

Resumen:

Este proyecto se propone investigar la representación de la indigenidad en sistemas de inteligencia artificial (IA), con un enfoque particular en la identificación de posibles sesgos y estereotipos. Utilizando herramientas de Análisis Crítico del Discurso (ACD) y Procesamiento del Lenguaje Natural (NLP), se analizarán las representaciones textuales y visuales generadas por modelos de IA como ChatGPT o DALL-E. Este estudio busca reflexionar sobre formas de representación justa y ética, y se desarrollará de manera interdisciplinaria en colaboración con expertos en inteligencia artificial.

La inteligencia artificial está ganando un peso significativo en la conformación del imaginario social contemporáneo. Las decisiones automatizadas y las representaciones generadas por sistemas de IA influyen cada vez más en nuestra percepción del mundo y de los diferentes grupos sociales. En este contexto, es fundamental examinar cómo la IA representa a los pueblos indígenas, un grupo históricamente marginado y sujeto a estereotipos y sesgos.

La creencia generalizada de que la IA es imparcial y justa se ha visto desafiada por investigaciones que demuestran lo contrario. Estudios recientes han revelado que los algoritmos de IA pueden perpetuar y acentuar sesgos preexistentes, resultando en representaciones desiguales y estereotipadas. Este proyecto se enmarca en el contexto de reflexiones globales sobre la ética en la IA, evidenciadas en documentos y acuerdos internacionales como la Declaración de Montreal para la Responsabilidad Ética Global de la Inteligencia Artificial (2018), las Directrices de la OCDE (2019) y los Principios de la IA de la UNESCO (2019). Estos documentos subrayan la importancia de principios éticos como la transparencia, la equidad, la inclusión y la rendición de cuentas en el desarrollo y la implementación de la IA.

Fundamentación teórica:

En el ámbito de los estudios culturales, la literatura académica ha demostrado repetidamente la persistencia de discursos estereotipados, sesgados y frecuentemente racistas sobre los pueblos indígenas de México. Ya sea en el cine, la televisión, la literatura, la prensa, la política o el material didáctico, estas representaciones suelen ofrecer una visión simplificada y exotizada de los indígenas, despojándolos de su realidad contemporánea y multidimensional. Gayatri Spivak (1998) destacó cómo estas representaciones actúan como una forma de violencia epistémica, que no solo es un residuo de la colonización, sino que continúa manifestándose en la educación y otros espacios públicos, imponiendo una imagen de los indígenas completamente ajena a su realidad cotidiana.

La noción de los indígenas como símbolos de un "pasado glorioso" de México es otra faceta problemática de estas representaciones. Según Zuñiga González (1998), los indígenas son frecuentemente idealizados como "indios puros", figuras que simbolizan los mitos fundadores y el patrimonio histórico del país. Esta representación abstracta es construida por y para la retórica nacional, y está completamente desvinculada de la historia y realidad contemporánea de estos pueblos. Los discursos dominantes sobre los indígenas suelen ser formulados desde la perspectiva de la sociedad mayoritaria, donde se exalta la cultura étnica mientras se expande el imaginario occidental, a menudo a través de nociones de autenticidad y exotismo. Estas categorías no solo refuerzan la percepción de los indígenas como un ente cultural estático, sino que también oscurecen la diversidad real y las transformaciones contemporáneas dentro de las comunidades indígenas.

Históricamente, las representaciones de los pueblos indígenas en medios y discursos académicos y populares han sido enmarcadas dentro de una visión esencialista, donde se les define por características culturales percibidas como fijas y folklóricas. Esta visión esencialista reduce las culturas indígenas a un conjunto estereotipado de imágenes y atributos como trajes tradicionales, danzas y artesanías, que son presentados como si fueran inmutables a lo largo del tiempo. El concepto de "geografías imaginadas", explorado por Edward Said (2023), es particularmente útil para entender cómo se construyen estos discursos. Las representaciones de los pueblos indígenas en México son a menudo proyecciones de deseos occidentales que hablan más de sus creadores que de los sujetos representados. Estos discursos sirven para perpetuar desigualdades y subordinaciones, al posicionar a los

indígenas dentro de narrativas estáticas y folklóricas que limitan su participación en la sociedad contemporánea.

