

Inteligencia artificial y datos masivos en archivos digitales sonoros y audiovisuales

Perla Olivia Rodríguez Reséndiz
Coordinadora



La presente obra está bajo una licencia de:
<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>



Atribución-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0)

Este es un resumen legible por humanos (y no un sustituto) de la [licencia](#). [Advertencia](#).

Usted es libre de:

Compartir — copiar y redistribuir el material en cualquier medio o formato

Adaptar — remezclar, transformar y construir a partir del material

La licenciante no puede revocar estas libertades en tanto usted siga los términos de la licencia

Bajo los siguientes términos:



Atribución — Usted debe dar [crédito de manera adecuada](#), brindar un enlace a la licencia, e [indicar si se han realizado cambios](#). Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciante.



NoComercial — Usted no puede hacer uso del material con [propósitos comerciales](#).



CompartirIgual — Si remezcla, transforma o crea a partir del material, debe distribuir su contribución bajo la [misma licencia](#) del original.

Inteligencia artificial y datos masivos
en archivos digitales sonoros
y audiovisuales

COLECCIÓN:
TECNOLOGÍAS DE LA INFORMACIÓN
INSTITUTO DE INVESTIGACIONES BIBLIOTECOLÓGICAS Y DE LA INFORMACIÓN

Inteligencia artificial y datos masivos
en archivos digitales sonoros
y audiovisuales

Coordinadora
Perla Olivia Rodríguez Reséndiz



Universidad Nacional Autónoma de México
2020

Investigación realizada gracias al programa DGAPA - PAPIIT IT400118

Q335 Inteligencia artificial y datos masivos en archivos digitales
I57 sonoros y audiovisuales / Coordinadora Perla Olivia Rodríguez
Reséndiz. - México: UNAM. Instituto de Investigaciones
Bibliotecológicas y de la Información, 2020.

xviii, 182 p. - (Tecnologías de la información)

ISBN: 978-607-30-3829-4

Investigación realizada gracias al programa

DGAPA - PAPIIT IT400118.

1. Inteligencia artificial - Procesamiento de datos. 2. Internet
de las cosas. 3. Archivos sonoros. 4. Big data. I. Rodríguez
Reséndiz, Perla Olivia, coordinadora. II. ser.

Diseño de portada: Oscar Fernando Arcos Casañas

Imágenes:

Envato Elements

(<https://elements.envato.com/es-419/>)

Primera edición, 2020

D.R. © UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Ciudad Universitaria, 04510, México D.F.

Impreso y hecho en México

ISBN: 978-607-30-3829-4

Publicación dictaminada

Contenido

PRESENTACIÓN

| | |
|-------------------------------------------------------------------------------------------------------------------------|----|
| APROXIMACIONES A LA INTELIGENCIA ARTIFICIAL Y A LOS DATOS MASIVOS EN ARCHIVOS DIGITALES SONOROS Y AUDIOVISUALES..... | ix |
| Perla Olivia Rodríguez Reséndiz | |

PRINCIPIOS CONCEPTUALES

| | |
|-----------------------------------------------------------------|----|
| EL PAPEL DE LA INTELIGENCIA ARTIFICIAL EN LA INDUSTRIA 4.0..... | 21 |
| Juan Humberto Sossa Azuel | |

| | |
|----------------------------------------------------------------------------------------------------|----|
| ONTOLOGÍA Y NARRATIVA EN LOS SISTEMAS DE INTELIGENCIA ARTIFICIAL, UNA MIRADA ANTROPOLÓGICA..... | 59 |
| Lily Díaz-Kommonen | |

| | |
|---------------------------------------------|----|
| BIG DATA BETWEEN PRIVACY AND COPYRIGHT..... | 77 |
| Amelia Bongarzone | |

| | |
|----------------------------------------------------------------------------------------|----|
| UNA ESTRATEGIA PARA GESTIONAR LOS BIG DATA EN ARCHIVOS SONOROS Y AUDIOVISUALES..... | 89 |
| Antonella Bongarzone | |

EXPERIENCIAS DE USO Y APLICACIÓN

| | |
|-------------------------------------------------------------------------------------------|-----|
| USING COMPUTATIONAL TOOLS AND EXPERTS TO IMPROVE ACCESS TO DIGITAL MEDIA ARCHIVES..... | 101 |
| Karen Cariani y David O Ives | |

| | |
|--------------------------------------------------------------------------------------------|-----|
| REFLEXIONES EN TORNO A LA INTELIGENCIA ARTIFICIAL: EL CASO DE LOS ARCHIVOS DE RTVE..... | 117 |
| Virginia Bazán-Gil | |

| | |
|-------------------------------------------------------------|-----|
| CATALOGACIÓN AUTOMÁTICA: MÁS CERCA QUE LEJOS..... | 133 |
| Alejandro Álvaro Ramírez Acosta y Juan Carlos Murillo Pabón | |

| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| APUNTES SOBRE LA INTELIGENCIA ARTIFICIAL Y SU APLICACIÓN EN LOS ARCHIVOS SONOROS..... | 149 |
| Georgina Sanabria Medina | |
| IDENTIFYING METADATA ELEMENTS IN PHOTOGRAPHIC REPOSITORIES BY SEMANTIC SEGMENTATION OF IMAGES WITH DEEP LEARNING..... | 159 |
| Sinaí López-Castillo, Isnardo Reducindo y Francisco Benita | |
| PROPUESTA METODOLÓGICA Y DE ANÁLISIS COMPUTACIONAL PARA IDENTIFICAR EL PROCESO FOTOGRÁFICO EN FOTOGRAFÍAS HISTÓRICAS DEL SIGLO XIX Y XX..... | 171 |
| Gustavo Lozano San Juan y Rodrigo Colín Rivera | |

Aproximaciones a la inteligencia artificial y a los datos masivos en archivos digitales sonoros y audiovisuales

PERLA OLIVIA RODRÍGUEZ RESÉNDIZ
Universidad Nacional Autónoma de México

Desde la invención de la tecnología de la información, a principios del siglo XIX, nuestra capacidad de generación de datos nunca había sido tan intensa y constante. Cada año se produce más información digital que el anterior. El uso y manejo de grandes volúmenes de información digital es un signo contemporáneo.

