

La inteligencia artificial y sus aplicaciones en las humanidades.

Miércoles de 10 a 12 horas.

Responsables:

Dr. Ernesto Priani Saisó, Facultad de Filosofía y Letras.

Dr. Victor Mireles Chávez, Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades.

Presentación

En los últimos tres años la inteligencia artificial basada en grandes modelos del lenguaje ha probado ser una herramienta de comunicación y automatización de tareas muy poderosa, que hoy tiene usos tanto a pequeña como a gran escala. Por otro lado, las humanidades digitales (HD) es un campo multidisciplinar en el que se refleja la transformación de las disciplinas humanísticas a partir de la digitalización de la cultura y la vida cotidiana.

A partir de la premisa de que los objetos de las humanidades son ahora digitales, ya sea por que nacen digitales o porque se digitalizan, las humanidades digitales llevan más de tres décadas explorando la aplicación de herramientas computacionales para investigar estos nuevos objetos digitales textuales y no textuales que forman bibliotecas, archivos, museos, por mencionar solo los más evidentes.

La cátedra extraordinaria "La inteligencia artificial y sus aplicaciones en las humanidades" se propone abordar, como su nombre lo indica, el uso que la inteligencia artificial tiene hoy en día para la investigación humanística en todas sus disciplinas. Aunque se trata todavía de prácticas experimentales consideramos relevante que los alumnos las conozcan y la exploren críticamente, para conocer su capacidad, sus limitaciones. Ante todo la mira está puesta en que las personas asistentes descubran cómo aplicaciones concretas de IA permite explorar archivos, no sólo incrementa la capacidad de exportación cuando contamos con una gran cantidad de datos a explorar, sino que además permite hacerlo de maneras distintas a las desarrolladas por humanos.

Si bien partimos de que la IA es una tecnología epistémica de modelado de la cultura, y explicaremos cuáles son las implicaciones de esta definición, la cátedra está orientada principalmente a que los alumnos comprendan y utilicen herramientas de IA para sus investigaciones tanto con documentos textuales, auditivos o visuales.

Objetivo general

El alumno conocerá a través del curso las metodologías necesarias, así como las herramientas de inteligencia artificial para investigar en humanidades.

Objetivos específicos

- Reconocer los inteligencia artificial como una tecnología epistémica de modelado de la cultura.
- Descubrir como se construye un corpus y un archivo digital.
- Explorar herramientas de IA para transcripción, codicología en y producción de metadatos.
- Utilizar herramientas de IA para el análisis de texto, audio y video.

Temario

Sesión 1. Introducción al curso

Sesión 2. La Inteligencia artificial como tecnología epistémica de modelado de la cultura.

El curso partirá de confrontar dos ideas sobre qué tipo de tecnología es la inteligencia artificial. La que la define como tecnología epistémica, con las implicaciones que esto tiene, y como la define como una tecnología cultural, porque se forma a partir de los datos de nuestra cultura y, por lo mismo, ofrece un modelo de nuestra cultura.

Sesión 3. Corpus digital para aplicar herramientas de IA

Para comenzar a trabajar con herramientas de inteligencia artificial debemos configurar primero un corpus -textual, de audio, video- a partir de formular una pregunta que deseamos responder. Esta sesión está destinada a revisar cómo formamos un corpus.

Sesión 4. Del corpus al archivo.

Una vez que se ha formado un córpus digital tenemos que formar un archivo. En esta sesión exploraremos cómo hacerlo a partir de conocer qué tipo de repositorios digitales existen, cómo se utilizan y se crean ontologías para catalogar los documentos contenidos en el archivo.

Sesión 5. Del archivo a la web.

La siguiente tarea en el curso es convertir nuestro archivo digital en un sitio web dónde alojarlo. En la sesión veremos qué tipo de software existe para alojarlos y

consultarlos y cuáles son las mejores estrategias para hacerlo, teniendo en mente que lo utilizaremos para aplicar herramientas digitales.

Sesión 6. Revisión de avances y preparación del archivo.

La primera tarea de evaluación es tener un archivo digital sobre el cual se trabajará con las herramientas de inteligencia artificial. Dedicaremos la sesión a revisar avances e intercambiar ideas con los participantes.

Sesión 7. Descifrar el pasado de manera automática.

La primera herramienta a conocer y aplicar es Traskribus. Una aplicación de inteligencia artificial para la transcripción de textos antiguos y modernos de forma automática, que puede ser usada tanto con textos antiguos, como actuales.

Sesión 8. Descifrar obras con texto e imágenes usando IA.

