

# La **IA** se suma al arsenal de los historiadores

para  
comprender  
el pasado

**MIT  
Technology  
Review**

Publicado por Opinno



MOIRA DONOVAN  
03 MAYO, 2023

Los historiadores del mañana utilizarán la informática para analizar cómo vivía la gente siglos atrás

**E**ra una tarde del año 1531 en Venecia. En el taller de un impresor, un aprendiz se afanaba en el diseño de una página destinada a ser parte de un libro de texto de astronomía. En la hoja, una densa línea de caracteres y una ilustración xilográfica de una cabeza de querubín observando formas nadando por el cosmos que representa un eclipse lunar.

Como todos los aspectos de la **producción literaria del siglo XVI**, era un proceso que llevaba mucho tiempo, pero que permitía que el conocimiento se difundiera con una rapidez sin precedentes.

500 años después, la producción de información es una bestia completamente diferente. Los **terabytes de imágenes, vídeo y texto en torrentes de datos digitales** circulan casi al instante y tienen que analizarse con la misma rapidez, permitiendo -y requiriendo- el entrenamiento de modelos de aprendizaje automático para ordenar ese flujo. Este cambio en la

producción de información tiene implicaciones para el futuro de todo, desde la creación artística hasta el desarrollo de fármacos.

Sin embargo, estos avances también permiten **analizar de forma diferente los datos del pasado**. Los historiadores ya han empezado a utilizar el aprendizaje automático, las redes neuronales profundas en particular, para examinar documentos históricos. Entre las que se incluyen las tablas astronómicas como las producidas en Venecia y otras ciudades de principios de la Edad Moderna, que han sido emborronadas tras pasar siglos en archivos emmohecidos o distorsionadas por el desliz de la mano de un impresor.

Los historiadores afirman que la aplicación de la informática actual al pasado más remoto ayuda a establecer conexiones en una franja más amplia de registro histórico. Sería imposible de otro modo, ya que corrige las **distorsiones que se producen al analizar la historia documento a documento**. Y también introduce sus propias distorsiones, como el riesgo de que el aprendizaje automático introduzca sesgos o falsificaciones en el registro histórico. Todo esto plantea un interrogante a los historiadores y otros especialistas que, como se suele decir, entienden el presente examinando la historia. Ya que las máquinas van a desempeñar un papel más importante en el futuro, ¿cuánto pasado debemos cederles?

#### LA COMPLEJIDAD DEL ANÁLISIS SINTÁCTICO

Los macrodatos han llegado a las humanidades a través de iniciativas para digitalizar un número cada vez mayor de documentos históricos, como la colección de millones de páginas de periódicos de la Biblioteca del Congreso (EE UU) y los registros judiciales de los archivos finlandeses, que se remontan al siglo XIX. Para los investigadores, es tanto un problema como una oportunidad: **hay más información, y hasta ahora no ha habido forma de cribarla**.

Este reto se ha superado con el desarrollo de herramientas informáticas, que ayudan a los estudiosos a analizar esta complejidad. En el año 2009, **Johannes Preiser-Kapeller**, profesor de la Academia Austriaca de Ciencias, estaba examinando un registro de decisiones de la Iglesia

los terabytes de imágenes, vídeo y texto en torrentes de datos digitales circulan casi al instante y tienen que analizarse con la misma rapidez, permitiendo -y requiriendo- el entrenamiento de modelos de aprendizaje automático para ordenar ese flujo.



bizantina del siglo XIV. Al darse cuenta de que, para dar sentido a cientos de documentos haría falta un estudio digital sistemático de las relaciones entre obispos, Preiser-Kapeller creó una base de datos de personas y utilizó un programa informático de análisis de redes para reconstruir sus conexiones.

Esta reconstrucción reveló **patrones de influencia encubiertos**, lo que llevó a Preiser-Kapeller a argumentar que los obispos que más hablaban en las reuniones no eran los más influyentes. Desde entonces, ha aplicado la técnica a otras redes, incluida la élite bizantina del siglo XIV, descubriendo formas en las que su tejido social se sostenía gracias a **las contribuciones ocultas de las mujeres**. «Pudimos identificar lo que ocurría fuera de la narrativa oficial, hasta cierto punto», apunta.