Datos masivos, en inglés *big data*, es el término contemporáneo que alude a la identificación, procesamiento y recuperación de grandes volúmenes de información gracias al uso de la Inteligencia Artificial (IA) (Kusnetzky, 2014).

La extracción de información y conocimiento para desarrollar otras tareas es uno de los retos más importantes de la sociedad (Xindong Wu et al., 2014). De ahí que la aplicación de la IA en grandes volúmenes de información sea considerada una herramienta necesaria para la sistematización y análisis de datos.

La IA tiene una amplia gama de aplicaciones para enormes cantidades de datos. Por ejemplo, la presentación de proyectos como el algoritmo DeepMind de Google (DeepMind, 2019) que aprendió cómo ganar 49 juegos de Atari; así como la posibilidad de que los algoritmos puedan reconocer el lenguaje y patrones escritos a mano, describir la información de fotos y videos. Incluso, que el contenido de noticias se genere automáticamente. También se evidencian prometedoras aplicaciones en la salud al reducir los costos de la investigación clínica y ayudar a médicos y pacientes en la mejor toma de decisiones.

Además, en la biología y en la genética se observan soluciones de relevancia social. De hecho, la era ómica, es concebida como el periodo histórico en el que “se da una visión global de los procesos biológicos basada en el análisis de un gran volumen de datos... para estudiar organismos que son ahora desconocidos, así como sus funciones, todo a través de su rastro genético” (Monleon-Getino 2016, 433).

Además, la relevancia de los datos masivos y de la inteligencia artificial en la economía mundial es incuestionable. Para 2020, se ha estimado que el crecimiento de la industria de los datos a nivel mundial será de 739 mil millones de euros, que equivalen al 4% del producto interno bruto (PIB) y a 10 millones de empleos, directos e indirectos, que dependerán de este sector (Pellicer y Fariza, 2019). Los datos se han convertido en un activo clave para la economía (Monleon-Getino, 2016).

Por ello, la resistencia a la dependencia de empresas como Amazon, Google o Microsoft, en materia de almacenamiento y gestión de grandes volúmenes de datos ocupa la agenda política a nivel internacional. A finales del 2019, la canciller alemana Ángela Merkel declaró “muchas compañías han subcontratado [la gestión, el almacenamiento y el procesamiento de] sus datos a empresas estadounidenses... no digo que sea malo, sólo quiero decir que los productos de valor añadido que proceden de ahí con la ayuda de la inteligencia artificial crearán dependencias que no estoy segura de que sean buenas” (Pellicer y Fariza 2019, s/p). En el Parlamento Europeo se ha señalado que los datos y la inteligencia artificial son elementos para la innovación y para resolver problemas sociales, desde la salud, agricultura, hasta la seguridad. Por ello, se asignaron 20 millones de euros anuales para el desarrollo de la inteligencia artificial (Pellicer y Fariza, 2019).

La Unión Europea diseñó una estrategia sobre *big data* para apoyar y acelerar la transición hacia una economía basada en los datos en el espacio europeo. La economía basada en datos estimulará la investigación y la innovación en general, mientras que lleva a más oportunidades de negocio y a un aumento de la disponibilidad de los conocimientos y el capital, en particular para las pequeñas y medianas empresas (PYME) (Monleon-Getino 2016, 431).

Es sabido que la manipulación en el procesamiento de grandes volúmenes de datos de ciudadanos puede incidir en el manejo de la

toma de decisiones de la sociedad y poner a la democracia en riesgo (Helbing et al. 2018). También se ha evidenciado que en el diseño de herramientas de IA pueden incidir perspectivas subjetivas individuales o de grupo que excluyan a personas por raza, creencias, preferencias o bien afiliaciones ideológicas. Este es, sin lugar a dudas, un ámbito de investigación que debe ser abordado con mayor detalle para establecer los fundamentos éticos del uso de la IA en grandes volúmenes de contenidos digitales.

Estamos ante una revolución de los datos fenómeno que comprende el movimiento de los datos abiertos, el incremento en la colaboración pública (*crowdsourcing*), las nuevas tecnologías de la información, la inteligencia artificial y el internet de las cosas que están transformando a la sociedad (ONU, 2017) y no podemos tener una actitud pasiva, es necesario formular una perspectiva crítica del uso masivo de los datos a través de herramientas de IA.

ARCHIVOS DIGITALES

La incorporación de la tecnología digital en los archivos sonoros y audiovisuales modificó la noción de documento analógico a ítem, objeto y documento digital. Asimismo, transformó en ciclo de vida digital a los procesos documentales físicos e intelectuales, a través de los cuales se preservaron durante décadas los soportes analógicos. Además, confirió dinamismo y vitalidad a las instituciones de la memoria destinadas al resguardo de este tipo de colecciones. Por ello, la curaduría, como nuevo proceso intelectual se ha instaurado en los archivos para dar valor y visibilidad a los materiales que durante años sólo se conservaban. Con ello, la generación de nuevos contenidos y servicios de información que no habían sido imaginados ocupó el interés de los archivistas.

También, se dio origen a la construcción de archivos digitales, que pueden ser definidos como complejos sistemas de información que preservan cuantiosos volúmenes de contenidos digitales. Los archivos digitales, son un término polivalente que expresa la tecnología, un conjunto de procesos, técnicas y tecnologías empleadas en la preservación digital y además, puede ser considerada la institución de

la memoria que tiene como misión la salvaguarda de objetos digitales. Además, de forma coloquial se nombra como archivo digital a los registros digitales unitarios que se realizan en diversos formatos y lenguajes.

