

Lineamientos para el uso de inteligencia artificial en contextos universitarios

Elaborada por [Juan David Gutiérrez](mailto:juagutie@uniandes.edu.co) (juagutie@uniandes.edu.co), profesor asociado de la [Escuela de Gobierno](#) de la Universidad de los Andes. Versión 5.0 (08/08/2023). [Licencia C.C. BY 4.0](#).

1. Introducción

- **Usamos herramientas de inteligencia artificial diariamente.** De manera consciente o inconsciente, prácticamente todos usamos algún tipo de [inteligencia artificial \(IA\)](#) en nuestra vida diaria, en particular sistemas basados en [modelos de lenguaje](#). Por ejemplo, usamos modelos de lenguaje cuando estamos escribiendo un mensaje de texto o un correo electrónico y la herramienta de “autocompletar” nos sugiere cómo terminar la oración. También usamos modelos de lenguaje para apoyarnos en la realización de actividades académicas, tales como sistemas de traducción automatizada o programas que detectan errores gramaticales, entre otros. Además, actualmente, todo tipo de plataformas web y aplicaciones móviles permiten acceder a [herramientas de IA para generar](#) textos y contenidos audiovisuales.
- **Algunas IA pueden apoyar procesos de aprendizaje, investigación, enseñanza, y otras actividades profesionales.** Existen usos adecuados de dichas tecnologías, que no solo pueden resultar útiles para apoyar procesos de [aprendizaje, enseñanza, e investigación](#) en la universidad sino también en el [plano profesional](#). Por ejemplo, diferentes herramientas de IA hoy en día contribuyen en las actividades de todas las [etapas de los procesos de política pública](#) y, en Colombia, [más de 50 entidades del sector público](#) han adoptado sistemas de decisión automatizada para apoyar la realización de sus funciones.
- **Pero ciertos usos de la IA son riesgosos, por tanto, su uso debe ser informado, transparente, ético y responsable.** Las herramientas de IA no son idóneas para realizar cualquier tipo de actividad y determinados tipos de uso pueden ser contraproducentes para el proceso pedagógico. Algunos usos de la IA pueden generar riesgos tanto para los usuarios como para terceros. Además, como se explica en el punto 4 de estos lineamientos, es indispensable ser conscientes de las implicaciones éticas, ambientales y de derechos humanos asociadas a la utilización de estas herramientas. Precisamente, debido a estos riesgos algunas universidades y profesores han publicado [lineamientos, principios, políticas, directrices y guías](#) sobre el uso de IA en el plano académico (ver listado en el punto 5 de estos lineamientos).
- **No deberíamos asumir que todos nuestros estudiantes son “nativos digitales” ni que quieren adoptar nuevas tecnologías.** La [actitud de las personas respecto de las tecnologías](#) es diversa y compleja, no se debe asumir que todos nuestros estudiantes tienen avidez por adoptar nuevas tecnologías. Por ejemplo, en los sondeos que he realizado en talleres con estudiantes sobre el uso de IA en la universidad, he encontrado que estos sistemas generan en ellos/as sentimientos muy diversos: miedo, intriga, emoción, incertidumbre, curiosidad, desconfianza, euforia, preocupación, asombro, expectativa, angustia, tranquilidad, confusión, fascinación, precaución, cuestionamiento, rechazo, desasosiego y confort. Además, la [capacidad de las personas para adoptar nuevas tecnologías](#) es heterogénea y depende de diferentes factores económicos, sociales y culturales, entre otros. Todo lo anterior implica que quienes enseñamos deberíamos implementar estrategias de enseñanza diferenciadas según las necesidades de nuestros estudiantes y que no debemos forzar el uso de nuevos sistemas de IA.