El trabajo de Preiser-Kapeller es otro ejemplo más de esta tendencia. Sin embargo, hasta hace poco, el aprendizaje automático a menudo era incapaz de extraer conclusiones de grandes colecciones de texto. Entre otras cosas porque ciertos aspectos de los documentos históricos (en el caso de Preiser-Kapeller, el griego mal manuscrito) los hacían **indescifrables para las máquinas**. Ahora, los avances en el aprendizaje profundo han empezado a abordar estas limitaciones, utilizando redes que imitan el cerebro humano para detectar patrones en conjuntos de datos grandes y complejos.

Hace casi 800 años, **Johannes de Sacrobosco**, astrónomo del siglo XIII, publicó el *Tractatus de sphaera*. Este era un tratado introductorio sobre el cosmos geocéntrico, que se convirtió en lectura obligatoria para los estudiantes universitarios de la Edad Moderna. Fue el libro de texto más difundido sobre cosmología geocéntrica, y perduró incluso después de que la revolución copernicana pusiera patas arriba la visión geocéntrica del cosmos en el siglo XVI.

El tratado también es la estrella de una colección digitalizada de **359 libros de texto de astronomía**, publicados entre **1472 y 1650: 76.000 páginas** que incluyen decenas de miles de ilustraciones científicas y tablas astronómicas. Matteo Valleriani, profesor del Instituto Max Planck de Historia de la Ciencia (Alemania), vio en este amplio conjunto de datos la oportunidad de trazar la evolución del



Ilustración digital hecha con IA en Adobe Firefly Beta.

**todo esto plantea una pregunta a los historiadores: Ya que las máquinas van a desempeñar un papel más importante en el futuro, ¿cuánto pasado debemos cederles?**

conocimiento europeo hacia una cosmovisión científica común. Pero también se dio cuenta de que discernir el patrón requería algo más que capacidades humanas. Así que Valleriani y un equipo de investigadores del Instituto de Fundamentos del Aprendizaje y los Datos de Berlín (**BIFOLD**, por sus siglas en alemán) recurrieron al aprendizaje automático.

Para ello, fue necesario dividir la colección en tres categorías: partes de texto (secciones de escritos sobre un tema concreto, con un principio y un final claros), **ilustraciones** científicas (que ayudaban a iluminar conceptos, como un eclipse lunar), y **tablas numéricas** (utilizadas para enseñar aspectos matemáticos de la astronomía).

Valleriani cuenta que, al principio, el texto desafiaba la interpretación algorítmica. Por un lado, la tipografía variaba mucho, ya que las imprentas de la Edad Moderna desarrollaban tipos únicos para sus libros y solían tener sus propios talleres metalúrgicos para fundir las letras. Esto significaba que un modelo que utilizara el **procesamiento del lenguaje natural** (NLP, por sus siglas en inglés) para leer el texto tendría que re-entrenarse para cada libro.

El idioma también planteaba un problema. Muchos textos estaban escritos en dialectos latinos específicos de cada región, a menudo irreconocibles para máquinas que no han sido entrenadas en lenguas históricas. «Cuando no se tiene el vocabulario para entrenar en segundo plano, es una gran limitación para el procesamiento del lenguaje natural», cuenta Valleriani. Esta es una de las razones por las que el PLN **funciona bien con lenguas dominantes como el inglés, pero es menos eficaz, por ejemplo, con el hebreo antiguo**.

En cambio, los investigadores extrajeron manualmente el texto de los materiales originales e identificaron enlaces individuales entre conjuntos de documentos. Por ejemplo, cuando un texto era imitado o traducido en otro libro. Estos datos se colocaron en un gráfico, que incorporó automáticamente esos enlaces individuales en una red que contenía todos los registros. Después, los investigadores utilizaron un gráfico para entrenar un método de aprendizaje automático que puede sugerir conexiones entre textos. Por último, quedaban los elementos visuales de los textos: **20.000 ilustraciones y 10.000 tablas**,

que los investigadores estudiaron mediante redes neuronales.

## EN EL PRESENTE

La visión por ordenador de imágenes históricas se enfrenta a retos similares a los de la PLN. Lauren Tilton, profesora asociada de Humanidades Digitales en la Universidad de Richmond (Virginia, EE UU), lo denomina sesgo “actualista”. Muchos modelos de IA se entrenan con conjuntos de datos de los últimos 15 años, explica Tilton, y los objetos que han aprendido a enumerar e identificar tienden a ser características de la vida actuales, como móviles o coches. A menudo, los ordenadores solo reconocen iteraciones contemporáneas de objetos que tienen una historia más larga, como los **iPhones y los Teslas, en lugar de centralitas y modelos T**. Por si fuera poco, los modelos suelen entrenarse con imágenes en color de alta resolución, no con las granuladas fotografías en blanco y negro del pasado, o con las primeras representaciones modernas del cosmos, de aspecto incoherente y degradadas por el paso del tiempo. Todo ello hace que la visión por ordenador sea menos precisa cuando se aplica a imágenes históricas.