A pesar de las amplias posibilidades que ofrece el contexto digital para los archivos digitales, existen desafíos aún no resueltos. Algunos son derivados de la incorporación de nuevas técnicas y tecnologías digitales y otros, heredados del trabajo con soportes analógicos.

Uno de estos problemas es la continua acumulación de contenidos digitales (media y metadata) producto tanto de la digitalización como de la producción de documentos de origen digital.

Lo que significa que digitalizar no es preservar. La digitalización es el medio para transferir contenidos de un soporte analógico a uno digital. La preservación digital es una tarea continua que no se debe interrumpir; basada en una visión sustentable y de largo plazo. Por ello, una vez que se digitalizan los contenidos requieren ser migrados de forma sistemática, como parte de un plan de preservación digital sustentable.

A esta problemática se suma la constante y creciente producción de contenidos sonoros y audiovisuales de origen digital. Una gran parte de este tipo de materiales desaparece, por desconocimiento de los métodos adecuados de preservación y porque, en algunos casos, se producen en formatos de compresión y propietarios, condiciones que no son recomendadas para garantizar su permanencia a largo plazo. Por ello, reiteradamente se ha advertido que los materiales de origen digital afrontan un alto riesgo de pérdida, mayor incluso que el de los soportes analógicos, cuyos contenidos no han sido digitalizados.

De todos los tipos de materiales de información, las grabaciones sonoras y audiovisuales son las que requieren mayor espacio de almacenamiento digital. Aun cuando los soportes para el almacenamiento masivo digital evolucionan como resultado de la inercia del mercado y, cada vez se pueden conservar mayores volúmenes de contenidos digitales, su costo y rápida obsolescencia tecnológica limitan su adquisición. Sólo los archivos que cuentan con políticas de preservación digital sustentable y cuyos recursos económicos están garantizados a largo plazo pueden instalar robustos sistemas de almacenamiento masivo digital con visión de largo plazo.

El volumen de los documentos digitales sonoros y audiovisuales que deben ser preservados supera los límites imaginados por archivistas y documentalistas sonoros y audiovisuales. La producción de los materiales de origen digital es constante y exponencial. El tratamiento documental demanda de saberes, técnicas y tecnologías para preservar cuantiosas cantidades de contenidos digitales sonoros y audiovisuales. Más allá de las limitaciones económicas y tecnológicas, el crecimiento en el volumen de contenidos digitales trae consigo otro tipo de problemas asociados al ciclo de vida digital y a la permanencia del contenido digital.

INTELIGENCIA ARTIFICIAL Y DATOS MASIVOS (*BIG DATA*)

De 2003 a 2008, se evidenció el crecimiento de los centros de datos en el mundo. En 2011, se reportaron 509,147 centros de datos en el mundo (Miller, R. 2011). En esta proyección se incluyen los centros de datos que preservan contenidos sonoros y audiovisuales. La incorporación del big data y de la inteligencia artificial en los archivos sonoros y audiovisuales es reciente si se considera que las primeras aplicaciones de la IA comenzaron en los años cincuenta del siglo pasado y el crecimiento del volumen de contenidos digitales es un rasgo, como se ha señalado, que caracteriza a la era contemporánea.

Los archivos sonoros y audiovisuales que han digitalizado sus colecciones o están llevando a cabo este proceso, acogen estas herramientas como un medio para resolver problemas asociados a la identificación y a la recuperación de la información. Debido al tiempo y complejidad que demanda el proceso de catalogación sonora y audiovisual, en muchos casos, sólo se cuenta con los datos que se pueden incorporar en la creación de inventarios. La catalogación es una tarea compleja porque difícilmente un archivista o documentalista puede asentar con precisión la amplia gama de materias que se resguardan en los archivos. Además, el grado de error en el asentamiento de los datos persiste a través de faltas de redacción y de ortografía, entre otras. Frente a este panorama, el uso de herramientas para identificar, transcribir y recuperar contenidos de forma automatizada, significa una alternativa en las tareas de preservación digital que deben afrontar los archivos.

En los primeros años del siglo XXI, se han puesto en marcha las primeras aproximaciones teóricas y desarrollos tecnológicos de la IA en grandes volúmenes de datos digitales que se resguardan en archivos sonoros y audiovisuales. Los resultados de las experiencias pioneras son muy valiosos. Como sucedió con las primeras experiencias de digitalización de archivos, el aprendizaje basado en el ensayo-error ha sido una pieza fundamental en el desarrollo y asentamiento de la digitalización como proceso tecnológico para transferir contenidos grabados en soportes analógicos a digitales.

Los científicos de las ciencias de la información, bibliotecología y archivología tienen, en el estudio e investigación de la inteligencia artificial y el big data, un ámbito emergente de conocimiento. Es necesario determinar los efectos sociales y culturales del big data y la inteligencia artificial. Además, es necesario conocer el impacto que estas tecnologías tendrán en las instituciones de la memoria que cada día incrementan el volumen de ítems digitales en bibliotecas, archivos y museos.

En este sentido, el libro *Inteligencia artificial y datos masivos* en archivos digitales sonoros y audiovisuales ofrece discusiones teóricas sobre el tema y presenta estudios de caso, que dan cuenta de los derroteros de los archivos sonoros y audiovisuales que han sido pioneros en la utilización de la IA en grandes volúmenes de datos digitales.

El capítulo “El papel de la inteligencia artificial en la Industria 4.0”, escrito por el Dr. Humberto Sossa, presenta el marco conceptual y el estado del arte de este ámbito del conocimiento que, de acuerdo con el autor, será detonante de la quinta revolución industrial, de la economía cíclica. El trabajo sitúa el desarrollo histórico de la IA a nivel internacional y ubica la posición de México, lo que constituye una aportación significativa si se considera que alentar la investigación científica en este campo incide en el desarrollo tecnológico, en la innovación y en la generación de nuevos mercados.