En la actualidad, estas mismas preocupaciones sobre la perpetuación de estereotipos y prejuicios en la representación de los pueblos indígenas se están trasladando a los estudios sobre inteligencia artificial. Desde la ingeniería de software, se observa un eco notable de estas problemáticas en diversas investigaciones. Varios autores han comenzado a identificar claramente cómo los sistemas de IA, especialmente los modelos de lenguaje y generativos, pueden reproducir y hasta amplificar los prejuicios existentes en los conjuntos de datos con los que son entrenados. Las investigaciones recientes, como las de Ananya (2024), han subrayado los límites éticos de la IA, resaltando cómo estas tecnologías pueden acentuar estereotipos de género, raciales y culturales, siendo simplistas en sus representaciones. Este reconocimiento señala una creciente conciencia de que, al igual que en los estudios culturales, en la inteligencia artificial también es imperativo abordar y mitigar estos sesgos para evitar la perpetuación de desigualdades y discriminaciones en el diseño y aplicación de estas tecnologías avanzadas.

La problemática central de este fenómeno reside en la forma en que estos sistemas de IA "aprenden" y se desarrollan. Al depender de grandes cantidades de datos recopilados de la sociedad humana, estos modelos inevitablemente capturan y reflejan las desigualdades y sesgos inherentes a esos datos. Esto se ve ejemplificado en casos documentados, como los señalados por Brack et al (2024), quienes han demostrado cómo la IA puede reforzar estereotipos de género, raza y cultura. En Zou & Schiebinger (2018), se evidenció que un algoritmo de reconocimiento de imágenes etiquetaba incorrectamente las imágenes de una boda tradicional india como "disfraces", revelando una comprensión superficial y errónea de contextos culturales no occidentales. Estos errores no son meramente anecdóticos, sino que señalan un problema sistémico dentro de la tecnología de IA. La reflexión sobre la ética de la IA también se extiende al ámbito de la sostenibilidad, como se observa en los trabajos de Antonio Luis Terrones Rodríguez (2022). Él plantea que la integración de principios éticos debe ir de la mano con un compromiso hacia la sostenibilidad, argumentando que las soluciones tecnológicas por sí solas no son suficientes para enfrentar los desafíos éticos contemporáneos.

La convergencia de estos enfoques subraya una tendencia creciente hacia una comprensión más crítica y matizada de la IA. No solo se trata de desarrollar tecnología más avanzada, sino de asegurar que esta tecnología funcione dentro de un marco ético que reconozca y rectifique sus propias limitaciones. Este cambio de paradigma hacia una IA éticamente consciente y culturalmente sensible es crucial para su desarrollo sostenible y justo, lo que a su vez garantiza que la tecnología sirva a toda la sociedad, respetando y fomentando la diversidad y la inclusión.

Justificación:

El impacto del discurso generado por la IA en las relaciones políticas y los tipos de representación social es innegable. La IA moldea nuestros imaginarios, pensamientos, comportamientos y cultura. Si estos datos no son curados adecuadamente para equilibrar o corregir sesgos, la IA puede perpetuar o incluso amplificar estereotipos negativos y representaciones sesgadas. Esto es crítico cuando las representaciones generadas por la IA pueden influir en la percepción pública sobre comunidades específicas, como los pueblos indígenas en México, contribuyendo a una comprensión distorsionada y estigmatizada de estas comunidades.

A través del análisis de un corpus de representaciones visuales y textuales generadas por IA, el proyecto evaluará de manera exhaustiva cómo estas imágenes y textos configuran los imaginarios sociales y culturales. Este análisis permitirá identificar patrones recurrentes y elementos discursivos que reflejen estereotipos y narrativas históricamente arraigadas, así como nuevas formas de representación que puedan surgir en la era digital. Al examinar tanto las representaciones explícitas como las implícitas, se buscará entender cómo la IA contribuye a la perpetuación o transformación de las percepciones sociales sobre los pueblos indígenas.