«Hablaremos con informáticos y nos dirán: ‘Ya hemos resuelto la detección de objetos’», afirma Tilton, y añade: «Y nosotros les diremos que, en realidad, **si coges un conjunto de fotografías de los años 30, verás que no está tan resuelto como pensamos**». Los modelos de aprendizaje profundo, que identifican patrones en grandes cantidades de datos, pueden ayudar porque son capaces de una mayor abstracción.

En el caso del proyecto **Sphaeraproyecto**, los investigadores de BIFOLD entrenaron una red neuronal para detectar, clasificar y agrupar (en función de la similitud) ilustraciones de textos modernos tempranos. Ahora, ese modelo es accesible para otros historiadores a través de la web pública **CorDeep**. Y también adoptaron un enfoque innovador para analizar otros datos. Por ejemplo, no podían compararse visualmente las distintas tablas de los cientos de libros de la colección porque **«una misma tabla puede imprimirse de mil maneras distintas»**, explica Valleriani. Así que los investigadores desarrollaron una arquitectura de red neuronal que detecta

y agrupa tablas similares en función de los números que contienen, e ignorando su diseño.

Hasta ahora, el proyecto ha arrojado algunos resultados sorprendentes. Uno de los patrones hallados en los datos permitió a los investigadores ver que, mientras Europa se fracturaba por motivos religiosos tras la Reforma protestante, el conocimiento científico se unificaba. Los textos científicos que se imprimían en lugares como **Wittenberg**, la ciudad protestante que se había convertido en un centro de innovación académica gracias al trabajo de los eruditos reformistas, eran imitados en centros neurálgicos como París y Venecia, antes de extenderse por todo el continente. La Reforma protestante es un tema bastante estudiado, afirma Valleriani, pero una perspectiva mediada por máquinas permitió a los investigadores ver algo nuevo: «Antes, esto no estaba nada claro». Los modelos aplicados a tablas e imágenes han empezado a devolver patrones similares.

Estas herramientas ofrecen posibilidades más significativas que el simple seguimiento de 10.000 tablas, afirma Valleriani. En su lugar, permiten a los investigadores extraer inferencias sobre la evolución del conocimiento a partir de patrones en registros grupales. Aunque, en realidad, solo hayan examinado un puñado de documentos. «Con mirar dos tablas, ya puedo sacar una gran conclusión sobre 200 años», apunta.

Las redes neuronales profundas también están desempeñando un papel en el examen de una historia aún más antigua. Descifrar inscripciones (conocido como epigrafía) y restaurar ejemplares dañados son tareas minuciosas, sobre todo, cuando los objetos inscritos se han trasladado o carecen de pistas contextuales. Por tanto, los historiadores especializados tienen que sacar conjeturas. Para ayudarles, **Yannis Assael**, investigador científico de DeepMind (Reino Unido); y **Thea Sommerschildt**, becaria posdoctoral de la Universidad Ca’ Foscari de Venecia (Italia), desarrollaron una red neuronal llamada Ithaca. Esta puede reconstruir partes perdidas de inscripciones, además de atribuir fechas y ubicaciones a los textos. Los investigadores afirman que este método de aprendizaje profundo, entrenado con un conjunto de datos de más de 78.000 inscripciones, es el primero que aborda la restauración y la atribución conjuntamente a través del aprendizaje a partir de grandes cantidades de datos.

**a menudo, los ordenadores solo reconocen iteraciones contemporáneas de objetos que tienen una historia más larga: como los iPhones y Teslas, en lugar de centralitas y Modelos T.**



Ilustración digital hecha con IA en Adobe Firefly Beta.



Simulación digital de la Venecia medieval.  
Ilustración digital hecha con IA en Adobe Firefly Beta.

la IA generativa  
podría utilizarse  
para hacer  
predicciones  
que rellenen los  
espacios en blanco  
de los registros  
históricos.