A continuación, la doctora Lily Díaz presenta, desde la perspectiva antropológica, el capítulo “Ontología y narrativa en los sistemas de inteligencia artificial, una mirada antropológica”. Ofrece una reflexión teórica en relación con los grandes volúmenes de datos y su recuperación e interpretación a través de la inteligencia artificial. Esta visión establece una primera discusión en torno a la relación del big

data y la inteligencia artificial que señala la necesidad de contar con métodos que permitan trabajar de forma creativa con datos de forma responsable e inclusiva.

El uso de la IA en grandes volúmenes de datos es, sin lugar a dudas, una fuente de progreso e innovación; no obstante, al mismo tiempo establece aspectos jurídicos que es necesario considerar. En el apartado “Big Data Between Privacy and Copyright” de Amelia Bongarzone, se exploran, desde una perspectiva jurídica, las posibilidades que ofrece el big data para elaborar modelos interpretativos, analíticos e incluso predictivos de acontecimientos y comportamientos humanos, difíciles de imaginar hasta hace poco, y que llaman la atención del derecho privado y suscitan la reflexión sobre las implicaciones que pueden afectar a la sociedad. La visión jurídica va más allá del ámbito patrimonial.

Antonella Bongarzone en “Una estrategia para gestionar los big data en archivos sonoros y audiovisuales”, introduce el tratamiento de documentos dinámicos, versus el estático. Establece que la catalogación, con base en las normas internacionales establecidas, es portadora de información sobre el contenido sonoro o visual. Más allá de este propósito también se reflejan el contexto histórico, sociológico, antropológico y cultural en que se produce dicho registro. Bajo esta consideración, el uso potencial de del big data, desde la perspectiva de las ciencias sociales, ofrece la posibilidad de ampliar el análisis y profundizar las descripciones superficiales con datos profundos del contexto del documento que se derivan precisamente del contexto histórico. La autora se centra en el significado y el valor del dato para que los grandes volúmenes de información no se conviertan en un mar incontrolado, sino que sean verdaderos vehículos de conocimiento.

En el capítulo “Using Computational Tools and Experts to Improve Access to Digital Media Archives”, escrito por Karen Cariani y David O Ives, se exponen las posibilidades de uso del aprendizaje automático, en inglés Machine Learning (ML) y la inteligencia artificial, en el American Archive of Public Broadcasting (AAPB), de Estados Unidos. El trabajo expone a partir de un estudio de caso la necesaria colaboración que debe existir entre las instituciones que preservan grandes cantidades de datos digitales sonoros y audiovisuales y los científicos que desarrollan y necesitan probar herramientas para el manejo de

grandes volúmenes de datos. Se sostiene la necesaria colaboración para mejorar la IA y el ML para obtener más y mejor calidad en los datos que se resguardan en archivos y bibliotecas.

Virginia Bazán-Gil ofrece la experiencia de la Radio y Televisión Española (RTVE) en la incorporación de la inteligencia artificial como una herramienta para dar visibilidad a los archivos. El capítulo “Reflexiones en torno a la inteligencia artificial: el caso de los archivos de RTVE”, presenta la experiencia de la incorporación de la IA para la generación automática de metadatos a través de la transcripción automática de voz a texto, así como de la segmentación y agrupación por hablantes y caras para el uso de contenidos en la producción de RTVE. El texto es una lección, necesaria para los archivos de radio y televisión, que, una vez concluido el proceso de digitalización de sus colecciones, afrontan el desafío de ensanchar las posibilidades de acceso. El apartado es muy claro y aleccionador, describe los alcances y limitaciones de la IA en un medio público de radio y televisión. Expresa cómo la incorporación de este tipo de tecnologías no debe ser una tarea aislada, sino resultado de la sinergia entre las diferentes áreas del medio.

En el apartado “Catalogación automática: más cerca que lejos”, escrito por Alejandro Álvaro Ramírez Acosta y Juan Carlos Murillo Pabón, se presentan los avances del Laboratorio de Radio e Inteligencia Artificial de Radio Televisión Nacional de Colombia (RTVC), que trabaja con los documentos del archivo sonoro de Señal Memoria. La aplicación de la IA en la experiencia colombiana está destinada a apoyar el proceso de catalogación de los archivos.

La evolución tecnológica es el punto de partida del capítulo “Apuntes sobre la Inteligencia Artificial y su aplicación en los archivos sonoros” de Georgina Sanabria. En este trabajo se determina, desde una perspectiva histórica, la evolución de la IA y sus posibles aplicaciones en los archivos sonoros. La autora destaca que, aun cuando el término IA es utilizado de forma recurrente en entornos comerciales, no se pueden omitir las posibilidades que ofrece esta tecnología en la búsqueda de soluciones a problemas sociales, uno de los cuales compete a la preservación de la memoria sonora de la humanidad.

La complejidad que implica la descripción de grandes volúmenes

de información afecta también a los archivos de fotografías. Esta problemática recurrente a los archivos sonoros y audiovisuales, impacta también en archivos de imagen fija como son los fotográficos. De ahí que en el capítulo “Identifying Metadata Elements in Photographic Repositories by Semantic Segmentation of Images with Deep Learning”, escrito por Sinaí López-Castillo, Isnardo Reducindo y Francisco Benita, se presente un método para la descripción automatizada de fotografías a través del uso de algoritmos de Deep Learning (DL).

El último capítulo “Propuesta metodológica y de análisis computacional para identificar el proceso fotográfico en fotografías históricas del siglo XIX y XX”, de la autoría de Gustavo Lozano y Rodrigo Colín Rivera, se ofrece una metodología para identificar a partir del uso de la IA la técnica fotográfica. Con ello, se busca proponer una herramienta a disposición de archivistas para la catalogación de fotografías.