- **Orientación para el uso de IA dentro y fuera del aula universitaria.** Estos lineamientos ofrecen una orientación básica para profesores/as y estudiantes universitarios sobre cómo usar IA, particularmente [modelos de lenguaje de gran escala](#), de manera informada, transparente, ética y responsable. Los retos del uso de IA en el ámbito educativo no se agotan en aspectos asociados con los procesos de aprendizaje e investigación académica. Por ejemplo, la adopción de IA para (semi)automatizar procesos de contratación y evaluación de personal o para la vigilancia de las instalaciones educativas también tiene implicaciones legales y éticas. Pero estos asuntos desbordan el alcance de estos lineamientos dado que las reglas y principios aquí planteados se enfocan exclusivamente en el uso de IA con fines educativos y académicos dentro y fuera del aula universitaria.
- **Definiciones de tres términos clave.** A continuación, encontrarán definiciones de trabajo de tres términos tecnológicos utilizados en estos lineamientos: **(A) [Sistema de Inteligencia Artificial \(IA\)](#):** “sistema basado en máquinas que puede, para un conjunto determinado de objetivos definidos por el ser humano, hacer predicciones, recomendaciones o tomar decisiones que influyen en entornos reales o virtuales. Los sistemas de IA están diseñados para funcionar con diversos niveles de autonomía”. **(B) [Modelo de lenguaje](#):** algoritmos que procesan, analizan y producen el lenguaje natural en forma escrita o verbal. **(C) [Modelo de lenguaje a gran escala \(Large Language Model - LLM\)](#):** algoritmos entrenados a partir de una vasta cantidad de información cuya principal función es predecir el hilo de palabras más probable dado el contexto que lo antecede o le sigue.

2. Objetivos de estos lineamientos

- **Contribuir con la alfabetización digital.** Contribuir con el aprendizaje de conocimientos básicos sobre el uso de IA en contextos universitarios –con énfasis en los LLMs. Este objetivo incluye la enseñanza sobre sus beneficios y riesgos, sus implicaciones para la sociedad, y el desarrollo de competencias digitales críticas.
- **Promover el uso responsable de IA.** Promover el uso responsable y ético de estas tecnologías en los procesos de aprendizaje y para su futuro uso en la vida profesional.
- **Prevenir la vulneración a la integridad académica.** Prevenir situaciones en las cuales los estudiantes consciente o inconscientemente incurran en deshonestidad académica.

3. Reglas de uso de IA dentro y fuera del aula

- **Regla general.** Está permitido el uso de IA como herramienta de soporte para realizar las diferentes actividades de aprendizaje. Los parámetros para el uso de estas herramientas que se describen a continuación distinguen entre herramientas de “bajo riesgo” y de “alto riesgo”, según los riesgos que la respectiva IA genere para el proceso pedagógico, para los usuarios y para terceros.
- **Reglas para uso de IA de bajo riesgo.** Podrán usar libremente herramientas de IA de bajo riesgo, es decir aquellas que les permitan corregir o revisar contenidos generados por los propios estudiantes o aquellas que les permitan recolectar y procesar datos. Por ejemplo, se consideran de bajo riesgo las herramientas de corrección gramatical, de traducción, de transcripción audio-texto, y de búsqueda de información en Internet, entre otras. Se recomienda que cuando dichas herramientas permitan realizar labores significativas (ej. traducción), se indique en la sección que corresponda (ej. metodología) sobre el uso de dichas herramientas. Si tienen dudas sobre cómo clasificar el riesgo de uso de una IA específica, sugiero consultar al profesor.

