompt la decisión que puede revertirla.

2.4 Homogeneización del registro y valor epistémico

La literatura sobre escritura asistida por IA documenta un fenómeno de homogeneización según el cual los textos producidos con estos modelos tienden a converger hacia los mismos patrones léxicos y retóricos, que reflejan las regularidades del entrenamiento más que la voz de cada autor (Doshi y Hauser, 2024). En matemática, la voz del estudiante no es un asunto de estilo, sino el conjunto de sus elecciones representacionales y de sus conversiones.

La homogeneización no opera solo sobre el producto final, sino también sobre el repertorio de elecciones que lo precede. Cuando la conversión entre registros se delega en el modelo, lo que se descarga no es un cálculo auxiliar, sino el proceso en el que Duval (2006) sitúa la comprensión. Por ello, la descarga cognitiva puede resultar razonable para tareas rutinarias, pero problemática cuando recae sobre la operación que constituye el aprendizaje (Risko y Gilbert, 2016). De ahí que diversas propuestas de alfabetización en IA insistan en que el usuario debe conservar la capacidad de evaluar y contrastar las salidas del modelo, y no solo la de obtenerlas (Ng et al., 2021).

Esta convergencia tiene un costo epistémico. En los términos de Artigue (2002), el valor pragmático de la herramienta, es decir, su eficacia para resolver la tarea, puede ser alto mientras su valor epistémico (su capacidad de sostener la construcción de significado) queda vacío si la instrumentalización no ocurre. Una respuesta correcta puede, paradójicamente, tener valor epistémico nulo.

2.5 Incidencia de la IA en las respuestas del estudiantado

El marco anterior permite examinar cómo inciden los modelos generativos sobre las respuestas del estudiantado.

La respuesta mediada por IA puede diluir la voz del estudiante, es decir, el modo particular en que pasa del enunciado a la figura, las hesitaciones y las conversiones fallidas que corrige. Considérese la consigna “determine y clasifique los extremos de $f(x) = x^3 - 3x$ ”. Una respuesta genuina de aprendiz suele ser desigual, por ejemplo, tantea valores, esboza la parábola de $f(x) = 3x^2 - 3$, se equivoca con el signo de la segunda derivada y se corrige; y esa producción imperfecta es densa en información diagnóstica. La respuesta generada por un modelo llega ya en forma canónica; plantea $f'(x) = 3x^2 - 3 = 0$, obtiene $x = \pm 1$, evalúa $f''(x) = 6x$ y concluye un máximo local en $x = -1$ y un mínimo local en $x = 1$, con prosa pulcra y genérica. El resultado es correcto y el registro intachable; pero la fricción semiótica, es decir, los procesos de representar y comunicar en acto, desaparece, y con ella el rastro del pensamiento.

El riesgo es doble. Por una parte, la conversión, motor cognitivo para Duval (2006), puede dejar de ejercitarse. Por otra, las respuestas pierden la heterogeneidad que les daba valor diagnóstico, como los errores productivos y las representaciones propias. En ese escenario, el aula se desplazaría hacia un registro único, el del modelo, justamente el desenlace que la reforma costarricense quiso evitar al rechazar las meras operaciones mecánicas (MEP, 2012). Conviene subrayar el estatus de esta afirmación, que es una hipótesis teórica fundada en el fenómeno documentado de homogeneización (Doshi y Hauser, 2024) y no una constatación empírica sobre aulas concretas.

A escala agregada, la evidencia preliminar muestra que la correlación entre el desempeño en tareas no supervisadas, donde la IA está disponible, y el de exámenes presenciales tiende a debilitarse a medida que el contenido se vuelve más conceptual, señal de que el rendimiento no supervisado puede reflejar fluidez sintética antes que comprensión interiorizada (Wang et al., 2025).

2.6 Incidencia de la IA en las respuestas del profesorado

Los modelos inciden también sobre el docente, en dos facetas que conviene distinguir. En la faceta evaluativa, las señales tradicionales de comprensión se debilitan. El docente, que antes podía interpretar la matemática de un trabajo mediante el seguimiento de las conversiones y la localización del punto en que se quebraba el razonamiento, se ve ahora empujado a determinar la autenticidad del texto. La detección automática de texto generado por IA es, sin embargo, poco fiable. Una evaluación de catorce herramientas concluyó que no son ni exactas ni confiables, que tienden a clasificar el texto como humano y que su desempeño empeora ante la paráfrasis (Weber-Wulff et al., 2023). Apoyar la evaluación en estos instrumentos resulta, por tanto, frágil y potencialmente injusto; de ahí que distintos sistemas

universitarios opten por enfoques formativos antes que prohibicionistas (UCR, 2025). La salida que se perfila no es un mejor detector, sino un rediseño de la evaluación hacia procesos observables como defensas orales, trabajo en clase, tareas que exijan justificar las conversiones (Walkington, 2025).

En la faceta productiva, las orientaciones institucionales alientan el uso de la inteligencia artificial para preparar clases y corregir ejercicios (MEP, 2026). Este uso ofrece beneficios reales, como una retroalimentación más abundante y descriptiva y herramientas capaces de afinar el discurso del aula (Kasneci et al., 2023; Walkington, 2025). Sin embargo, el riesgo de homogeneización también opera en este ámbito, y la calidad de la retroalimentación automática no está garantizada. Un estudio comparativo halló que las sugerencias generadas por IA para resolver problemas matemáticos resultaban menos precisas que las elaboradas por docentes (Jia et al., 2024).

Si los problemas, las explicaciones y la retroalimentación provienen del mismo modelo, el registro del aula puede converger hacia una única forma de expresión. En ese escenario, el docente corre el riesgo de dejar de ser una fuente distintiva de lenguaje matemático y convertirse, sin advertirlo, en transmisor del registro del modelo.

Este panorama debe equilibrarse con la evidencia de efectos positivos y diferenciados. Un metaanálisis de 35 estudios experimentales reportó un efecto positivo moderado de ChatGPT sobre los resultados de aprendizaje, con variaciones según la materia, la duración y la modalidad instruccional (Wu et al., 2026). Asimismo, un ensayo aleatorizado halló que un tutor de IA diseñado con andamiaje pedagógico explícito superó a una clase de aprendizaje activo en física universitaria (Kestin et al., 2025).

El estudio de mayor interés para este trabajo es el de Bastani et al. (2025), realizado en matemática escolar. En este estudio, el acceso a un modelo sin salvaguardas mejoró el desempeño durante la práctica, pero lo redujo una vez retirada la herramienta. En cambio, una versión cuyo comportamiento fue configurado mediante prompts orientados a preservar el razonamiento mitigó ese deterioro.