Este libro es un trabajo pionero en el uso de la inteligencia artificial y los datos masivos en archivos digitales sonoros y audiovisuales, su lectura puede contribuir a comprender los alcances actuales y futuros de estas tecnologías en la preservación digital sustentable.

REFERENCIAS

DeepMind. (2019). <https://deepmind.com/>

Kusnetzky, D. (2014). *What is «Big Data?»* ZDNet. <https://www.zdnet.com/article/what-is-big-data/>.

Monleon-Getino, A. (2016). *El impacto del Big-data en la Sociedad de la Información. Significado y utilidad*. *Historia y Comunicación Social*, 20(2), 137-155. https://doi.org/10.5209/rev_HICS.2015.v20.n2.51392.

Pellicer y Fariza. (2019). *Europa ultima un plan para dar la batalla en el negocio de los datos* | Economía | El País. https://elpais.com/economia/2019/11/16/actualidad/1573926886_318836.html.

Xindong Wu, Xingquan Zhu, Gong-Qing Wu, & Wei Ding. (2014). Data Mining with Big Data. *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, vol. 26, núm. 1, JANUARY 2014.

PRINCIPIOS CONCEPTUALES

El papel de la inteligencia artificial en la Industria 4.0

JUAN HUMBERTO SOSSA AZUELA
Instituto Politécnico Nacional

INTRODUCCIÓN

La Inteligencia Artificial (IA) es, hoy en día, una disciplina con un nivel de madurez muy alto. No se puede poner en tela de duda su potencial, aplicabilidad e impacto en nuestra sociedad. Prácticamente todos en nuestras casas, oficinas de trabajo, negocios, escuelas, etcétera, sin darnos cuenta, utilizamos algún dispositivo “inteligente” que nos hace la vida más fácil, extiende nuestras capacidades, nos libera de actividades engorrosas, nos da seguridad. En resumen, nos brinda una mejor calidad de vida.

Los avances logrados hasta el momento, en esta área del conocimiento, permiten que un gran número de personas, hogares y empresas, cuenten con sistemas capaces de auxiliarlos en la toma de decisiones atinadas, y en situaciones que antes hubiera sido muy difícil, incluso imposible de lograr. Entre otros, podemos referirnos a los sistemas aconsejadores, los llamados chatbots, etcétera. En ambos casos, se trata de un tipo particular de robot de servicio, no necesariamente electromecánico, que permite seleccionar, por ejemplo, el mejor lugar para comer, la película a ver en la siguiente semana, el próximo libro a comprar; todo esto de acuerdo con los gustos personales, o simplemente completar un trámite en alguna entidad del gobierno.

En el caso particular de los robots asistenciales, se trata de máquinas con capacidades especiales para poder interactuar con los seres

humanos, atender algunas de sus necesidades y, por ejemplo, auxiliar a una persona con una incapacidad, en su alimentación, baño, etcétera.

En las fábricas, los robots inteligentes cooperantes (*cobots*), auxilian a los seres humanos en la realización de las tareas difíciles con menor riesgo y con una eficiencia y eficacia que no sería posible lograr con únicamente grupos de personas.

En materia de vigilancia y seguridad, grupos de robots, aéreos, submarinos y terrestres, ya monitorean los accesos, las acciones y actividades que pudieran ocurrir en espacios cerrados y abiertos. Un uso inmediato de los robots inteligentes voladores podría ser para rastrear el robo de combustible.

La IA, en muchos países es un campo de investigación y aplicación con bastante avance: China y Estados Unidos, por ejemplo. Esta tecnología coadyuva al desarrollo sustentable de estos países. Constituye un espacio de grandes oportunidades para el desarrollo científico, pero sobre todo tecnológico e innovador, para la apertura de nuevos mercados y de nuevos negocios.

La IA ha transitado por diferentes concepciones e intereses. En sus primeros años (la década de los cincuenta), la investigación en IA pretendía reproducir procesos de inteligencia humana para dar soluciones a problemas formulados en ambientes controlados, por ejemplo, en prueba de teoremas y la solución de juegos estratégicos con reglas previamente definidas.

En los últimos años la IA se ha expandido y diversificado hasta ser considerada como una rama de las ciencias orientada a la creación de máquinas inteligentes, con habilidades para aprender, adaptarse y actuar con autonomía. Además, el nivel de madurez en IA le confiere ya la calidad de una ingeniería, igual como sucedió, en su momento, con la Electrónica, que al principio era un campo de investigación de los físicos. Con el devenir del tiempo se transformó en una ingeniería, lo que coadyuvó en la creación de carreras tan solicitadas y útiles como la de Ingeniero en Electrónica e Ingeniero en Comunicaciones y Electrónica.

Los problemas que actualmente la IA aborda están relacionados con el manejo de grandes volúmenes de datos para derivar información útil que permita la generación automática de algoritmos y, con ello, la

solución de problemas complejos asociados al razonamiento, la percepción, la planeación, el aprendizaje y la habilidad de manipular objetos.

La IA se ha convertido en un campo multidisciplinario, con un amplio potencial para generar beneficios para la sociedad. Se requiere articular grupos de investigación interdisciplinaria que reúnan a investigadores, técnicos académicos y estudiantes, capaces de desarrollar fundamentos y metodologías orientadas a la modelación de fenómenos y a la resolución de problemas sociales, en campos tan vastos como la medicina, la seguridad, la movilidad y el transporte, la previsión y la gestión de desastres naturales y la formación de capital humano, para beneficio de todos los sectores, público, social y privado.

En lo general, las acciones encaminadas a potenciar el desarrollo y uso de la IA se llevan a cabo de manera individualizada, descoordinada y desarticulada. Se cuenta con pocos programas de licenciatura y de posgrado especializados en IA, aún así, estos planes de estudio permitirían formar los recursos humanos altamente capacitados que el mercado laboral requiere actualmente. Existe pocos consorcios o clústeres que aglutinen, coordinen y, por tanto, potencialicen, los esfuerzos individuales logrados hasta el momento.