Hasta ahora, según Assael y Sommerschild, el método está arrojando luz sobre las inscripciones de los decretos de un importante periodo de la Atenas clásica. Durante mucho tiempo, estos se han atribuido entre los años **446 y 445 a.C.**, un periodo que algunos historiadores han puesto en duda. Como prueba, los investigadores entrenaron el modelo con un conjunto de datos que no contenía la inscripción en cuestión, luego le pidieron que analizara el texto de los decretos, y el resultado ofreció una datación diferente. «La fecha media predicha por Ítaca para los decretos es el año **421 a.C.**, lo que concuerda con los avances más recientes en datación y muestra cómo el aprendizaje automático puede contribuir a los debates en torno a uno de los momentos más significativos de la historia griega», señalaron por correo electrónico.

### MÁQUINAS DEL TIEMPO

Otros proyectos proponen utilizar el aprendizaje automático para extraer inferencias aún más amplias sobre el pasado. Esta fue la motivación de la Máquina del Tiempo de Venecia, una de las varias *máquinas del tiempo* locales que se han creado en toda Europa para reconstruir la historia local, a partir de registros digitalizados. Los archivos estatales venecianos abarcan 1.000 años de historia repartidos en 80 kilómetros de estanterías. Y el objetivo de los investigadores era digitalizar estos registros, muchos nunca habían sido examinados por historiadores modernos. Utilizarían redes de aprendizaje profundo para extraer información y, rastreando nombres que se repiten en otros documentos, reconstruir los lazos que unían a los venecianos.

Frédéric Kaplan, presidente de la Organización Máquina del Tiempo (TMO, por sus siglas en inglés), afirma que el proyecto ya ha digitalizado suficientes documentos administrativos de la ciudad como para captar la textura de Venecia en siglos pasados. Esto permite ir edificio por edificio, e identificar a las familias que vivieron allí en distintos momentos. «Son cientos de miles de documentos que hay que digitalizar para alcanzar este tipo de flexibilidad. Esto nunca se había hecho antes», explica Kaplan.

Sin embargo, cuando se trata de la última promesa del proyecto, una simulación digital

de la Venecia medieval hasta el nivel de barrio, mediante redes reconstruidas por inteligencia artificial, los historiadores como Johannes Preiser-Kapeller, profesor de la Academia Austriaca de Ciencias que dirigió el estudio de los obispos bizantinos, afirman que el proyecto no ha sido capaz de cumplir porque el modelo no puede entender qué conexiones son significativas.

Preiser-Kapeller ha hecho su propio experimento utilizando la detección automática para desarrollar redes a partir de documentos: extrae información de la red con un algoritmo, en lugar de que un experto extraiga información para alimentar la red, como en su trabajo sobre los obispos. Y asegura que produce mucha «complejidad artificial», pero nada que sirva para la interpretación histórica. El algoritmo fue incapaz de distinguir los casos donde los nombres de dos personas aparecían en el mismo padrón de contribuyentes o figuraban en un certificado de matrimonio, así que, como dice Preiser-Kapeller, **«lo que se obtiene no tiene valor explicativo»**. Es una limitación que los historiadores han puesto de relieve con el aprendizaje automático, similar al que la gente ha señalado sobre los grandes modelos lingüísticos como **ChatGPT**: en última instancia, los modelos no entienden lo que leen y pueden llegar a conclusiones absurdas.

Kaplan afirma que, con las fuentes disponibles en la actualidad, es necesaria la interpretación humana para proporcionar contexto, aunque esto podría cambiar una vez que un número suficiente de documentos históricos sean legibles por máquinas.

No obstante, imagina una aplicación del aprendizaje automático más transformadora y potencialmente problemática. La IA generativa podría utilizarse para hacer predicciones que rellenen los espacios en blanco de los registros históricos, por ejemplo, sobre el número de aprendices en el taller de un artesano veneciano. Y a partir de datos agregados, no registros individuales, que podrían ser inexactos o incompletos. Esto puede aportar **perspectivas no elitistas, pero va en contra de la práctica histórica habitual**, donde las conclusiones se basan en las pruebas disponibles.

Sin embargo, una preocupación más inmediata la plantean las **redes neuronales**, que crean registros falsos.

## ¿ES REAL?

En YouTube, los espectadores ahora pueden ver a **Richard Nixon pronunciar un discurso que había sido escrito por si el alunizaje de 1969 acababa en desastre**, pero que nunca tuvo que pronunciarse. Los investigadores crearon el deepfake para mostrar cómo la IA podría afectar a nuestro sentido compartido de la historia. En cuestión de segundos, se pueden generar imágenes falsas de grandes acontecimientos históricos, por ejemplo, el desembarco del **Día D**. Tal y como Dan Cohen, profesor de historia de **Northeastern**, comentó recientemente junto a los alumnos de una clase dedicada a explorar el modo en que los medios digitales y la tecnología están dando forma al estudio histórico: «Las fotos son convincentes. Puedes meter a un montón de gente en una playa con un tanque y una ametralladora, y parece perfecto».