La lectura conjunta de estos resultados no contradice la tesis de este artículo, sino que la precisa. El efecto de la IA sobre el aprendizaje no constituye un atributo inherente de la tecnología, sino que depende del diseño de la interacción, el cual se define, en buena medida, mediante el prompt. Existen incluso indicios de que el andamiaje mediado por IA puede favorecer procesos considerados vulnerables en este trabajo, como la resolución de problemas matemáticos (Zhou et al., 2025). Esta evidencia refuerza la necesidad de criterios de diseño como los propuestos en el apartado 4.1, pero también delimita sus condiciones de éxito. La tipología presupone la intervención de un docente que selecciona, revisa y depura las salidas del modelo, así como un uso deliberado de la herramienta; por ello, cabe esperar que resulte menos efectiva en contextos de uso masivo y no supervisado.

2.7 El caso costarricense: política, currículo e instituciones

A manera de caso, conviene leer en conjunto los tres niveles en que Costa Rica ha abordado la IA en educación, pues ilustran tanto una fortaleza como un punto ciego de interés general.



En el nivel nacional, la Estrategia Nacional de Inteligencia Artificial 2024–2027 fija principios de uso ético, supervisión humana, transparencia y equidad, e incluye la integración de habilidades de IA en todos los niveles educativos (MICITT, 2024). En el nivel escolar, el Ministerio de Educación Pública ha traducido esa estrategia en orientaciones para el aula (MEP, 2026), sobre la base de un currículo que, desde 2012, ya había situado representar y comunicar entre sus procesos centrales (MEP, 2012; Ruiz, 2013). En el nivel universitario, instituciones como la Universidad de Costa Rica han emitido lineamientos y marcos de gobernanza (UCR, 2025), con una coordinación interuniversitaria que ha promovido espacios comunes de reflexión y formación docente.

Leída en conjunto, esta arquitectura de tres niveles resulta valiosa; sin embargo, presenta un punto ciego que interesa a este trabajo. El énfasis dominante de los instrumentos recae sobre la ética, la gobernanza, la protección de datos y la integridad académica, mientras que la dimensión semiótica, es decir, qué ocurre con los procesos de representar y comunicar cuando la respuesta se produce a través de un modelo, permanece subtematizada. La lección es transferible, pues una política de IA en educación matemática que se limite a regular el uso, sin preguntarse por su efecto sobre los procesos que dan sentido al aprendizaje, atiende la superficie del problema y deja intacto su núcleo.

3. ABORDAJE METODOLÓGICO

Este artículo se construye como una propuesta teórico-metodológica fundamentada y no como un estudio empírico. Su elaboración combina tres procedimientos. El primero es una revisión narrativa y crítica de la literatura sobre IA generativa en educación matemática, lenguaje y registros de representación, génesis instrumental y evaluación. La revisión tuvo un carácter selectivo y argumentativo, no sistemático; no se aplicó un protocolo de elegibilidad ni se aspiró a la exhaustividad, de modo que no pretende la reproducibilidad de una revisión sistemática. Las fuentes se localizaron mediante búsquedas realizadas durante 2026 en Google Scholar y SciELO, con una cobertura que privilegió el período 2019–2026 para la literatura sobre IA en educación y que se remonta a las obras fundacionales para los fundamentos semióticos y didácticos. Se priorizaron trabajos revisados por pares y documentos institucionales oficiales; el único preprint incluido se declara como tal y se acompaña de fuentes consolidadas. La selección respondió a la pertinencia conceptual respecto de las dimensiones del lenguaje matemático, y la síntesis se orientó a construir argumentos e identificar la dimensión menos atendida del problema.

El segundo procedimiento es una lectura documental de los marcos que regulan la IA en el sistema educativo costarricense, como la Estrategia Nacional de Inteligencia Artificial (MICITT, 2024), las orientaciones del Ministerio de Educación Pública (MEP, 2026) y los lineamientos universitarios (UCR, 2025), con el fin de situar el problema en un caso concreto.

El tercer procedimiento es la derivación conceptual de la tipología, que no se fijó de antemano, sino que se obtuvo en tres fases. En la primera se identificaron, a partir del marco teórico, las dimensiones del lenguaje matemático vulnerables ante la mediación de un modelo generativo. En la segunda fase, cada dimensión se asoció con

el riesgo específico introducido por la inteligencia artificial y con la decisión de diseño del prompt orientada a mitigarlo. Así, los registros y las conversiones (Duval, 2006) se vincularon con el tránsito entre registros; el sistema escrito de la matemática (Pimm, 1987, 1995), con el uso riguroso de la notación y su adecuada representación; la fluidez sintética (Wang et al., 2025), con la verificación explícita; la diversidad del lenguaje matemático (Alcalá, 2002), con la adecuación del registro al receptor; y el contrato didáctico (Brousseau, 1997), con la distinción entre problema y ejercicio. En la tercera fase, las categorías se contrastaron de manera iterativa con ejemplos redactados para distintas áreas del currículo (Anexo A), lo que permitió depurar sus fronteras y comprobar que cada riesgo identificado quedara cubierto. La Tabla 1 hace explícita la correspondencia resultante. El número de categorías, seis, responde a las dimensiones que el marco identifica como vulnerables y no a una cifra fijada a priori.

Conviene precisar que la tipología se ofrece como un marco inicial y no como un dispositivo empíricamente validado. Su derivación no incluyó la codificación de datos empíricos ni la consulta previa a personas expertas. Asimismo, la experiencia docente presentada en el apartado 4.2 cumple una función estrictamente ilustrativa, pues muestra la plausibilidad de la propuesta, pero no aporta evidencia sistemática. El carácter preliminar y aún no validado empíricamente de la tipología delimita el alcance de la propuesta y define una agenda de validación en tres vías complementarias.

La primera consiste en estudios de aula que documenten el efecto de cada categoría sobre el razonamiento, la representación y la comunicación matemática del estudiantado. La segunda, en análisis de fiabilidad interevaluador que examinen si docentes distintos clasifican los mismos prompts en las mismas categorías. La tercera, en una consulta a especialistas mediante la técnica Delphi que valore la suficiencia del conjunto y la delimitación entre categorías, así como su contraste con otras taxonomías de diseño de prompts.

4. RESULTADOS/DISCUSIÓN

4.1 Una tipología de prompts para preservar la actividad semiótica

Si el riesgo se sitúa en el modo en que se producen las respuestas, la respuesta pedagógica debe situarse en el mismo lugar: el prompt, entendido como la instrucción con la que el docente o el estudiante orienta al modelo. El prompt es el punto en el que se decide si el recurso resultante ejercita o suprime la actividad semiótica, y donde se juega la diferencia entre instrumentación e instrumentalización. Como soporte de redacción, esta sección adopta el modelo de las “Seis R” (rol, resultado, receptor, restricciones, recursos y revisión), propuesto por Ramírez Bogantes et al. (2026) para el diseño de recursos digitales de matemática en secundaria; la sexta R, la revisión, y la curaduría docente posterior operan como el control humano que el marco considera innegociable. A partir de ese soporte se proponen seis categorías de prompts en las que el manejo del lenguaje matemático es la decisión central; los ejemplos completos se ofrecen en el Anexo A.