Los documentos antiguos (como códices prehispánicos o manuscritos medievales) presentan una combinación de texto e imágenes. Estructurar dicha información permite la búsqueda rápida en acervos, el análisis cuantitativo de su contenido, y el empleo de técnicas simbólicas para su estudio. Se presentarán algunas herramientas para clasificar texto manuscrito, tipos de script y marginalia, así como para agrupar documentos según estilos, colores, y otras características.

Sesión 9. Una mirada a la biblioteca desde la IA.

La catalogación de acervos de acuerdo a los distintos esquemas de metadatos que se han desarrollado en los últimos 100 años, siempre ha representado una tarea intensiva, y los resultados siempre han estado sujetos a las idiosincrasias y tiempos de los catalogadores. Aquí se abordan ejemplos de como los desarrollos recientes en IA pueden ayudar en la catalogación, por medio del descubrimiento de metadatos en documentos textuales e imágenes, asistidos por metadatos de origen y por herramientas de conversión entre esquemas.

Sesión 10 Análisis de corpus: crear Grafos de conocimiento.

Existen representaciones formales del conocimiento que son más tratables por máquinas y por colectivos humanos que el texto. En esta sesión se abordará la creación de grafos (gráficas/redes) de conocimiento, que son representaciones de las relaciones entre entidades como personas, lugares o eventos. Los grados se pueden hacer semi-automáticamente a partir de un corpus de texto, permitiendo su subsecuente integración con otros grafos, y su procesamiento por medio de sistemas de razonamiento automático.

Sesión 11 Análisis de corpus: Identificar los temas más frecuentes.

Incluso seleccionando cuidadosamente un corpus, los textos en él contenidos tratarán una diversidad de temas. Esto es aún más relevante en corpora generales como aquellos provenientes de fuentes noticiosas. En esta sesión se abordarán técnicas para identificar los temas tratados en un corpus y para etiquetar segmentos de texto relevantes a cada uno, exponiendo a la vez la noción de descomposición en tópicos latentes.

Sesión 12 Análisis de corpus: identificación de los sentimientos

En muchas aplicaciones académicas e industriales es de interés identificar la disposición de las personas autoras de un texto con respecto a un tema. Las técnicas conocidas como “análisis de sentimiento”, aplicadas a documentos individuales, brindan una dimensión nueva para analizar cuantitativamente un corpus, y permiten vincular los documentos con el contexto sociopolítico en que se produjeron.

Sesión 13 Revisión de avances y uso de herramientas.

La segunda etapa de evaluación es aplicar las herramientas de IA al archivo digital previamente elaborado. Dedicaremos la sesión a revisar avances e intercambiar ideas con los participantes sobre la aplicación de las herramientas y sus resultados.

Sesión 14 IA para análisis de archivos de audio

Los paisajes sonoros, así como los discursos transmitidos oralmente, capturan gran cantidad de información que es susceptible de análisis y cuantificación. Esta sesión explorará técnicas diversas de conversión de audio a texto, de búsqueda y similitud sonora, y de extracción de información musical.

Sesión 15 IA para análisis de archivos de video

Los videos generados para propósitos periodísticos o documentales pueden contener información de interés para las ciencias sociales y las humanidades. En esta sesión se mostrarán algunos ejemplos de análisis de videos para ligarlos con otras fuentes documentales y complementar cuantitativamente su estudio.

Sesión 16 Conclusiones

Ultima sesión del curso donde se hará un recuento de lo visto, y un balance de las aplicaciones, y sus resultados.

Modelo pedagógico

A lo largo del curso se combinarán dos formas de enseñanza complementarias. Por un lado tendremos exposiciones en las sesiones conceptuales de la primera y segunda parte, y por otro funcionaremos a manera de taller, en las sesiones de la

tercera y cuarta parte, donde examinaremos aplicaciones y límites de la inteligencia artificial. La idea es que durante el curso no sólo se desarrollen reflexiones teórico-críticas, sino también aplicaciones prácticas de la herramienta para explorar sus posibilidades, características y limitaciones.

Formas de evaluación

Para acreditar el curso se deberá contar con el 80% de asistencias, y haber entregado tres trabajos para ser evaluados que serán los siguientes: elaboración de un corpus a partir de los materiales que el alumno quiera investigar. Aplicación de herramientas de IA para el análisis del corpus formado. Entrega de un documento con interpretación de los análisis y sus conclusiones.

Bibliografía

Mira Mira, José (2008). "Aspectos conceptuales de la Inteligencia Artificial y la Ingeniería del Conocimiento". En José Tomás Palma Méndez, Roque Marín Morales. *Inteligencia Artificial: Métodos, técnicas y aplicaciones*. Madrid: McGRAW-HILL.