Cohen señala, como ejemplo, la forma en que **Joseph Stalin** ordenó borrar a sus enemigos de los libros de historia. Falsar la historia no es nada nuevo, pero la escala y la velocidad con que pueden crearse falsificaciones son impresionantes, y el problema va más allá de las imágenes. La IA generativa puede crear textos que parezcan discursos parlamentarios de la época victoriana, como ha hecho Cohen con sus alumnos. Al generar manuscritos o tipografías históricas, también podría crear lo que parece un convincente registro histórico escrito.

Mientras tanto, chatbots de IA como **Character.ai** e **Historical Figures Chat** permiten a los usuarios simular interacciones con personajes históricos. Los estudiosos han expresado su preocupación por estos chatbots que pueden, por ejemplo, hacer que algunos individuos parezcan **menos racistas y más arrepentidos de lo que fueron**.

En otras palabras, existe el riesgo de que la inteligencia artificial, desde los chatbots históricos a los modelos que hacen predicciones basadas en registros históricos, se equivoque mucho. Algunos de estos errores son anacronismos benignos: una consulta a **Aristóteles** en el chat Character.ai sobre su opinión sobre **las mujeres** (a las que veía como inferiores) devolvió como respuesta que **«no deberían tener redes sociales»**. Pero otros pueden ser más graves, sobre todo, cuando se mezclan en una colección

de documentos demasiado grande para que un historiador pueda revisarlos individualmente; o si los hace circular alguien interesado en una interpretación particular de la historia.

Aunque no haya un engaño deliberado, a algunos académicos les preocupa que los historiadores utilicen herramientas para las que no están formados. «Existe un gran riesgo: **como humanistas o historiadores estamos subtratando el análisis a otro campo, o quizá a una máquina**», afirma Abraham Gibson, profesor de Historia de la Universidad de Texas (San Antonio, EE UU). Gibson afirma que, hasta hace poco, los historiadores con los que hablaba no veían la importancia de la inteligencia artificial para su trabajo; pero se han dado cada vez más cuenta de que podrían llegar a ceder parte de la interpretación de la historia a una caja negra.

Este problema no es exclusivo de la historia: incluso los desarrolladores de sistemas de aprendizaje automático a veces tienen dificultades para entender cómo funcionan. Afortunadamente, algunos de los métodos diseñados están estructurados pensando en los historiadores para ofrecer mayor transparencia. **Ithaca** produce una serie de hipótesis clasificadas por probabilidad, y los investigadores de BIFOLD están trabajando en la interpretación de sus modelos con IA explicable, que pretende revelar qué entradas contribuyen más a las predicciones. Los historiadores aseguran que ellos mismos promueven la transparencia animando a la gente a ver el aprendizaje automático con desapego crítico: como una herramienta útil pero falible, igual que las personas.

## LOS HISTORIADORES DEL MAÑANA

Aunque persiste el escepticismo hacia esta nueva tecnología, el campo de estudio la está adoptando gradualmente, y Valleriani cree que con **el tiempo disminuirá el número de historiadores que rechacen los métodos digitales**. En su opinión, la preocupación de los académicos por la ética de la IA no es tanto una razón para no utilizar el aprendizaje automático como una oportunidad para que las humanidades contribuyan a su desarrollo.

Emmanuel Le Roy Ladurie, historiador francés, escribió en el año 1968 una respuesta al trabajo de los historiadores que habían empezado a experimentar con la historia computacional para investigar cuestiones como los patrones de votación del Parlamento británico en la década de 1840: **«El historiador del mañana será un programador, o no existirá»**. </>



Moirá Donovan es una periodista científica independiente afincada en Halifax, Nueva Escocia.

El artículo original «La IA se suma al arsenal de los historiadores para comprender el pasado» pertenece a la edición digital de *MIT Technology Review*.

Los contenidos bajo el sello *MIT Technology Review* están protegidos enteramente por copyright. Ningún material puede ser reimpresso parcial o totalmente sin autorización.

Si quisiera syndicar el contenido de la revista *MIT Technology Review*, por favor contáctenos.

E-mail: [redaccion@technologyreview.com](mailto:redaccion@technologyreview.com)

Tel: +34 911 284 864