La primera categoría atiende la notación rigurosa y desambiguada. Responde al criterio de rigor disciplinar y a la advertencia sobre la especificidad del sistema escrito de la matemática (Pimm, 1995), pues los modelos producen con frecuencia una

notación que aparenta corrección, pero la incumple parcialmente; por ejemplo, una base de logaritmo omitida, un nombre de variable que cambia entre ítems o una propiedad falsa presentada como válida. Un prompt de esta categoría fija la notación de antemano y nombra las confusiones que se busca evitar. El principal riesgo de curaduría es que el vocabulario experto aparezca usado con imprecisión, simulando un rigor ausente.

La segunda categoría persigue el tránsito entre registros (Duval, 2006). Busca que el recurso articule simultáneamente al menos dos registros, como el algebraico, el gráfico, el tabular o el verbal, y que haga visible la conversión entre ellos, proceso en el que se evidencia la comprensión. Un explorador que solo muestra una representación incumple el criterio; en cambio, uno que actualiza varios registros de forma sincronizada lo encarna. El riesgo más frecuente es la desincronización entre registros o la incoherencia entre el registro verbal y el gráfico.

La tercera categoría regula el renderizado de la notación. La matemática en texto plano resulta inaccesible, pues el sistema escrito matemático no se reduce a una cadena de caracteres (Pimm, 1987, 1995). Un prompt de esta categoría exige que la simbología se exprese en un lenguaje de marcado matemático y se renderice correctamente, y anticipa los puntos de falla técnica. La fidelidad notacional es, también ella, una condición del lenguaje matemático.

La cuarta categoría introduce la verificación explícita como salvaguarda frente a la fluidez sintética. Dado que los modelos pueden adoptar las marcas formales del rigor sin cumplirlo, por ejemplo, un procedimiento ordenado cuyo último paso no se deduce de los anteriores, y que sus respuestas no siempre son precisas (Jia et al., 2024), un prompt de esta categoría obliga a verificar cada resultado y advierte dónde es probable que el procedimiento falle. La curaduría debe vigilar la coherencia lógica entre pasos que el registro experto fluido tiende a encubrir.

La quinta categoría calibra el registro y la densidad simbólica según el receptor, en la línea de la matemática como apropiación progresiva del lenguaje y de la atención a la diversidad del aula (Alcalá, 2002). El mismo contenido admite distintas dosis de formalización. Para algunas adecuaciones conviene anclar primero la lengua natural y reducir la densidad de símbolos; para el trabajo con altas capacidades, conviene aumentar la formalización y exigir generalización. El riesgo es doble, pues consiste en introducir símbolos sin anclaje verbal o en confundir “mayor dificultad” con “más formalización”.

La sexta categoría procura que el recurso plantee un problema y no un ejercicio. Apoyada en la distinción que vertebra el enfoque costarricense (Ruiz, 2013) y en la noción de contrato didáctico (Brousseau, 1997), busca tareas no rutinarias que movilicen los procesos de representar y comunicar, de modo que el estudiante decida cómo organizar los datos, represente en más de un registro y justifique por escrito su conclusión. El riesgo característico es que el modelo degrade el problema a ejercicio, indicando de antemano el procedimiento que debía construirse.

Las Tablas 1 y 2 presentan, respectivamente, la correspondencia entre las construcciones teóricas y las seis categorías de prompts, así como una síntesis de estas categorías y sus ejemplos. Esta relación no es biunívoca, pues una misma construcción

teórica puede atenderse mediante varias categorías y una categoría puede responder a más de un riesgo. La tipología puede situarse, además, frente a otras taxonomías de diseño de prompts. Los marcos de uso general, como el modelo CLEAR, orientado a la alfabetización informacional mediante prompts concisos, lógicos, explícitos, adaptativos y reflexivos (Lo, 2023), o el propio soporte de las Seis R (Ramírez Bogantes et al., 2026), organizan la forma de la instrucción y son transversales a las disciplinas, pero no se pronuncian sobre el contenido matemático de la interacción. La tipología aquí propuesta se distingue porque su criterio organizador no es la estructura del prompt, sino la dimensión del lenguaje matemático que cada categoría protege, a saber, la notación, la conversión entre registros, el renderizado, la verificación, el registro según el receptor y el estatus de problema. En ese sentido, no compete con los marcos generales, sino que los especializa para la educación matemática.

Tabla 1. Correspondencia entre construcciones teóricas y categorías de prompts

Construcción teórica	Riesgo asociado	Categoría de prompt
Registros de representación y conversión (Duval, 2006)	Recurso de un solo registro; sin conversión	C2 · Tránsito entre registros
Representar y comunicar como procesos curriculares (MEP, 2012; Ruiz, 2013)	La respuesta mediada por IA suprime la actividad semiótica	C6 · Problema, no ejercicio
Fluidez sintética (Wang et al., 2025; Pimm, 1995)	Producto pulido sin razonamiento	C4 · Verificación explícita
Contrato didáctico; ejercicio frente a problema (Brousseau, 1997; Ruiz, 2013)	La IA degrada el problema a ejercicio	C6 · Problema, no ejercicio
Rigor disciplinar y sistema escrito (Pimm, 1987, 1995)	Notación que aparenta corrección, pero la incumple	C1 y C3 · Notación y renderizado
Matemática como lenguaje y diversidad (Alcalá, 2002)	Registro único; densidad simbólica inadecuada	C5 · Registro calibrado al receptor
Instrumentalización frente a instrumentación (Vérillon y Rabardel, 1995; Artigue, 2002)	El modelo opera como prótesis, no como andamiaje	Las seis categorías + curaduría docente

Fuente: Elaboración propia.

Tabla 2. Síntesis de las seis categorías de prompts y sus ejemplos

Categoría	Eje de lenguaje matemático	Ejemplo (área del Programa MEP)
C1	Notación rigurosa y desambiguada	Logaritmos (Relaciones y álgebra)
C2	Tránsito entre registros	Sistemas 2×2 (Relaciones y álgebra)
C3	Renderizado de la notación	Radicales y exponentes (Números)
C4	Verificación explícita	Ecuaciones con radicales (Relaciones y álgebra)
C5	Registro calibrado al receptor	Perímetro y área (Geometría)
C6	Problema, no ejercicio	Frecuencia relativa (Estadística y probabilidad)

Fuente: Elaboración propia.