Además, se desarrollarán recomendaciones y estrategias específicas para mitigar los sesgos en los sistemas de IA. Esto incluirá la propuesta de mejoras en los datos de entrenamiento, la implementación de algoritmos más inclusivos y la creación de directrices éticas claras que guíen el desarrollo y uso de estas tecnologías. Las estrategias se enfocarán en promover una representación más justa y equitativa, asegurando que las voces y experiencias de las comunidades indígenas sean adecuadamente reflejadas y respetadas en las producciones de

IA. Se contemplarán también prácticas de revisión y validación continua para mantener la integridad y equidad en las representaciones generadas.

La colaboración interdisciplinaria será clave para abordar estas cuestiones desde múltiples perspectivas. Expertos en inteligencia artificial y estudios culturales trabajarán conjuntamente, aportando sus conocimientos y experiencias para crear un enfoque integral y ético. Esta colaboración permitirá una comprensión más profunda y matizada de los desafíos y oportunidades en la representación de los pueblos indígenas en la IA, y garantizará que las soluciones propuestas sean tanto técnicamente viables como culturalmente sensibles y respetuosas. El trabajo conjunto buscará no solo identificar y corregir los sesgos existentes, sino también establecer un marco de desarrollo de IA que promueva la diversidad, la inclusión y la justicia social.

Preguntas:

1. ¿De qué manera los sistemas de IA representan a los pueblos indígenas mexicanos en sus producciones textuales y visuales?
2. ¿Qué estereotipos y sesgos, tanto históricos como contemporáneos, se reflejan en las representaciones generadas por los sistemas de IA sobre los pueblos indígenas mexicanos?
3. ¿Cómo influyen la estructura de los datos de entrenamiento, el proceso de entrenamiento, el etiquetamiento, las actividades de crowdsourcing y la retroalimentación con el usuario en las representaciones de los pueblos indígenas generadas por la IA?

Hipótesis:

1. Las representaciones generadas por sistemas de IA sobre los indígenas mexicanos perpetúan estereotipos históricos y folclóricos, contribuyendo a una comprensión distorsionada de estas comunidades.
2. Las representaciones textuales y visuales de la indigenidad en sistemas de IA están influenciadas por asimetrías de poder en los datos de entrenamiento y en el proceso de desarrollo de la IA.
3. La colaboración interdisciplinaria y la implementación de estrategias de mitigación pueden mejorar la representación de los pueblos indígenas en los sistemas de inteligencia artificial.

Objetivos:

1. Analizar de qué manera los sistemas de IA representan a los pueblos indígenas mexicanos en sus producciones textuales y visuales.
2. Identificar y analizar los sesgos y estereotipos en las representaciones textuales y visuales de la indigenidad generadas por sistemas de IA.
3. Evaluar la configuración de imaginarios sociales derivadas de estas representaciones.
4. Desarrollar recomendaciones y estrategias para mitigar los sesgos en los sistemas de inteligencia artificial, promoviendo una representación más justa y equitativa.
5. Fomentar la colaboración interdisciplinaria para abordar los sesgos en la IA desde múltiples perspectivas.

Metodología:Corpus y Herramientas:

El corpus de este estudio se formará utilizando plataformas de inteligencia. Se solicitará la creación de imágenes y descripciones textuales de personas indígenas mexicanas. Estas representaciones se almacenarán y se harán accesibles para su análisis.

Análisis Cuantitativo y Cualitativo:

1. Organización preliminar del corpus visual:

Descripción cuantitativa de aspectos como distribución por género y categoría etaria, tipo y estilo de iconografía utilizada, tipo de plano y paisajes representados, y expresiones faciales de las personas retratadas.

2. Análisis iconográfico interpretativo:

Focalización en las connotaciones históricas y culturales de los elementos presentes en las imágenes.

3. Análisis detallado del corpus textual:

Utilización de spaCy para el procesamiento del lenguaje natural (NLP), incluyendo la tokenización, categorización gramatical de textos y análisis de dependencias.

Análisis Crítico del Discurso (ACD):

El Análisis Crítico del Discurso (ACD) es una metodología esencial en nuestro proyecto para investigar la representación de la indigenidad en los sistemas de inteligencia artificial. Esta metodología, desarrollada y perfeccionada por académicos como Teun A. van Dijk (1997, 2013, 2016), Norman Fairclough (1989, 1995) y Ruth Wodak (2000), se centra en la comprensión de cómo el discurso se configura y perpetúa en contextos sociales específicos, y cómo este discurso reproduce relaciones de poder y desigualdades sociales.