Los avances alcanzados hasta ahora, apuntan en varias direcciones, muchas veces opuestas, porque si bien, los resultados obtenidos son buenos, no tienen el impacto social y económico deseado. Es necesario poner en operación más programas de licenciatura, posgrado, redes y consorcios, que potencien la asimilación y el uso de esta tecnología. En nuestro país, acciones de este tipo, permitirían identificar y localizar a los grupos que llevan a cabo labores en este campo, y a partir de ahí, organizarlos para definir nichos de oportunidad y coordinar eventos relacionados, a fin de atraer fondos y, lo más importante, proponer soluciones adecuadas a quienes lo demanden.

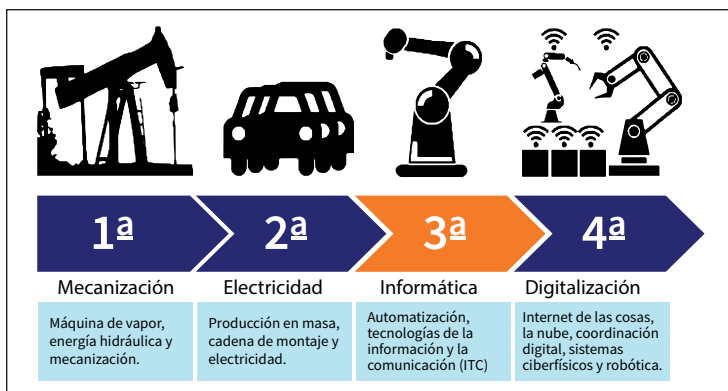
Para resolver los problemas complejos de los sectores de la sociedad es necesario transitar de, llevar a cabo investigaciones aisladas realizadas por investigadores con acceso a infraestructura limitada con pertinencia regional pero poco impacto nacional, a grupos con altos desempeños y masas críticas, de pertinencia regional e impacto nacional, y cuyos investigadores tengan acceso a infraestructura y redes de clase mundial.

En este capítulo, después de hacer mención de varios temas, como el estado del arte relativo a la IA, algunas definiciones del término, los tipos de IA en estudio actualmente, los enfoques para la construcción de máquinas inteligentes, los momentos clave y un breve recuento histórico de este fulgurante tema de investigación, se muestran las potenciales aplicaciones de esta tecnología, además de explicar cómo está siendo usada en el mundo y en nuestro país. También se enlistan algunos ejemplos ilustrativos y se abordan los retos para su implementación, así como, algunas consideraciones para las empresas y la ética en su uso. Después, se habla un poco del porqué, hoy por hoy, es el tiempo de la IA y de los países líderes en su desarrollo y uso. Antes de terminar se presenta un conjunto de datos útiles y una reflexión. Finalmente, se formulan las conclusiones.

ESTADO DEL ARTE

La IA, actualmente, es una tecnología que está influyendo fuertemente en los puestos de trabajo de prácticamente todas las industrias. Durante siglos se ha visto a la automatización como algo maligno que amenaza nuestros empleos y altera el statu quo. Esto ha sido notorio, particularmente, durante la primera, segunda y tercera revoluciones industriales (Figura 1).

Figura 1. Etapas o revoluciones por las cuales ha avanzado el proceso de la automatización.



La introducción de la automatización en todas estas revoluciones ha impulsado a los trabajadores a aprender nuevas habilidades, para adaptarse a las nuevas circunstancias. La automatización sigue creciendo y generando beneficios para las empresas en todo el mundo. Esto, hoy, se hace patente en la llamada Revolución 4.0 o Industria 4.0, donde de manera natural se conjuntan las tecnologías: 1) Internet de las cosas (IoT) y sistemas ciberfísicos, 2) grandes datos, minería de datos y analítica de datos, 3) simulación y fabricación aditiva (impresión 3D), 4) sistemas de integración horizontal y vertical, 5) ciberseguridad, 6) realidad virtual y la realidad aumentada, 7) cómputo en la nube y 8) robótica autónoma y colaborativa (*cobots*).

Sin duda, la automatización cambiará el lugar de trabajo de muchas maneras. Hará obsoletos algunos trabajos, hará que aparezcan otros. Para obtener un trabajo de tecnología, un aspirante debe desarrollar un nivel sofisticado de habilidades en temas difíciles.

En esta nueva revolución industrial surge un nuevo actor: la inteligencia artificial (IA). Ésta puede ayudar mucho, entre otras cosas, facilitar el proceso de aprendizaje y simplificar tareas computacionales al quitar parte de la carga a su homólogo humano.

La IA es, sin duda, una de las tendencias definitorias de nuestros días. En los últimos 10 años, diversos tipos de máquinas (programas y robots) han sido entrenadas para resolver problemas cada vez más complejos. Hoy, las máquinas son capaces de llevar a cabo tareas solo atribuibles, en otros tiempos, al ser humano, por ejemplo, en la identificación de personas en medio de la multitud [Co kun, 2017; Syafeeza, 2014; Wang, 2020], el guiado autónomo de autos [Davis, 2018; Gadam, 2018] y el intervenir en juegos complejos como el ajedrez, el Go, el Shogi [Ross, 2018], el póker [Brown, 2017]. StarCraft II [Jones, 2019], entre otros. Estos ejemplos bastan para demostrar que algunas veces las máquinas hacen mejor las cosas que nosotros los seres humanos.

Las máquinas pensantes no son nuevas por supuesto, datan de hace más de 75 años. Entre los programas más antiguos se encuentran el juego de damas, diseñado por Christopher Strachey [Galvan, 2019], y el de ajedrez, de Dietrich Prinz [Jodra, 2015]. Ambos, desarrollados en la computadora Ferranti Mark 1 de la Universidad de Manchester en 1951.