4.2 Una experiencia de aula que ilustra la tipología

Esta sección no presenta resultados empíricos ni una validación de la tipología, sino una experiencia de aula que permite ilustrarla mediante una lectura retrospectiva. La tipología se derivó previamente en el plano conceptual; corresponde ahora mostrar cómo algunas de sus categorías se manifestaron en una situación educativa concreta.

Para ello, se recurre a una experiencia desarrollada durante el primer semestre de 2026 en el curso Ecuaciones Diferenciales (EM4610) de la carrera de Enseñanza de la Matemática con Entornos Tecnológicos del Instituto Tecnológico de Costa Rica. Participó un grupo de once estudiantes y se incorporó un componente de uso de IA generativa organizado en cinco actividades breves, realizadas en clase, una por sesión y en parejas o tríos.

Cada actividad se vinculó con un tema del programa y con una herramienta distinta: un resumen en audio con NotebookLM sobre ecuaciones separables y homogéneas; un artefacto interactivo con Claude sobre técnicas de primer orden; un material de estudio redactado conjuntamente con un editor de documentos asistido por IA sobre ecuaciones lineales homogéneas de orden superior; una resolución comparativa entre varios modelos sobre transformada de Laplace; y un informe de síntesis elaborado con una IA de búsqueda sobre las funciones de Heaviside y la delta de Dirac.

El diseño se apoyó en los modelos TPACK (Mishra y Koehler, 2006) y SAMR (Puentedura, 2006). La rúbrica de evaluación no premiaba la mera obtención de la respuesta, sino el criterio mostrado por el estudiante al utilizar la herramienta, como la detección de errores, la reflexión pedagógica y la incorporación de aportes propios. La copia de la salida del modelo recibía una calificación nula (Acuña Chacón, 2026).

Este criterio coincide con el principio rector de la tipología: determinar si el uso de la IA ejercita o suprime la actividad semiótica. Por ello, la experiencia constituye un terreno propicio para ilustrar las categorías. Los episodios que se presentan a continuación provienen de las cinco actividades; se han omitido los nombres del estudiantado y los grupos se identifican de forma genérica.

Actividad 1: resumen en audio

La primera actividad permitió observar tensiones relacionadas con el registro calibrado al receptor (C5). El resumen en audio explicaba los conceptos con fluidez, pero no formulaba ninguna ecuación concreta y omitía las condiciones de existencia y unicidad, así como los dominios de validez.

Ante esta limitación, un grupo propuso una tarea en la que los estudiantes debían contrastar el lenguaje divulgativo del audio con el rigor propio del discurso matemático. Esta propuesta coincide con el propósito de C5, pues exige adecuar el registro al receptor sin renunciar a la precisión conceptual.

Actividad 2: artefacto interactivo

La verificación explícita (C4) y la notación rigurosa (C1) se hicieron visibles en torno a un clasificador de ecuaciones de primer orden generado con Claude. Al examinarlo,

el grupo descubrió que el artefacto no operaba directamente sobre el objeto matemático, sino sobre patrones de texto.

Una ecuación que era simultáneamente separable y homogénea se rotulaba únicamente como separable «por su forma». Además, categorías que no se excluyen entre sí, como lineal y exacta, se trataban como incompatibles. Al repetir la clasificación, los pasos e incluso la categoría asignada cambiaban, y en algunos desarrollos desaparecían términos.

El episodio ilustra la tesis central de este trabajo: una salida fluida puede aparentar corrección mientras incumple el sistema notacional y conceptual que afirma manejar. La corrección no provino de un prompt aislado más afortunado, sino de la curaduría docente, que permitió revisar y reorientar la interacción.

El renderizado de la notación (C3) también apareció en esta actividad. Una pareja exigió en su prompt que toda la matemática se escribiera en LaTeX, pero que se mostrara compuesta mediante una biblioteca de renderizado, de modo que el usuario «nunca viera el código crudo». Esta formulación constituye una manifestación espontánea de C3.

Actividad 3: material de estudio co-redactado

En la tercera actividad se manifestó el problema inverso relacionado con el renderizado de la notación. Al exportar a PDF el material generado en el editor de documentos, las expresiones matemáticas aparecían como código sin compilar y las definiciones de las variables quedaban separadas del texto. Solo después de utilizar otra vía de exportación se recuperó una notación legible.

El episodio confirmó que el sistema escrito de la matemática no se reduce a una cadena de caracteres (Pimm, 1987, 1995). La presentación gráfica de los símbolos, su relación con el texto y la definición explícita de las variables forman parte de la producción del significado matemático.

El documento también incluía espacios para que el estudiante respondiera preguntas y elaborara un ejemplo propio. Al exigir una producción que el aprendiz debía completar, revisar y corregir, la actividad recuperaba la voz que una respuesta cerrada del modelo tiende a borrar. De este modo, se aproximaba a la categoría de problema, no ejercicio (C6).

Actividad 4: resolución comparativa

En la cuarta actividad, los estudiantes resolvieron un mismo problema de transformada de Laplace con dos modelos y mediante un procedimiento manual. Posteriormente, construyeron una tabla en la que compararon la notación, los pasos y los errores de cada solución.

Los estudiantes observaron que uno de los modelos resolvía «de manera más algorítmica, como si ya se entendiera el tema». La ausencia de un desarrollo suficientemente explícito dificultaba comprobar la existencia de errores. Esta observación expresa, en palabras del propio estudiantado, la fluidez sintética y el riesgo

de homogeneización del registro descritos en la literatura (Wang et al., 2025; Doshi y Hauser, 2024).

El grupo concluyó que la IA podía ayudar a avanzar en la resolución, pero que no convenía confiar en sus respuestas sin contrastarlas con el libro o con los materiales proporcionados por el docente. Esta conclusión refleja la verificación explícita (C4) y la curaduría que el marco considera innegociables.

Actividad 5: informe de síntesis

La quinta actividad consistió en elaborar, mediante una IA de búsqueda, un informe de síntesis sobre aplicaciones de las funciones de Heaviside y la delta de Dirac. En este caso, la verificación explícita (C4) se extendió desde el procedimiento matemático hacia la revisión de las fuentes.

La consigna exigía acceder al menos a dos de las fuentes citadas por la herramienta y confirmar que la información fuera fidedigna. Un grupo comprobó que sus cuatro fuentes respaldaban las afirmaciones presentadas. Otro encontró que una referencia atribuida a un texto clásico de ecuaciones diferenciales no sustentaba directamente el modelo específico que pretendía justificar. La obra presentaba el concepto general y un circuito diferente, pero no la aplicación citada.

Este hallazgo prolonga la tesis central hacia el registro de la citación. Una salida puede aparentar respaldo bibliográfico con la misma fluidez con la que aparenta corrección notacional.