El ACD no solo se enfoca en el contenido explícito del texto, sino también en las estructuras subyacentes que moldean y sostienen estas narrativas. En el contexto de nuestro estudio, el ACD se empleará para desentrañar cómo los sistemas de IA, como ChatGPT y DALL-E, generan representaciones textuales y visuales de los pueblos indígenas mexicanos y cuáles son los sesgos y estereotipos implicados en estas representaciones.

El proceso de ACD en este proyecto incluye varias etapas clave:

- Selección del corpus

El corpus estará compuesto por textos e imágenes generados en respuesta a prompts específicos relacionados con la indigenidad mexicana. Estas representaciones se analizarán para identificar patrones discursivos recurrentes.

- Análisis contextual

El primer paso en el ACD es situar el discurso en su contexto social y político. Esto implica examinar el trasfondo histórico y contemporáneo de la representación de los pueblos indígenas en México. Se revisarán estudios previos sobre estereotipos y representaciones mediáticas de los indígenas para comprender cómo estos discursos se han desarrollado y transformado a lo largo del tiempo.

- Identificación de macronarrativas:

Utilizando las técnicas desarrolladas por Van Dijk y Fairclough, identificaremos las macronarrativas predominantes en el corpus. Esto incluye examinar cómo se estructuran los relatos sobre los pueblos indígenas, qué temas y símbolos son recurrentes, y cómo se representan las relaciones de poder. Por ejemplo, se analizará si los indígenas son representados como figuras del pasado, con una conexión estática a tradiciones folclóricas, o si se les presenta en contextos contemporáneos.

- Análisis de actores y posicionamientos

El ACD también se enfocará en los actores involucrados en la creación de estos discursos y sus posicionamientos. Esto incluye no solo los desarrolladores de IA y sus equipos, sino también los datos de entrenamiento utilizados, que a menudo reflejan las perspectivas y sesgos de la sociedad mayoritaria. Se explorará cómo estos actores influyen en las representaciones producidas por los sistemas de IA.

- Reglas de interacción y relaciones de poder

Analizaremos las reglas de interacción que rigen cómo diferentes plataformas de IA generan respuestas. Esto incluye cómo los algoritmos responden a diferentes prompts y cómo manejan las solicitudes que pueden contener sesgos implícitos. Se evaluará si estas interacciones refuerzan o desafían las relaciones de poder existentes y cómo podrían ser modificadas para promover una representación más justa.

- Evaluación crítica y propuestas de mejora

Finalmente, el ACD se utilizará para evaluar críticamente los hallazgos y desarrollar recomendaciones para mitigar los sesgos identificados. Esto puede incluir sugerencias para mejorar los conjuntos de datos de entrenamiento, ajustar los algoritmos y desarrollar pautas éticas para la representación de grupos marginados en la IA.

El uso del ACD en este proyecto permitirá una comprensión profunda de cómo los discursos generados por la IA no solo reflejan, sino también configuran, las percepciones sociales sobre los pueblos indígenas. Al desenmascarar las estructuras de poder y los sesgos implícitos en estas representaciones, podremos avanzar hacia una IA más ética y equitativa.

Cronograma de trabajo:

Actividades	Meses					
	Jun	Jul	Agos	Sept	Oct	Nov
Revisión bibliográfica						
Revisión del panorama de diferentes modelos de IA y su representación de la indigenidad						
Delimitación del corpus y aplicación de herramientas de ACD y NLP para análisis de datos						

Presentación de avances en el XXXV Encuentro Nacional e internacional AMIC. Inteligencia Artificial y Procesos Comunicativos: desde las prácticas innovadoras, hasta las dimensiones éticas y de carácter legal (Puebla)						
Redacción de resultados y publicación de capítulo en libro digital						