- **Reglas para uso de IA de alto riesgo.** Son de alto riesgo, las herramientas de IA que se usen para generar texto sintético. Es decir, los LLMs que pueden accederse a través de *chatbots*, tales como *ChatGPT*, *Bard*, *Bing*, y *Claude*, etc. Solo podrán incluir en sus entregas individuales y colectivas contenidos generados por IA cuando cumplan con cuatro requisitos:
 - i. **Uso informado.** Que previo a su uso hayan investigado quién o qué empresa desarrolló la herramienta, cómo fue desarrollada, cómo funciona, qué funciones puede desempeñar, cuáles son sus términos de uso, qué hace la empresa con la información que recolecta del usuario, y qué limitaciones y/o riesgos presenta para el usuario y/o para terceros.
 - ii. **Uso transparente.** Que indiquen detallada y expresamente qué herramienta usaron y cómo la usaron en la sección de metodología o en la que consideren pertinente. Este es un requisito que actualmente solicitan revistas científicas y casas editoriales, tales como [Nature](#) y [Cambridge University Press](#), a quienes entregan manuscritos para revisión de pares.
 - iii. **Uso ético.** En relación con el uso del LLMs, que sea posible distinguir qué fue escrito o producido directamente por ustedes y qué fue generado por una herramienta de IA. Respecto de este punto, aplican las reglas de citación generales, por ejemplo, usar comillas si incluyen apartes literales además de la referencia correspondiente. Para citar texto sintético producido por LLMs, pueden consultar las [reglas](#) sugeridas por el equipo de estilo de APA. Las infracciones a esta regla, particularmente respecto del segundo requisito, serán abordadas bajo un enfoque similar al que aplica en casos de plagio. Además, *ChatGPT* no es una persona, luego no puede ser considerado su coautor. Algunas [revistas científicas](#) ya han actualizado sus políticas de publicación para aclarar que los modelos de lenguaje como *ChatGPT* no satisfacen sus criterios de atribución de “autoría”. Por otra parte, el uso ético de sistemas de IA –en general, no solo LLMs– debería estar informado por cuatro [principios desarrollados por la bioética](#): beneficencia (consecución de beneficios), respeto a la autonomía de las personas (incluye la obtención de consentimiento si hay terceros involucrados), justicia (especialmente la no discriminación) y no maleficencia (no producir daño a otros de manera intencionada o por imprudencia).
 - iv. **Uso responsable.** Se recomienda que el uso de estas IA se circunscriba a etapas tempranas de la investigación, para inspirar o sugerir rumbos, no para producir contenido que luego sea incluido en sus entregas. En todo caso, si optan por usar información obtenida a través de un LLM, deben acreditar que contrastaron rigurosamente dicha información con otras fuentes confiables dado que estos sistemas tienden a ofrecer información [imprecisa, errónea, y falsa](#). El uso responsable también implica evitar que la utilización del sistema cause daños al usuario o a terceros. Por ejemplo, no se recomienda incluir en las consultas realizadas a los *chatbots* datos personales o información confidencial propia o ajena porque una vez introducida dicha información se pierde control de quien puede acceder a la misma y se genera el riesgo de que sea conocida por terceros no autorizados.
- **Usos plausibles de los LLMs:**
 - i. **Explorar.** Explorar nuevos asuntos y buscar inspiración, es decir, usar el *chatbot* en las etapas más tempranas de la investigación.

- ii. **Automatizar tareas básicas.** Apoyar ejercicios que impliquen el reconocimiento y/o reproducción de patrones en textos, por ejemplo: traducir, resumir, modificar el tono y estilo del texto –por ejemplo, para que la redacción de texto sea más clara o asequible a determinadas audiencias–, corregir ortografía y gramática, conversión audio–texto, programación, etc. Pero tampoco pueden fiarse porque deben revisar con cuidado el *output* del LLM.
 - iii. **Escritura trivial.** Apoyar tareas de escritura poco importantes. Para decidir si algo es trivial o no, piensen en la siguiente heurística: ¿qué ocurriría si mi interlocutor se entera que utilicé un LLM para escribir el mensaje?
 - iv. **Sparring.** Probar argumentos o ideas propias a través de la interacción con el *chatbot*, por ejemplo, solicitando al sistema que responda con críticas o contraargumentos a los planteamientos del estudiante.
- **Usos no recomendados de los LLMs:**
 - i. **Usar el *chatbot* como principal o única fuente para buscar información factual o técnica.** No lo recomiendo, [los LLMs no son herramientas de búsqueda fiables](#). Además, los buscadores de Internet son más eficaces y confiables para esa tarea. En todo caso, dado que hacer búsquedas en los *chatbots* puede ser percibido como una manera de obtener información más fácil y rápida, recomiendo que la información obtenida por ese medio siempre sea cuidadosamente contrastada con fuentes confiables.
 - ii. **Uso de texto sintético para asuntos sustanciales.** No lo recomiendo por la tendencia de los LLMs a producir respuestas con información [imprecisa, errónea y falsa](#). En todo caso, si lo hiciera, debe contrastarse cada punto con otras fuentes confiables.
 - iii. **Realizar cálculos matemáticos.** Los LLMs [no están entrenados para realizar cálculos matemáticos](#) y, por ende, con frecuencia fallan en responder operaciones aritméticas sencillas.
 - **Acompañamiento del profesor y desarrollo de competencias digitales.** El profesor dispondrá tiempo de la clase para explicar qué es la IA, qué son los modelos de lenguaje a partir de los cuales funcionan herramientas como *ChatGPT*, qué oportunidades y riesgos generan para los trabajos académicos y profesionales, y cuáles son las implicaciones éticas, ambientales y de derechos humanos asociadas al uso de estas herramientas. El profesor siempre estará disponible para aclarar el alcance de estos lineamientos, para discutir y co-crear estos lineamientos, y para resolver preguntas puntuales sobre el uso de IA. Además, el profesor procurará realizar actividades que contribuyan con la alfabetización digital de los estudiantes, en particular para el desarrollo y fortalecimiento de las [competencias digitales](#) tales como las habilidades para evaluar críticamente los beneficios y riesgos de las nuevas tecnologías.
 - **Supervisión.** El profesor usará diferentes fuentes para identificar potenciales situaciones de deshonestidad académica a través del uso de IA, pero no dependerá de las herramientas disponibles para detectar texto sintético dado que actualmente ninguna de esas herramientas es fiable. En todo caso para evitar riesgos asociados a la integridad académica, la principal estrategia que usará el profesor será [pedagógica](#) no punitiva.