Alvarado, R. (2023) "AI as an Epistemic Technology". *Sci Eng Ethics* 29, 32. <https://doi.org/10.1007/s11948-023-00451-3>

Drucker, J. (2021). *The digital humanities coursebook: an introduction to digital methods for research and scholarship*. Routledge.

Gopnik, Alison (2022, Julio 13). *Large Language Models as a Cultural Technology* (Video) YouTube <https://www.youtube.com/live/k7rPtFLH6yw>

Hawkins, A. (2022). Archives, linked data and the digital humanities: increasing access to digitised and born-digital archives via the semantic web. *Archival Science*, 22(3), 319-344.

Hayles, Katherine. (2025) *Bacteria to AI. Human Futures with our Nonhuman Symbionts*. Chicago University Press

Underwood, Ted (2024, Octubre 23) *Why AI Needs the Humanities as a Partner* . (Video) YouTube <https://www.youtube.com/watch?v=olUQ28tFdfM>

Herramientas

<https://www.transkribus.org/>

<https://omeka.org/>

<https://wordpress.com/es/>

<https://new.drupal.org/home>

<https://colab.research.google.com/>

<https://jena.apache.org/>

Coordinadores

Ernesto Priani Saisó

Ernesto Priani Saisó es filósofo y humanista digital. Profesor de tiempo completo en la Facultad de Filosofía y Letras de la UNAM. Sus intereses actuales son la filosofía del renacimiento, particularmente el renacimiento florentino, con la exploración de la utilización de dispositivos tecnológicos para la investigación en humanidades. Es director del proyecto de edición y laboratorio de investigación digital Biblioteca Digital del Pensamiento Novohispano. De manera reciente publicó *Los instrumentos de la noche. Fragmentos para una historia de lo soñado* y *Una Mirada al Libro electrónico*, escrito junto con Isabel Galina. Coordinó la traducción de las 900 tesis de Pico della Mirandola publicada por la UNAM, y ha escrito y publicado artículos sobre filosofía y humanidades digitales en los que destaca la entrada sobre Ramón Llull en la Stanford Encyclopedia of Philosophy. Es presidente de la Red de Humanidades Digitales, miembro de la Alianza de Organizaciones de Humanidades Digitales (ADHO) y del Virtual Humanities Lab de la Universidad de Brown.

Víctor Mireles

Víctor Mireles es Doctor en Ciencias por la Universidad Libre de Berlín, y Maestro y Licenciado en Ciencias de la Computación por la UNAM. Investigador titular del Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades (CEIICH) en temas injusticias epistémicas en la Inteligencia Artificial, Procesamiento del Lenguaje Natural y Grafos de Conocimiento. Ha participado en numerosos proyectos financiados por la Comisión Europea, la Agencia Austriaca de Investigación y el Conahcyt (hoy SECIHTI) y publicado más de 30 artículos en revistas y memorias de conferencias arbitradas, así como 3 capítulos de libro.

Perfiles de expositores

Miriam Peña. Confirmada.

Especialista en Humanidades Digitales y estudios literarios, con particular interés en el desarrollo de modelos para la publicación enriquecida y en Digital Scholarship (Pedagogía digital) para la formación de recursos humanos en las Humanidades Digitales. Es investigadora principal del proyecto Libros de Todos, Libros para Todos (LTxLT): Visibilizar las Colecciones Patrimoniales de la BNDM Usando Humanidades Digitales y del proyecto (por concluir) Bibliografía Mexicana y Humanidades Digitales Aplicadas, donde desarrolla protocolos para la preservación digital por medio de la publicación enriquecida y la remediación de objetos.

Isabel Galina. Confirmada.

Especialista en Humanidades Digitales, trabaja en los siguientes proyectos de investigación: El patrimonio bibliográfico digital. Criterios para su definición y manejo

y La preservación del patrimonio bibliográfico digital. Imparte las materias Automatización de bibliotecas (Biblioteca digital) en el posgrado en Bibliotecología y Estudios de la Información, así como Las edades del libro: historias, perspectivas e impacto en los estudios literarios en el posgrado de letras. Es miembro fundadora de la Red de Humanidades Digitales (RedHD) y del Grupo de Preservación Digital(GPD), además de ser editora de la serie Digital Research in the Arts and the Humanities (Routledge). Recientemente coordinó el proyecto de investigación Colecciones Digitales Inclusivas. Análisis Crítico de los Procesos de Creación de las Colecciones Digitales Mexicanas (PAPIIT IG400322).

Jonathan Giron. Confirmado.