1. Revisión bibliográfica (Junio - Julio 2024): Durante estos meses se llevará a cabo una revisión exhaustiva de la literatura existente sobre la representación de la indigenidad en IA, el análisis crítico del discurso y el procesamiento del lenguaje natural. Se identificarán las principales teorías y estudios previos que sustenten el marco teórico de la investigación.
2. Revisión del panorama de diferentes modelos de IA y su representación de la indigenidad (Junio - Julio 2024): Se realizará una evaluación de diversos modelos de IA, para comprender cómo representan a las comunidades indígenas mexicanas. Esto incluirá el análisis de la metodología de entrenamiento y los datos utilizados.
3. Delimitación del corpus y aplicación de herramientas de ACD y NLP para análisis de datos (Agosto - Octubre 2024): Se seleccionará un corpus representativo de textos e imágenes generados por IA. Se aplicarán herramientas de Análisis Crítico del Discurso (ACD) y Procesamiento del Lenguaje Natural (NLP) para analizar los sesgos y estereotipos presentes en las representaciones.
4. Presentación de avances en el XXXV Encuentro Nacional e Internacional AMIC (Octubre 2024): se presentarán los avances de la investigación en este encuentro, centrado en la Inteligencia Artificial y Procesos Comunicativos. Se obtendrá retroalimentación de colegas y expertos en el campo.
5. Redacción de resultados y publicación de capítulo en libro digital (Noviembre 2024): se procederá a la redacción de los resultados de la investigación, que se mandará como propuesta para la publicación de un capítulo en un libro digital. Este capítulo incluirá las recomendaciones y estrategias para una representación más justa y equitativa en los sistemas de IA.

Solicitud de presupuesto

- | | |
|---------------------------------------|--------|
| 1. Toner TN420 para impresora Brother | \$3000 |
| 2. Memoria USB (1 unidad) | \$2000 |

La adquisición de tóner es esencial para la impresión de documentos de trabajo, informes preliminares y finales, así como materiales para presentaciones. Este insumo es indispensable para la documentación y difusión de los hallazgos del proyecto, asegurando que la información esté disponible en formatos accesibles para todos los participantes y partes interesadas.

La memoria USB es esencial para el almacenamiento y la transferencia segura de grandes cantidades de datos y documentos del proyecto, asegurando la protección y portabilidad de la información crítica.

Este presupuesto ha sido diseñado cuidadosamente para asegurar que el recurso solicitado contribuya de manera directa y significativa al cumplimiento de los objetivos del proyecto, promoviendo una gestión eficiente y responsable de los fondos disponibles. Aseguramos que se ha considerado la austeridad, alineándonos con las normas establecidas por la Universidad de Guadalajara para garantizar el uso prudente y efectivo de los recursos financieros.

Criterios éticos:

La investigación se llevará a cabo con estricta adherencia a principios éticos, asegurando la protección de los derechos y la dignidad de las comunidades indígenas representadas en los sistemas de IA. Se adoptarán las siguientes medidas:

1. Transparencia: La investigación se conducirá de manera transparente. Los resultados serán compartidos con las comunidades indígenas y otros interesados, promoviendo un diálogo abierto y constructivo.
2. Consentimiento y participación informada: aunque el estudio no implica la recopilación de datos de individuos, se promoverá la de representantes de las comunidades indígenas en la revisión y validación de los resultados. Esto asegurará que las voces y perspectivas de las comunidades sean escuchadas y respetadas.

3. Equidad y justicia: La investigación se comprometerá a abordar y mitigar cualquier forma de sesgo y discriminación en los sistemas de IA. Esto incluye un enfoque crítico en la identificación y rectificación de sesgos en los datos de entrenamiento y en las representaciones generadas.

4. Protección de datos: Se implementarán medidas estrictas de seguridad y protección de datos para garantizar la confidencialidad y privacidad de cualquier información sensible manejada durante el proyecto.

Esta investigación busca no solo identificar y analizar los sesgos presentes en las representaciones de la indigenidad generadas por IA, sino también promover un uso más ético y equitativo de estas tecnologías, respetando y valorando la diversidad cultural.

Actividades de divulgación:

- Presentación de ponencia en el XXXV Encuentro Nacional e internacional AMIC. Inteligencia Artificial y Procesos Comunicativos: desde las prácticas innovadoras, hasta las dimensiones éticas y de carácter legal (Puebla) donde se discutirán los avances y hallazgos preliminares del proyecto.
- Publicación de un capítulo académico en un libro digital sobre las representaciones de la indigenidad por IA.
- Desarrollo, desde la colaboración interdisciplinar, de estrategias y recomendaciones para mitigar estos sesgos en sistemas de inteligencia artificial.