4. Por qué es necesario el uso informado, transparente, ético y responsable de IA en la universidad

- **El uso de herramientas de IA debe ser informado, transparente, ético y responsable por –al menos– cuatro tipos de motivos:** (1) porque los LLMs no siempre son confiables; (2) porque hay riesgos de que ciertos usos afecten negativamente los procesos de aprendizaje; (3) por el riesgo de que los usuarios traten la IA como si fuera un humano; y, (4) porque el uso de las herramientas tiene implicaciones éticas y de derechos humanos debido a la manera como fueron desarrolladas y/o porque algunas herramientas pueden replicar o amplificar problemáticas sociales tales como la discriminación. A continuación, se explica cada de los cuatro motivos en detalle:
- **En primer lugar**, porque sus [respuestas NO siempre son confiables](#) a pesar de que , por ejemplo, una IA generadora produce textos que tienen [apariencia de ser convincentes](#). Los sistemas basados en LLMs como *ChatGPT* [no funcionan con la precisión](#) de otras herramientas usadas en ambientes de aprendizaje como las calculadoras. De hecho, *ChatGPT* tiende a incluir en sus respuestas [información falsa o fantasiosa](#). El *chatbot* incorporado por Microsoft a su [explorador de Internet, Bing](#), y el [chatbot de Google, Bard](#), también han presentado los [mismos tipos de problemas](#). Estos sistemas [no distinguen lo verdadero de lo falso](#). ¿Por qué ocurre esto? Los LLMs hilan palabras a partir de inferencias probabilísticas de los datos con los cuales fueron entrenados, pero no tienen la capacidad de entender lo que producen ni asocian significados a las palabras que emiten (son “[loros estocásticos](#)”).
- **En segundo lugar**, porque el uso de herramientas generadoras de texto puede desestimular la motivación de los estudiantes para [escribir y pensar por su propia cuenta](#). Vale la pena reiterar que las actividades de aprendizaje de esta asignatura buscan desarrollar sus habilidades cognitivas y que estos lineamientos buscan evitar que algunas [herramientas de IA se conviertan en mecanismos de plagio automatizadas](#).
- **En tercer lugar**, por el riesgo de que los usuarios, consciente o inconscientemente, traten la conducta de la IA como si fuera humana ([Efecto Eliza](#)). Por ejemplo, las herramientas basadas en [LLMs no “entienden” los textos que producen](#), imitan patrones de lenguaje a partir de la síntesis de grandes volúmenes de datos con base en los cuales generan las secuencias de palabras. Este problema de [antropomorfizar las máquinas](#) se exacerba por cuanto algunos *chatbots* han producido [respuestas violentas o peligrosas que podrían llevar a las personas a realizar conductas dañinas](#).
- **En cuarto lugar**, por las implicaciones éticas y de derechos humanos asociadas al uso de ciertos sistemas de IA dado que: algunas herramientas tienden a reproducir o amplificar [estereotipos derogatorios](#) y [discriminatorios](#) asociados al [género](#), raza, etnia o discapacidad; las tecnologías podrían haber sido desarrolladas a partir de la [violación masiva de derechos de autor](#); nuevas formas de [colonialismo a través de la extracción no consentida de información](#) de comunidades históricamente marginalizadas; [algunas herramientas](#) habrían sido desarrolladas en [contextos de explotación laboral](#); el desarrollo y operación de dichos sistemas genera una [huella de carbono considerable](#); y, la potencial [violación a los derechos de privacidad y de protección de datos personales](#) de quienes las usan. Para profundizar en los daños asociados al uso de IA Generativa, recomiendo consultar el reporte “[Ghost in the machine: Addressing the consumer harms of generative AI](#)” (2023) del *Norwegian Consumer Council* y el reporte “[Generating Harms: Generative AI’s Impact & Paths Forward](#)” del *Electronic Privacy Information Center*.