En esta actividad también reapareció la tensión entre el registro divulgativo y el rigor matemático (C5). Un grupo observó que la síntesis mencionaba la utilidad del método de Heaviside, pero no formulaba ni resolvía explícitamente la ecuación asociada. Asimismo, se manifestó la necesidad de curar el formato (C3), pues la herramienta de síntesis visual incumplió las especificaciones de diseño y el grupo debió migrar a otra aplicación para obtener el documento solicitado. De los cuatro grupos participantes en esta sesión de cierre, tres entregaron el informe.

La Tabla 3 resume la correspondencia observada entre las actividades y las categorías ilustradas. Conviene, sin embargo, precisar su alcance. La experiencia no constituye una validación empírica de la tipología. Las actividades se diseñaron a partir de TPACK y SAMR, y no con el propósito de poner a prueba las seis categorías. Por tanto, la lectura realizada es ilustrativa y retrospectiva.

Además, el número de estudiantes es reducido, la experiencia corresponde a un único grupo y la categoría de tránsito entre registros (C2) quedó menos representada que las restantes. El aporte de la experiencia consiste en ofrecer evidencia de plausibilidad: las categorías permiten nombrar decisiones y dificultades que emergen de manera espontánea en el trabajo real de futuros docentes. Asimismo, proporciona un punto de partida concreto para el proceso de validación que se plantea más adelante.

Tabla 3. Correspondencia entre las actividades del rubro y las categorías ilustradas

Actividad (herramienta)	Tema del programa	Categorías ilustradas
1 · Resumen en audio (NotebookLM)	ED separables y homogéneas	C5; tensión fluidez-rigor
2 · Artefacto interactivo (Claude)	Técnicas de primer orden	C1, C3, C4; instrumentalización
3 · Material co-redactado (editor con IA)	ED lineales de orden superior	C3, C6; voz del estudiante
4 · Resolución comparativa (varios modelos)	Transformada de Laplace	C1, C3, C4; curaduría docente
5 · Informe de síntesis (IA de búsqueda)	Heaviside y Delta de Dirac	C3, C4, C5; verificación de fuentes

Fuente: Elaboración propia.

4.3 Implicaciones para la práctica y la formación docente

De lo anterior se desprende una orientación que no se reduce a la falsa disyuntiva entre prohibir y permitir. Las guías institucionales ganarían en profundidad si, además de regular el uso ético, formularan cómo preservar los procesos de representar y comunicar. El criterio rector no debería ser solo si el uso es honesto y seguro, sino si ejercita o suprime la actividad semiótica del estudiante. En el plano didáctico, conviene privilegiar la instrumentalización sobre la instrumentación, es decir, enseñar a interrogar al modelo, a leer críticamente sus salidas y a contrastarlas con definiciones formales, y rediseñar la evaluación hacia procesos observables, especialmente en los cursos de servicio masivos. Finalmente, la retroalimentación apoyada en IA debe permanecer específica y anclada en el trabajo de cada estudiante, no delegarse por completo.

4.4 Limitaciones

Este trabajo presenta limitaciones que conviene explicitar. La primera es su naturaleza teórico-metodológica, pues la tipología es una propuesta conceptual aún no validada en el aula, por lo que sus efectos sobre el razonamiento, la representación y la comunicación deben verificarse empíricamente. La experiencia ilustrativa del apartado 4.2, basada en un único grupo reducido y de lectura retrospectiva, no supe esa validación. La segunda es que parte de la evidencia sobre la fluidez sintética proviene de literatura muy reciente y, en algún caso, preliminar (Wang et al., 2025), lo que aconseja prudencia interpretativa. La tercera es que el caso analizado es costarricense; aunque la lección se plantea como transferible, su traslado a otros contextos requiere atender sus particularidades. Estas limitaciones definen, a la vez, una agenda de investigación que comprende estudios de aula sobre el efecto de las categorías de prompts en la actividad semiótica, análisis de fiabilidad interevaluador, consultas a especialistas y análisis comparados de los marcos institucionales que regulan la IA en educación matemática.

5. CONCLUSIONES

Este trabajo se propuso fundamentar, desde una perspectiva teórico-metodológica, la incidencia de los modelos generativos sobre las respuestas matemáticas de estudiantes y docentes, y derivar una tipología de prompts orientada a

preservar la actividad semiótica. Ambos propósitos convergen en una misma constatación conceptual. El medio de la matemática es el lenguaje, de modo que aprender matemática es, en gran medida, aprender a moverse entre registros y a apropiarse de un sistema notacional. Los modelos generativos se han instalado como coautores de ese lenguaje a ambos lados del aula, y moldean tanto las respuestas del estudiantado como las del profesorado. El riesgo no proviene del error de la máquina, sino de su corrección uniforme, que amenaza con disolver la fricción semiótica de la que vive el aprendizaje.

Antes que prohibir o permitir, la tarea consiste en mantener viva y humana la actividad semiótica, y en traducir esa convicción en decisiones de diseño como la tipología aquí propuesta. Esa convicción encuentra apoyo en la evidencia reciente, pues los efectos de la IA sobre el aprendizaje matemático se muestran positivos o perjudiciales según el diseño de la interacción (Bastani et al., 2025; Kestin et al., 2025), lo que confirma que el prompt, y no la herramienta en abstracto, es el lugar pedagógicamente decisivo. Queda pendiente su validación en el aula y su articulación con los marcos institucionales, de modo que la política no solo regule el uso de la IA, sino que proteja las condiciones en que la actividad matemática conserva su valor epistémico. En último término, quien moldea el lenguaje matemático moldea la matemática que se aprende.

DECLARACIÓN ÉTICA Y DE CONSENTIMIENTO INFORMADO

La experiencia documentada se desarrolló en el marco de la actividad docente ordinaria del curso. El estudiantado otorgó su consentimiento informado para el uso, con fines exclusivamente académicos y de forma anonimizada, de sus producciones. No se reportan nombres, números de carné ni otros datos que permitan la identificación de las personas participantes.

DECLARACIÓN DE CONTRIBUCIONES DE LAS PERSONAS AUTORAS

RAC concibió la idea presentada, desarrolló el marco teórico-metodológico, diseñó e implementó el rubro de uso de IA en el curso, derivó la tipología de prompts, analizó los episodios de aula y redactó, revisó y aprobó la versión final del manuscrito.

DECLARACIÓN DE DISPONIBILIDAD DE DATOS

Los materiales que respaldan los hallazgos de este estudio, las consignas del rubro y las producciones anonimizadas del estudiantado están disponibles, previa solicitud razonable y con resguardo de la confidencialidad, por parte del autor correspondiente, RAC.