Colaboraciones interdisciplinarias:

Para asegurar un enfoque integral y ético, este proyecto se desarrollará en colaboración con expertos en inteligencia artificial (para entender y mitigar los sesgos en los algoritmos y datos de entrenamiento) e investigadores en estudios culturales (para analizar las representaciones y su impacto en los imaginarios sociales).

- Eduardo Ulises Moya Sánchez. Director de Inteligencia Artificial del Gobierno de Jalisco.
- Renato Salomón Arroyo Duarte. Ingeniero de datos y coordinación de análisis estratégico en el Gobierno de Jalisco.

- Víctor Manuel Téllez Lozano. Profesor investigador de la Universidad de Guadalajara y coordinador de la Maestría en Estudios Mesoamericanos.

Referencias

- Ananya. (2024). AI image generators often give racist and sexist results: can they be fixed? *Nature*, 627(8005), 722-725.
<https://www.nature.com/articles/d41586-024-00674-9>
- Berg H, Hall S, Bhalgat Y, Kirk H, Shtedritski A, Bain M. (2022). A Prompt Array Keeps the Bias Away: Debiasing Vision-Language Models with Adversarial Learning. In: Proceedings of the Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the International Joint Conference on Natural Language Processing (AACL-IJCNLP); 2022. p. 806–822.
<https://doi.org/10.48550/arXiv.2203.11933>
- Bourdieu, P, Passeron, J. C. (2001) *La reproducción: elementos para una teoría del sistema de enseñanza*. España: Editorial Popular.
- Brack, M., Friedrich, F., Hintersdorf, D., Struppek, L., Schramowski, P., & Kersting, K. (2024). SEGA: Instructing text-to-image models using semantic guidance. *Advances in Neural Information Processing Systems*, 36.
<https://doi.org/10.48550/arXiv.2301.12247>
- Fairclough, N. (1989). *Language and power*. Longman Group UK limited.
- Fairclough, N. (1995). *Critical discourse analysis*. Edinburgh, Longman.
- Ferrante, E. (2021). Inteligencia artificial y sesgos algorítmicos ¿Por qué deberían importarnos?. *Nueva sociedad*, (294), 27-36.
<https://nuso.org/articulo/inteligencia-artificial-y-sesgos-algoritmicos/>
- Foucault, M. (1971). *L'ordre du discours*. Paris : Gallimard. 88pp.
- Jäger, Siegfried (2003). “Discurso y conocimiento: aspectos teóricos y metodológicos de la crítica del discurso y del análisis de dispositivos”. En Wodak, Ruth y Meyer, Michael. *Métodos de análisis crítico del discurso*. Gedisa editorial. Barcelona.
- Marañón, F., & Muñiz, C. (2012). Estereotipos mediáticos de los indígenas. Análisis de las representaciones en programas de ficción y entretenimiento de televisoras en Nuevo León. *Razón y Palabra*, (80), 45-23.