Licenciado en Filosofía por la Universidad la Salle, diplomado en Edición Digital Académica por la Universidad Nacional de Educación a Distancia de España y maestro en Comunicación y Humanidades Digitales por la Universidad del Claustro de Sor Juana.Coordinador Editoria en el Instituto de Investigaciones sobre la Universidad y la Educación (IISUE) de la Universidad Nacional Autónoma de México (UNAM). Trabaja principalmente en edición digital, comunicación académica y humanidades digitales. Es secretario del Comité Ejecutivo de la Red de Humanidades Digitales, integrante del Grupo de Investigación de Escritos de Mujeres de la UNAM y del Programa Patrimonio/Publicaciones Digitales organizado por el IISUE y el Instituto de Investigaciones Bibliográficas.

Olly Ackerman

Es investigadora asociada del Departamento de Estudios Históricos y Culturales de la Universidad Libre de Berlín. Es especialista en manuscritos árabes e islam Shi'i, y ha estudiado la vida social de los acervos documentales en el sureste asiático. Sus publicaciones incluyen The Bohra Manuscript Treasury as a Sacred Site of Philology: a Study in Social Codicology y la monografía A Neo-Fatimid Treasury of Books in Gujarat. Arabic Manuscripts among the Alawi Bohras of South Asia

Elke Koppen

Elke Köppen es licenciada en Sociología por la Universidad de Bielefeld, Alemania y maestra y doctora en Bibliotecología y Estudios de la Información por la Universidad Nacional Autónoma de México. Es investigadora del Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades de la misma universidad donde es integrante del Programa de Investigación Estudios Visuales. Ha publicado artículos, capítulos de libro y libros coordinados sobre información visual, archivos fotográficos, imágenes científicas graffiti y fotografía.

Helena Gómez Adorno

La Dra Gómez Adorno es investigadora en el Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas de la UNAM. Tiene un doctorado en Ciencias de la Computación por el IPN. Sus intereses de investigación giran alrededor del procesamiento de lenguaje natural y el minado de textos. Ha trabajado en tareas como

similitud semántica, atribución de autoría, clasificación de texto, así como representaciones de texto basadas en gráficas.

Ximena Gutierrez

Es investigadora especializada en lingüística computacional e IA con un enfoque interdisciplinario para profundizar en el estudio del lenguaje humano. Sus líneas de investigación cubren el procesamiento del lenguaje natural en entornos multilingües, la lingüística cuantitativa, y el desarrollo de tecnologías para lenguas minorizadas de México. Ha realizado estudios en diversas universidades internacionales, recientemente fue investigadora posdoctoral en la Universidad de Zürich, Suiza, donde se especializó en aproximaciones para modelar la complejidad y la diversidad lingüística utilizando métodos estadísticos y computacionales. Actualmente se incorporó como investigadora asociada al Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades (CEIICH) de la UNAM, donde trabaja en la interfaz entre las humanidades y el área de la inteligencia artificial, lo anterior dentro del programa Macrodatos, Inteligencia Artificial e Internet.

Jorge Luis Perez Gonzalez

Doctor y Maestro en Ciencias en Ingeniería Biomédica por la Universidad Autónoma Metropolitana Unidad Iztapalapa. Actualmente es investigador asociado de tiempo completo en el Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas unidad Mérida, UNAM. Perteneció al Sistema Nacional de Investigadoras e Investigadores Nivel 1 y al Programa de Estímulos al Desempeño del Personal Académico de Tiempo Completo Nivel C. Ha efectuado estancias de investigación y colaboraciones con la Universidad Técnica de Munich Instituto Tecnológico de Massachusetts. Cuenta con más de 40 publicaciones conjuntando artículos científicos y memorias internacionales. Sus líneas de investigación son procesamiento digital de imágenes y señales, reconocimiento de patrones y aprendizaje computacional multimodal.

Domingo Meri

Domingo Mery es Profesor Titular del Departamento de Informática de la UC. Obtuvo el Diploma (M.Sc.) en Ingeniería Eléctrica por la Universidad Técnica de Karlsruhe, Alemania, y el Doctorado con distinción en la Universidad Técnica de Berlín. Actualmente es Profesor Titular del Departamento de Ciencias de la Computación de la Pontificia Universidad Católica de Chile (UC), Chile. Se ha desempeñado en puestos de investigación y dirección en diversas instituciones en Estados Unidos, Alemania y Chile. Sus intereses de investigación incluyen el procesamiento de imágenes para la detección de fallos en piezas de fundición de aluminio, imágenes de rayos X, programación en tiempo real y visión por computador. Es autor de más de 90 publicaciones técnicas SCI y más de 100 ponencias en congresos.