AGRADECIMIENTOS

El autor agradece a las personas estudiantes del curso EM4610 (Ecuaciones Diferenciales, I semestre 2026) de la carrera de Enseñanza de la Matemática con Entornos Tecnológicos del Instituto Tecnológico de Costa Rica, cuyas producciones hicieron posible la experiencia ilustrativa que aquí se documenta.

6. REFERENCIAS

- Acuña Chacón, R. (2026). *Uso de IA generativa (5 %): actividades de diseño de modelos interactivos y resolución de problemas con IA* [Programa del curso EM4610, Ecuaciones Diferenciales]. Escuela de Matemática, Instituto Tecnológico de Costa Rica.
- Alcalá, M. (2002). *La construcción del lenguaje matemático*. Graó.
- Artigue, M. (2002). Learning mathematics in a CAS environment: The genesis of a reflection about instrumentation and the dialectics between technical and conceptual work. *International Journal of Computers for Mathematical Learning*, 7(3), 245–274. <https://doi.org/10.1023/A:1022103903080>
- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö. y Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26), e2422633122. <https://doi.org/10.1073/pnas.2422633122>
- Brousseau, G. (1997). *Theory of didactical situations in mathematics: Didactique des mathématiques, 1970–1990*. Kluwer Academic Publishers.
- Cobbe, K., Kosaraju, V., Bavarian, M., Chen, M., Jun, H., Kaiser, L., Plappert, M., Tworek, J., Hilton, J., Nakano, R., Hesse, C. y Schulman, J. (2021). *Training verifiers to solve math word problems* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2110.14168>
- Doshi, A. R. y Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28), eadn5290. <https://doi.org/10.1126/sciadv.adn5290>
- Duval, R. (2006). Un tema crucial en la educación matemática: la habilidad para cambiar el registro de representación. *La Gaceta de la Real Sociedad Matemática Española*, 9(1), 143–168.
- Hendrycks, D., Burns, C., Kadavath, S., Arora, A., Basart, S., Tang, E., Song, D. y Steinhardt, J. (2021). Measuring mathematical problem solving with the MATH dataset. En J. Vanschoren y S. Yeung (Eds.), *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks* (Vol. 1). Neural Information Processing Systems Foundation. <https://datasets-benchmarks-proceedings.neurips.cc/paper/2021/hash/be83ab3ecd0db773eb2dc1b0a17836a1-Abstract-round2.html>
- Holmes, W. y Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4), 542–570. <https://doi.org/10.1111/ejed.12533>
- Jia, J., Wang, T., Zhang, Y. y Wang, G. (2024). The comparison of general tips for mathematical problem solving generated by generative AI with those generated by human teachers. *Asia Pacific Journal of Education*, 44(1), 8–27. <https://doi.org/10.1080/02188791.2023.2286920>

- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kestin, G., Miller, K., Klales, A., Milbourne, T. y Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports*, 15, 17458. <https://doi.org/10.1038/s41598-025-97652-6>
- Lo, L. S. (2023). The CLEAR path: A framework for enhancing information literacy through prompt engineering. *The Journal of Academic Librarianship*, 49(4), 102720. <https://doi.org/10.1016/j.acalib.2023.102720>
- Luckin, R., Holmes, W., Griffiths, M. y Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.
- Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones [MICITT]. (2024). *Estrategia Nacional de Inteligencia Artificial 2024–2027*. MICITT. https://www.micitt.go.cr/gobierno_digital/inteligencia_artificial
- Ministerio de Educación Pública. (2012). *Programas de estudio de Matemáticas: I y II ciclo de la Educación Primaria, III ciclo de Educación General Básica y Educación Diversificada*. <https://www.mep.go.cr/sites/default/files/media/matematica.pdf>
- Ministerio de Educación Pública. (2026). *Guía docente para el uso de la inteligencia artificial en el aula*. <https://www.mep.go.cr/sites/default/files/2026-02/GuiaDocenteIA.pdf>
- Mishra, P. y Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. NCTM.
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W. y Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Pimm, D. (1987). *Speaking mathematically: Communication in mathematics classrooms*. Routledge & Kegan Paul.
- Pimm, D. (1995). *Symbols and meanings in school mathematics*. Routledge.
- Puentedura, R. R. (2006). *Transformation, technology, and education*. Hippasus. <http://hippasus.com/resources/tte/>

- Qian, Y. (2025). Pedagogical applications of generative AI in higher education: A systematic review of the field. *TechTrends*, 69, 1–18. <https://doi.org/10.1007/s11528-025-01100-1>
- Ramírez Bogantes, M., Gutiérrez Montenegro, M., Acuña Chacón, R. y Meza Cascante, L. G. (2026). *De la idea al aula. Diseñar experiencias de aprendizaje con IA: prompts para crear recursos interactivos en secundaria*. Escuela de Matemática, Instituto Tecnológico de Costa Rica. <https://hdl.handle.net/2238/18937>
- Risko, E. F. y Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. <https://doi.org/10.1016/j.tics.2016.06.007>
- Ruiz, Á. (2013). La reforma de la educación matemática en Costa Rica. Perspectiva de la praxis. *Cuadernos de Investigación y Formación en Educación Matemática*, 8(número especial). <https://revistas.ucr.ac.cr/index.php/cifem>
- Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity.
- Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching*, 77, 20–26.
- Universidad de Costa Rica. (2025). *Lineamientos generales para la implementación de la inteligencia artificial en la Universidad de Costa Rica* (Resolución de Rectoría R-469-2025). UCR. <https://www.cu.ucr.ac.cr/gacetas/2025/a70-2025.pdf>
- Vérillon, P. y Rabardel, P. (1995). Cognition and artifacts: A contribution to the study of thought in relation to instrumented activity. *European Journal of Psychology of Education*, 10(1), 77–101. <https://doi.org/10.1007/BF03172796>
- Walkington, C. (2025). The implications of generative artificial intelligence for mathematics education. *School Science and Mathematics*. Publicación anticipada en línea. <https://doi.org/10.1111/ssm.18356>
- Wang, S., Xia, Q. y Ye, Q. (2025). *Synthetic fluency and epistemic offloading in undergraduate mathematics in the age of AI* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2512.21045>
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P. y Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, 19(1), 26. <https://doi.org/10.1007/s40979-023-00146-z>
- Wu, X., Zhu, P., Zhang, J., Yin, M. y Wang, Y. (2026). ChatGPT's impact on student learning outcomes: A meta-analysis of 35 experimental studies. *Humanities and Social Sciences Communications*, 13, 684. <https://doi.org/10.1057/s41599-026-07019-z>
- Zawacki-Richter, O., Marín, V. I., Bond, M. y Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education*, 16, 39. <https://doi.org/10.1186/s41239-019-0171-0>

Zhou, R., He, X., Fan, Q., Li, Y., Li, Y., Xiao, X. y Fang, J. (2025). Exploring ChatGPT-facilitated scaffolding in undergraduates' mathematical problem solving. *Journal of Computer Assisted Learning*, 41(4), e70077. <https://doi.org/10.1111/jcal.70077>

ANEXO A. EJEMPLOS DE PROMPTS BAJO EL MODELO DE LAS SEIS R

Los siguientes ejemplos ilustran las seis categorías del apartado 4.1. Son ejemplos originales; varían el tema, el área del Programa de Estudios del MEP y el contexto. La estructura sigue las Seis R: [ROL], [RESULTADO], [RECEPTOR], [RESTRICCIONES], [RECURSOS] y [REVISIÓN].