5. Otros recursos para profesores/as

- **Otros lineamientos sobre uso de IA en la universidad.** Recomiendo consultar lineamientos, políticas, y principios de otras universidades, profesores/as y otras organizaciones, tales como: [University of Washington](#), [Hertie School](#), [University of Tartu](#), [Yale University](#), [University of Helsinki](#), [Technical University of Munich](#), [University College London](#), [UNESCO](#), [Departamento de Educación del Reino Unido](#), [Boston University](#), [Warwick University](#), [Universidad de Buenos Aires \(IALAB\)](#), [Universidad del Desarrollo](#), [Russell Group](#), y [listado de Lance Eaton](#).
- **Literatura introductoria sobre LLMs.** (A) Para una introducción sobre cómo fueron desarrollados y cómo funcionan los LLMs (en particular, GPT-3), les recomiendo este [texto de Timothy B. Lee y Sean Trott](#). (B) Para una explicación en Castellano sobre cómo funcionan los LLMs y cuáles son sus principales riesgos, recomiendo el texto "[El riesgo de los grandes modelos lingüísticos: convertirse en loros estocásticos](#)", por Esther Sánchez García y Michael Gasser, que resume el artículo "[On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?](#)" de Emily M. Bender, Timnit Gebru, Angelina McMillan-Major y Shmargaret Shmitchell.
- **Recursos para la enseñanza.** (A) Quienes estén interesados en los retos que los LLMs generan para enseñar a escribir en las universidades les sugiero consultar el documento titulado "[Adaptando la escritura universitaria a la era de los Modelos de Lenguaje a Gran Escala como ChatGPT: Próximos pasos para educadores](#)" de Anna Mills y Lauren M. E. Goodlad. (B) Con el mismo fin, también sugiero los materiales contenidos en "[AI Text Generators and Teaching Writing: Starting Points for Inquiry](#)" de Anna Mills. El artículo de Anna Mills, Maha Bali y Lance Eaton titulado "[How do we respond to generative AI in education? Open educational practices give us a framework for an ongoing process](#)" documenta y reflexiona cómo profesores de diferentes universidades del mundo han adoptado nuevas prácticas pedagógicas para responder a los retos generados por la IA Generativa. (C) Para una reflexión sobre cómo integrar IA en procesos de formación en ciencias humanas, recomiendo este [blog post](#) de Bastien Bosa. (D) Para una reflexión sobre qué estrategias pedagógicas pueden adoptarse para [evaluar estudiantes "en la era de ChatGPT"](#), recomiendo el artículo de M. Paz Prendes Espinosa, María del Mar Sánchez Vera y Víctor González Calatayud. (E) MIT ofrece de manera gratuita su programa [DAILY Curriculum](#) que contiene materiales para enseñar conocimientos básicos sobre IA a adolescentes.
- **Aproximación crítica al uso de IA en entornos académicos.** (A) Recomiendo consulta el [listado de recursos de la profesora Maha Bali](#) sobre desarrollo de alfabetización crítica en IA. (B) Para una aproximación crítica sobre el uso de LLMs para el proceso de investigación y revisión bibliográfica, recomiendo el texto de "[Nadie escribiendo y nadie leyendo: los generadores de texto con inteligencia artificial y la ciencia que queremos](#)" de Ariel Guersenzvaig y Javier Sánchez Monedero. (C) También sugiero leer una [reflexión](#) de los profesores Mohammad Hosseini, Lex Bouter, y Kristi Holmes en la cual argumentan que no es deseable adoptar herramientas de IA en la educación sin reflexionar antes sobre sus efectos y sesgos y sin haber adoptado medidas que mitiguen sus riesgos. (D) El artículo "[How ChatGPT robs students of motivation to write and think for themselves](#)" de Naomi S. Baron discute cómo el uso de LLMs puede afectar negativamente los procesos creativos de estudiantes. (E) El [blog post](#) de Dimitrinka Atanasova reflexiona sobre cómo los LLMs pueden reducir las barreras para que estudiantes e investigadores escriban en Inglés (cuando este no es su primer idioma), pero que sus efectos respecto de la inclusividad cultural en la educación superior son más ambiguos. (F) También es valiosa la propuesta de Kathryn Conrad que aboga por una "[Carta de Derechos de la IA para la Educación](#)" que propone derechos para