Actualmente son enunciadas de manera continua las siguientes preguntas: ¿sistemas basados en IA nos desplazarán de nuestros trabajos?, ¿sistemas basados en IA nos sobrepasarán en actividades recreativas como los juegos y en otras creativas que por muchos años han dado significado a nuestras vidas? Como ejemplo, está la referencia del pintor basado en IA que inventa nuevos estilos de arte [Marti, 2017].

Figuras públicas como Stephen Hawking y Elon Musk, en su momento, fueron muy lejos al dejar ver al mundo un posible IA apocalíptico, donde máquinas súper inteligentes pudieran pisotearnos siguiendo sus insondables metas [Musk, 2017].

¿QUÉ ES LA INTELIGENCIA ARTIFICIAL?

Antes de definir el término inteligencia artificial, se debe primero hablar acerca de lo qué es la inteligencia natural o simplemente la inteligencia, término vago, cuya definición ha dependido del individuo involucrado en un determinado momento, incluso, de las circunstancias.

Platón, por ejemplo, decía que la inteligencia es la actividad que permite adquirir ciencia, lo cual, por supuesto, es muy restrictivo. Él propuso dos tipos de inteligencia, mientras que Howard Gardner hace referencia de nueve [Garner, 2018], y la psicología de doce [Muenste, 2020]. Se puede decir que la inteligencia es una capacidad mental muy general que implica: 1) habilidad para razonar, 2) planificar, 3) resolver problemas, 4) pensar de forma abstracta, 5) comprender ideas complejas, 6) aprender con rapidez, y 7) aprender de la experiencia. No supone el mero aprendizaje de un texto, una habilidad académica específica, o resolver un test de forma habilidosa. Refleja una capacidad amplia y profunda para comprender el entorno y ser capaz de capturar el significado de las cosas y darles un sentido, o ingeniárselas, para en un momento dado, saber qué hacer.

Con base en lo anterior, la inteligencia artificial pudiera ser definida como la ciencia e ingeniería de las máquinas que actúan de manera inteligente. En este sentido, una máquina es inteligente cuando es capaz de tomar decisiones apropiadas en circunstancias inciertas.

Otra manera de ver lo mismo, es cuando una máquina –que se dice ser inteligente– es capaz de aprender a mejorar su comportamiento con base en sus experiencias.

TIPOS DE IA

De acuerdo con estudios como los referenciados [Bostrom, 2016; Godoy, 2020; López 2019, Perez, 2019], se pueden identificar cinco tipos de IA:

Inteligencia General (IG) o Inteligencia Fuerte. Llamada también Inteligencia Artificial General (IAG) (comparada con la inteligencia humana), N. Bostrom. Consiste en enseñarle a la máquina a aprender, a razonar y a planificar. No cubre todo el campo de la inteligencia, faltaría agregar a la máquina una mente, así como una conciencia, lo que implicaría estar consciente y dotarla de un carácter propio, lo cual no es fácil de simular. Una posible solución para resolver este problema sería hacer ingeniería inversa del cerebro. Muchos piensan que un 90% de una inteligencia comparable con la inteligencia humana (inteligencia artificial con nivel humano, IANH) podría alcanzarse durante este siglo, aproximadamente en el año 2075.

Súper Inteligencia Artificial (SIA). Un paso más allá es la llamada súper inteligencia artificial, que podría lograrse relativamente rápido, una vez que la IAG haya sido alcanzada, supuestamente porque los mayores obstáculos han sido sobrellevados.

Inteligencia Artificial Débil (IAD) o aplicada. Se limita a afrontar tareas específicas, enfocadas a ayudar al ser humano. No intenta simular el rango completo de las habilidades cognitivas humanas. Este es el tipo de IA que actualmente se está desarrollando en los ámbitos académico y empresarial.

Después de la IAG y la SIA se encuentran otros dos tipos de inteligencia: **La Súper Inteligencia-Súper Consciente IA (SI-SC-IA)**, y aquella proveniente de los seres de nuestra imaginación.

ENFOQUES PARA LA CONSTRUCCIÓN DE MÁQUINAS INTELIGENTES

En la construcción de máquinas inteligentes se habla de dos enfoques: el de arriba hacia abajo, y el de abajo hacia arriba.

En el marco del enfoque de arriba hacia abajo (también llamado formal), el diseño de una máquina inteligente involucra la solución de los siguientes dos problemas: 1) el del Reconocimiento de Patrones (RP), y 2) el del Sentido Común (SC). Si, por ejemplo, se quiere diseñar y poner en operación un robot que sirva como mesero en un restaurante, éste deberá ser capaz de reconocer patrones, por ejemplo, palabras, objetos, personas, etcétera.

El segundo problema, el del sentido común, es todavía más fundamental. Una persona, por ejemplo, sabe lo siguiente: 1) el agua está mojada, 2) las madres son más viejas que las hijas, 3) a los animales no les gusta el dolor, 4) uno no regresa después de la muerte, y 5) el tiempo no va hacia atrás. Pero, ¿no hay todavía matemáticas que expresen estos hechos!

Sabemos que cualquier niño aprende de su realidad. Las máquinas apenas empiezan a experimentar esto. En general hacen lo que se les ha programado de antemano. Ha habido proyectos para poner en operación un sistema que tome en cuenta todas las leyes del sentido común. Fue el proyecto CYC, que inició en 1984 por Douglas Lenat, de la empresa Cycorp. El proyecto CYC fue el equivalente al proyecto Manhattan. El objetivo del proyecto CYC era alcanzar el momento en el cual una máquina, por sí misma, fuera capaz de digerir más información al leer revistas y libros. Se predijo que en 10 años CYC sería capaz de alcanzar un 50% de conciencia. Hoy se sabe que CYC estuvo muy lejos pese al gran número de líneas de código escritas. La última versión de CYC abarca unos 47,000 conceptos y 506,000 hechos. El intentar producir un programa que incluya las leyes del sentido común es aparentemente inalcanzable, ya que hay demasiadas leyes. Es bien sabido que los seres humanos lo hacemos sin esfuerzo [Cope-land, Lenat, 1990].