http://www.razonypalabra.org.mx/N/N80/V80/20_MaranonMuniz_V80.pdf

- Muñiz, E. (2021). Miradas encarnadas: Las nuevas formas de la discriminación racial. En J. Típa, S. Velasco Cruz, & U. Nuño Gutiérrez (Eds.), *Expresiones contemporáneas de los racismos en México: Cuerpos, medios y educación*. (pp. 11-34). Universidad de Guadalajara.
- Ouyang L, Wu J, Jiang X, Almeida D, Wainwright CL, Mishkin P, et al. Training language models to follow instructions with human feedback. (2022). En: Proceedings of the Advances in Neural Information Processing Systems: Annual Conference on Neural Information Processing Systems (NeurIPS).
- Pardo abril, N. G. (2012). Análisis crítico del discurso: Conceptualización y desarrollo. *Cuadernos de lingüística hispánica*, (19), 41-62. https://revistas.uptc.edu.co/index.php/linguistica_hispanica/article/view/447
- Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin.
- Said, E. W. (2023). Orientalism. In *Social theory re-wired* (pp. 362-374). Routledge.
- Spennemann, D. H. (2023). ChatGPT and the generation of digitally born “knowledge”: How does a generative AI language model interpret cultural heritage values?. *Knowledge*, 3(3), 480-512. <https://doi.org/10.3390/knowledge3030032>
- Spivak, G. C. (1988). Can the subaltern speak? En C. Nelson & L. Grossberg (Eds.), *Marxism and the Interpretation of Culture* (pp. 271-313). University of Illinois Press. Stecher.
- Terrones Rodríguez, A. (2022). Ética para la inteligencia artificial sostenible. *Arbor*, 198(806), a683-a683. <https://doi.org/10.3989/arbor.2022.806013>
- Típa, J. (2020). Estudios sobre la discriminación y el racismo en los contenidos de los medios de comunicación en México. *Revista Iberoamericana de Comunicación*, 38, 149-182.
- Típa, J., Velasco Cruz, S., & Nuño Gutiérrez, U. (Eds.). (2021). *Expresiones contemporáneas de los racismos en México: Cuerpos, medios y educación*. Universidad de Guadalajara.

<https://doi.org/10.1000/ric.2020.38.149-182>

- Van-Dijk, T. A. (2016). Análisis crítico del discurso. *Revista Austral de Ciencias Sociales*, (30), 203-222.
<https://doi.org/10.4206/rev.austral.cienc.soc.2016.n30-10>
- Van Dijk, T. (2013). Racismo cotidiano y política intercultural. *Manual para el diseño de políticas interculturales*, 114-120.
- Van Dijk, T. (1997). *Racismo y análisis críticos de los medios*. Paidós comunicación, Barcelona.
- Wang, A., Ramaswamy, V. V., & Russakovsky, O. (2022). Towards intersectionality in machine learning: Including more identities, handling underrepresentation, and performing evaluation. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 336-349).
<https://doi.org/10.1145/3531146.3533101>
- Weerts, H., Pfisterer, F., Feurer, M., Eggenberger, K., Bergman, E., Awad, N., ... & Hutter, F. (2024). Can fairness be automated? Guidelines and opportunities for fairness-aware AutoML. *Journal of Artificial Intelligence Research*, 79, 639-677.
<https://doi.org/10.1613/jair.1.12139>
- Wodak, R., Meyer, M., & Titscher, S. (2000). Methods of text and discourse analysis: In search of meaning. *Methods of Text and Discourse Analysis*, 1-288.
- Žižek, S. (1998). Multiculturalismo o la lógica cultural del capitalismo multinacional. *Estudios culturales. Reflexiones sobre el multiculturalismo*, 137-188.
- Zou, J., & Schiebinger, L. (2018). Design AI so that it's fair. *Nature*, 559(7714), 324-326.
- Zuñiga González, V. A. (1998). De cómo hablamos de los indígenas los mexicanos (no indígenas) en el México actual. *Región y sociedad*, 9(15).
<https://doi.org/10.22198/rys.1998.15.a833>

Referencias de las Declaraciones /Encuentros internacionales sobre IA y ética:

- Comisión Europea. (2019). Directrices éticas para una inteligencia artificial confiable. Bruselas, Bélgica. Recuperado de <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

- Foro Mundial sobre el Impacto Social de la Inteligencia Artificial. (2018). Declaración de Montreal para la responsabilidad ética global de la inteligencia artificial. Montreal, Canadá. Recuperado de <https://www.montrealdeclaration-responsibleai.com/>
- G7. (2019). Declaración de principios sobre la inteligencia artificial del G7. Biarritz, Francia. Recuperado de <https://www.consilium.europa.eu/en/press/press-releases/2019/08/26/g7-presidency-2019-france-agenda-and-outcomes/>
- Organización para la Cooperación y el Desarrollo Económicos. (2019). Recomendación del Consejo sobre Inteligencia Artificial. París, Francia. Recuperado de <https://www.oecd.org/going-digital/ai/principles/>
- UNESCO. (2019). La UNESCO adopta los principios éticos para la inteligencia artificial. París, Francia. Recuperado de <https://en.unesco.org/news/unesco-adoption-ethics-ai>
-