profesores y estudiantes que sirvan de base para la formulación de políticas y medidas protectoras.

6. Lineamientos abiertos al cambio

- La IA es un conjunto de herramientas que se transforma rápidamente, razón por la cual estos lineamientos permanecerán abiertos a futuras evaluaciones, modificaciones y revisiones.
- A lo largo del semestre académico el profesor abrirá espacios para discutir estos lineamientos con los/las estudiantes y, de ser necesario, podrán introducirse modificaciones a partir de ejercicios de co-creación.
- Agradezco a mis colegas y demás personas que aportaron sus comentarios y críticas a versiones iniciales de este documento y también a quienes han participado en las charlas y talleres que he realizado sobre el uso de modelos de lenguaje de gran escala en contextos universitarios.

7. Cambios introducidos a la presente versión de los lineamientos

- Introduce cambios de forma y de fondo respecto de la versión anterior de este instrumento ([v. 4.2](#)) que resumo a continuación.
- **Cambios de forma: (A)** La anterior versión se denominaba “Política de uso de inteligencia artificial en la asignatura”, el nuevo título es más amplio puesto que el instrumento es útil para ámbitos académicos que van más allá de las actividades en clase (ej. actividades de investigación). **(B)** Introduce nuevas fuentes que pueden consultarse en forma de hipervínculos a lo largo del texto.
- **Cambios de fondo: (A)** En el punto 1 de la nueva versión de los lineamientos incluí un glosario de términos clave y en el punto 3 introduje una nueva sub-sección con usos plausibles y usos no recomendados de los LLMs por parte de estudiantes. **(B)** Además, en esta versión incluí una nueva sección de recursos para profesores/as que incluye: ejemplos de otros lineamientos, literatura básica sobre LLMs, recursos para la enseñanza, y literatura crítica sobre uso de IA en contextos educacionales. **(C)** Por otra parte, la anterior versión de este instrumento indicaba que el profesor usaría sistemas de detección de texto sintético, pero he eliminado esas alusiones en la presente versión porque los sistemas actualmente disponibles producen un [alto nivel de falsos positivos y negativos](#). Por este motivo, por ejemplo, el *Teaching Center* de la Universidad de Pittsburgh publicó un [comunicado de prensa](#) en el cual informaba que no avalaba ni promovía ninguna de las herramientas actualmente disponibles.

8. Uso de estos lineamientos en sus clases

- Estos lineamientos están publicados bajo una [licencia Creative Commons Atribución 4.0](#), luego si usted es un/a educador/a interesado/a en formular sus propias políticas o directrices, puede compartir y adaptar estos lineamientos siempre que realice la correspondiente atribución e indique si realizó cambios.
- Agradezco a quienes quieran enviarme sus comentarios, críticas y sugerencias que por favor me escriban a juagutie@uniandes.edu.co.