A1. Notación rigurosa y desambiguada (Relaciones y álgebra, décimo año)

[ROL] Diseñador de recursos de matemática para secundaria, con criterio de rigor en notación algebraica. [RESULTADO] Práctica guiada en HTML autocontenido sobre propiedades de los logaritmos y ecuaciones logarítmicas sencillas. [RECEPTOR] Décimo año, 15-16 años. Conocen la definición de logaritmo como exponente. [RESTRICCIONES] Notación obligatoria y consistente: base siempre explícita ($\log_2(x)$, nunca “ $\log x$ ”); distinguir $\log$ de $\ln$; misma incógnita (x) en todos los ítems. Recuadro de ERROR FRECUENTE con la igualdad falsa $\log(a+b)=\log a+\log b$ frente a la propiedad $\log_b(a \cdot c)=\log_b a+\log_b c$. [RECURSOS] Habilidad MEP (Relaciones y álgebra): aplicar propiedades de los logaritmos. Ejemplos verificados: $\log_2(8)=3$; $\log_3(x)=4 \rightarrow x=81$; $\log_5(x)+\log_5(2)=\log_5(10) \rightarrow x=5$. [REVISIÓN] Ninguna expresión “ $\log$ ” sin base; $\log$ y $\ln$ nunca como sinónimos; misma incógnita en todos los ítems; recuadro de error correcto.

A2. Tránsito entre registros (Relaciones y álgebra, décimo año)

[ROL] Diseñador pedagógico atento a la articulación de registros de representación. [RESULTADO] Explorador en HTML autocontenido de un sistema de dos ecuaciones lineales con dos incógnitas y solución única. [RECEPTOR] Décimo año. Conocen la ecuación de la recta y el plano cartesiano. [RESTRICCIONES] Al mover los coeficientes con deslizadores, se actualizan a la vez CUATRO registros: algebraico ($ax+by=c$), gráfico (las dos rectas y su intersección), tabular (tres pares por recta) y verbal (frase que describe la solución). Los casos sin solución o con infinitas se describen bien. [RECURSOS] Habilidad MEP: resolver sistemas y relacionarlos con la intersección gráfica. Caso inicial: $x+y=5$ y $x-y=1$, solución (3, 2). [REVISIÓN] Al mover un deslizador cambian los cuatro registros; la frase verbal coincide con el gráfico.

A3. Renderizado correcto de la notación (Números, noveno año)

[ROL] Programador educativo de matemática para secundaria. [RESULTADO] Práctica guiada en HTML autocontenido sobre radicales y exponentes racionales, con la notación renderizada. [RECEPTOR] Noveno año. Conocen potencias de exponente entero. [RESTRICCIONES] Toda expresión en LaTeX renderizado por KaTeX (CDN oficial), nunca como texto plano; delimitadores ... correctamente cerrados; el colón costarricense como símbolo de moneda fuera de modo matemático; si se inyectan ítems dinámicamente, reinvocar el renderizado de KaTeX tras insertarlos. [RECURSOS] Habilidad MEP (Números): convertir entre radicales y exponentes racionales. Ejemplos verificados: raíz cúbica de x al cuadrado = $x^{(2/3)}$; raíz de $50 = 5 \cdot \text{raíz de } 2$; $8^{(2/3)} = 4$. [REVISIÓN] Ningún delimitador abierto; raíces y exponentes compuestos; la notación se renderiza también en los ítems generados dinámicamente.

A4. Verificación explícita (Relaciones y álgebra, décimo año)

[ROL] Diseñador de matemática con criterio de verificación estricta de los procedimientos. [RESULTADO] Práctica guiada en HTML autocontenido sobre ecuaciones con radicales, con énfasis en soluciones extrañas. [RECEPTOR] Décimo año. Resuelven ecuaciones lineales y cuadráticas básicas. [RESTRICCIONES] Cada ejercicio muestra la verificación de cada candidato sustituyéndolo en la ecuación original y declara si es válida o extraña. Incluir al menos un ejercicio con solución extraña y explicar por qué surge. Recuadro con dos errores típicos: dividir entre una expresión que puede ser cero, y olvidar verificar las soluciones. [RECURSOS] Habilidad MEP: resolver ecuaciones con radicales y validar soluciones. Ejemplos verificados: raíz de $(x+3)=x-3 \rightarrow$ candidatas $x=1$ y $x=6$, solo $x=6$ válida; raíz de $(2x+1)=3 \rightarrow x=4$, válida. [REVISIÓN] Cada solución se sustituye en la ecuación original; la extraña se identifica como tal; cada paso se sigue del anterior.

A5. Registro calibrado al receptor (Geometría, octavo año)

[ROL] Diseñador de matemática sensible a la adecuación del registro al estudiantado. [RESULTADO] DOS versiones de una misma práctica guiada en HTML autocontenido sobre perímetro y área de figuras compuestas. [RECEPTOR] Octavo año. Versión A: estudiante con discalculia. Versión B: estudiante con altas capacidades. [RESTRICCIONES] Versión A (densidad simbólica reducida): cada paso primero en lenguaje natural y luego la fórmula; una operación por renglón; apoyo visual de la figura. Versión B (formalización ampliada): notación general ($A=(b \cdot h)/2$ con b y h parámetros), expresar el área como suma algebraica y generalizar con medidas literales. [RECURSOS] Habilidad MEP (Geometría): perímetro y área por descomposición. Figura verificada: rectángulo $6 \text{ cm} \times 4 \text{ cm}$ con triángulo de base 6 cm y altura 3 cm ; área total $24 + 9 = 33 \text{ cm}^2$. [REVISIÓN] En A, una operación por renglón; en B, generalización literal correcta; ambas con el mismo resultado numérico.

A6. Problema, no ejercicio (Estadística y probabilidad, sétimo año)

[ROL] Diseñador pedagógico con enfoque de resolución de problemas. [RESULTADO] Un problema (no un ejercicio) en HTML autocontenido sobre frecuencia relativa, que pida representar los datos de más de una forma y comunicar una conclusión por escrito. [RECEPTOR] Sétimo año. Comparan fracciones y porcentajes sencillos. [RESTRICCIONES] Debe ser un problema: el procedimiento no se indica de antemano. Dos registros: tabla de frecuencias (absoluta y relativa) y gráfico de barras; y una conclusión en lenguaje natural. Espacio para justificar; la retroalimentación valora la justificación, no solo el número. [RECURSOS] Habilidad MEP (Estadística y probabilidad): calcular e interpretar frecuencias relativas. Contexto verificado: encuesta a 25 estudiantes sobre el medio de transporte: 10 a pie, 8 en bus, 5 en bicicleta, 2 en carro. Pregunta: ¿qué medio supera la cuarta parte del grupo y cómo lo justificas? [REVISIÓN] Las frecuencias relativas suman 1 (o 100 %); la tarea no revela el procedimiento; la conclusión exige leer los dos registros.

ANEXO B. PROMPTS REALES DE LA EXPERIENCIA DE AULA (ANONIMIZADOS)

A diferencia del Anexo A, los siguientes son los prompts efectivamente redactados por el estudiantado del curso EM4610 durante las actividades del apartado 4.2. Se reproducen de forma textual —con cortes señalados mediante [...] cuando son extensos— y no contienen datos personales. Cada uno se acompaña de la categoría que ilustra y del aprendizaje que dejó.

B1. Artefacto interactivo de primer orden (actividad 2). Ilustra C1, C3 y C4.

Actúa como un desarrollador experto en React y diseña un artefacto interactivo completamente funcional en un solo archivo, orientado al aprendizaje de ecuaciones diferenciales de primer orden. [...] El artefacto debe integrar dos funcionalidades principales [...]: por un lado, un generador automático de ejercicios con soluciones paso a paso ocultas, y por otro, un detector de errores en soluciones matemáticas. [...] Es muy importante que todas las ecuaciones matemáticas y expresiones se escriban utilizando LaTeX, pero que no se muestren como código LaTeX. En su lugar, deben renderizarse visualmente como matemáticas correctamente formateadas en pantalla (por ejemplo, usando una librería como KaTeX o MathJax). El usuario nunca debe ver el código LaTeX crudo [...]. La generación de ejercicios debe basarse en coeficientes aleatorios sencillos (enteros pequeños) para evitar complejidad innecesaria [...]. En el módulo de detector de errores, [...] estos errores deben ser realistas y representar fallos comunes de los estudiantes, como errores algebraicos, uso incorrecto del factor integrante, mala separación de variables o errores en la integración. [...] Se recomienda el uso de colores con significado pedagógico: verde para indicar respuestas correctas, rojo para errores y azul para acciones o botones interactivos. [...]

La exigencia de renderizar la notación —«el usuario nunca debe ver el código LaTeX crudo»— es una formulación espontánea de C3; el módulo de detección de errores y la

lista de fallos típicos materializan C4. Un prompt correctivo posterior pidió que «el nuevo código [se hiciera] en formato HTML» para resolver una falla en la generación de ejercicios.

B2. Clasificador de ecuaciones de primer orden (actividad 2). Ilustra C1 y C4, y la tesis central.

Hola Claude. Puedes diseñar un clasificador interactivo HTML en donde se le permita al usuario ingresar una ecuación diferencial de primer orden y que el artefacto indique si la ecuación diferencial se debe resolver con los siguientes métodos: separable, homogénea, exacta, Bernoulli, Clairaut, lineales, no exactas y casi homogéneas. Además, según el método que indique, muestra un procedimiento general [...]. Para el input de la ecuación queremos que sea una casilla en donde el usuario pueda ingresar CUALQUIER forma de ED de primer orden y que el artefacto la entienda y sugiera el método [...]. En caso de que el usuario ingrese algo que no corresponda a una ecuación diferencial de primer orden puedes indicar que no es una ED de primer orden.

El grupo descubrió que el artefacto «se generó utilizando regex [...], lo que identifica patrones de texto, no realiza cálculos matemáticos»: una ecuación a la vez separable y homogénea se rotulaba solo como separable, las categorías lineal y exacta se trataban como excluyentes, y al reejecutar la clasificación cambiaban los pasos y «desaparecían números» (por ejemplo, con la ecuación $dy/dx = x + y + xy$). Es el caso más nítido de la tesis del artículo y de la necesidad de verificación explícita (C4).

B3. Material de estudio co-redactado, orden superior (actividad 3). Ilustra C5 y C6.

Necesito que desarrolles un resumen teórico-práctico sobre el tema de ecuaciones diferenciales lineales homogéneas de orden superior con coeficientes constantes que incluya la ecuación auxiliar, tipos de raíces, formas de la solución general y un ejemplo resuelto. Debes desarrollarlo de tal manera que funcione como un material de estudio para un estudiante que faltó a la clase en la cual se explicó dicho tema. Entonces necesito que se desarrolle un borrador que permita al usuario (estudiantes) editar, corregir y ampliar la información en dicho recurso.

El destinatario explícito —«un estudiante que faltó a la clase»— calibra el registro (C5); pedir un borrador para editar, corregir y ampliar, con espacios para que el estudiante responda preguntas y proponga un ejemplo propio, acerca la tarea a la categoría de problema, no ejercicio (C6). Al exportar el material a PDF, la notación se filtró como código LaTeX sin compilar (C3) y solo se recuperó legible al reexportar por otra vía.

B4. Batalla de IA con la transformada de Laplace (actividad 4). Ilustra C3 y C4.

Resuelve mediante la Transformada de Laplace el siguiente problema de valor inicial: $3y''(t) + y'(t) - y(t) = t \cdot U(t - 1)$, con condiciones iniciales $y(0) = 0$, $y'(0) = 1$. Muestra paso a paso: 1. la transformada de cada término; 2. el uso de propiedades de Laplace (derivadas, traslación, función Heaviside); 3. la ecuación en el dominio de s ; 4. la solución $Y(s)$; 5. la transformada inversa para obtener $y(t)$; 6. verifica que la solución cumple las condiciones iniciales. En cada paso, incluye una breve justificación de por qué se aplica esa propiedad o procedimiento.

El mismo problema se resolvió con dos modelos y a mano, y se contrastó en un cuadro. El grupo observó que uno de los modelos resolvía «de manera [...] más algorítmica, haciendo una resolución como si se entendiera ya el tema», lo que volvía «más difícil verificar si hay errores por la falta de desarrollo»: una caracterización, en sus propias palabras, de la fluidez sintética. La exigencia final de entregar la respuesta «en formato LaTeX» vuelve sobre C3, y la conclusión del grupo —contrastar siempre con «libros o el material provisto por el docente»— sobre C4 y la curaduría docente.