Para los robots es muy difícil decir cuando una puerta está abierta y cuando lo está una ventana. Esto no impide, sin embargo, que haya grandes proyectos en esta dirección, por ejemplo, el auto autónomo.

En resumen, el enfoque de arriba hacia abajo presenta muchas limitaciones para la construcción de robots. Con base en ello, surge el llamado enfoque de abajo hacia arriba, cuyos seguidores tratan de imitar la manera en la que los bebés aprenden.

Uno de los promotores más reconocidos del enfoque (abajo-arriba) es el profesor Rodney Brooks del MIT [Brooks, 1989]. En lugar de usar programas de cómputo elaborados para calcular las posiciones precisas, los “insectoides” de Brooks, se valen de prueba y error para coordinar los movimientos de sus patas con mucho menos poder de cómputo. Genghis, por ejemplo, es capaz de navegar sin la ayuda de un sistema de cómputo centralizado. Varios de los diseños del Dr. Brooks son usados por los robots de la NASA para recolectar y analizar materiales en el planeta rojo. Otro proyecto de Brooks es COG, a través del cual intenta crear un robot con la inteligencia de un niño de seis años. A COG no se le programó ninguna inteligencia. Usa sus ojos para enfocar a un humano que le enseña con ejemplos.

Uno de los ingredientes que, muchos piensan, debiera incluir una máquina son las emociones. De hecho, muchos científicos consideran que una máquina requiere experimentar emociones para que sea eficiente.

Ejemplos de emociones que una máquina pudiera experimentar: querer algo, celos, vergüenza, remordimiento y soledad. Es muy probable que los robots avanzados vengán equipados con emociones. Incluso, hay quienes piensan que el sentir podría ser un puente crucial entre la inteligencia y la conciencia.

Uno puede preguntarse por qué las emociones son importantes. Sabemos que en el caso de los seres humanos, las emociones guían la toma de decisiones. Las máquinas del futuro, muy probablemente necesitarán de emociones para fijarse metas y así tomar decisiones, de otra forma, quedarían paralizadas ante un mar de posibilidades.

MOMENTOS CLAVE QUE HAN DADO FORMA A LA IA

Históricamente, la IA ha pasado por muchos momentos importantes que le han ido dando la forma que ahora tiene. Los siguientes ocho son los más importantes [Simonite, 2018]:

1. 1956. El proyecto de investigación de Dartmouth en Inteligencia Artificial, que tuvo lugar en el verano, y donde se acuñó el nombre de un nuevo campo relacionado con el quehacer de *software* habilidoso, como los humanos.
2. 1965. Joseph Weisenbaum, en el MIT, crea Eliza, el primer chatbot útil como psicoterapeuta.
3. 1975. En Stanford se desarrolla META-DENDRAL, *software* útil para interpretar análisis químicos, el cual es publicado en una revista.
4. 1987. Por primera vez una Van Mercedes, equipada con dos cámaras y varias computadoras, navega sola en terreno alemán, unos 20 km a más de 55 mph, proyecto dirigido por el ingeniero Ernst Dickmanns.
5. 1997. La computadora Deep Blue desarrollada por IBM derrota al campeón mundial de ajedrez, Garry Kasparov.
6. 2004. El Pentágono establece el proyecto DARPA Grand Challenge, una carrera para autos autónomos que da pie a esta industria.
7. 2012. Los investigadores encuentran un nicho en lo que se conoce como aprendizaje profundo, lo que abre un interés muy fuerte en IA y muestra que sus ideas pueden hacer más preciso el reconocimiento del habla y de las imágenes.
8. 2014. AlphaGo, creado por Google y DeepMind, derrota al campeón del mundo del entonces muy difícil juego del Go.

BREVE HISTORIA DE LA IA

Un recuento histórico de la IA, con los hechos más relevantes [NAT-GEO, 2019]:

- 384-322 a. C.: Las ideas más básicas se remontan a Aristóteles (384-322 a. C.). Fue el primero en describir un conjunto de reglas que describen una parte del funcionamiento de la mente para obtener conclusiones racionales.
- 250 a. C.: Ctesibio de Alejandría (250 a. C.) construye la primera máquina auto-controlada a través de un regulador del flujo de agua (racional, pero sin razonamiento).

- 1315: Ramon Llull en su libro *Ars magna generalis et ultima* tiene la idea de que el razonamiento podía ser efectuado de manera artificial.
- 1936: Alan Turing diseña formalmente la Máquina Universal que demuestra la viabilidad de un dispositivo físico para implementar cualquier cómputo formalmente definido.
- 1943: Warren McCulloch y Walter Pitts presentan su modelo de neurona artificial.
- 1955: Herbert Simon, Allen Newell y J. C. Shaw, desarrollan el primer lenguaje de programación orientado a la resolución de problemas, el IPL-11. Un año más tarde desarrollan el Logic Theorist, el cual era capaz de demostrar teoremas matemáticos.
- 1956: Se inventa el término Inteligencia Artificial por John McCarthy, Marvin Minsky y Claude Shannon en la Conferencia de Dartmouth.
- 1957: Newell y Simon continúan su trabajo con el desarrollo del General Problem Solver (GPS). GPS era un sistema orientado a la resolución de problemas.
- 1958: John McCarthy desarrolla en el Instituto de Tecnología de Massachusetts (MIT) LISP. Su nombre se deriva de LISt Processor. LISP fue el primer lenguaje para procesamiento simbólico.
- 1959: Rosenblatt introduce el Perceptron.
- 1959-1961: Robert K. Lindsay desarrolla Sad Sam, un programa para la lectura de oraciones en inglés y la inferencia de conclusiones a partir de su interpretación.
- 1963: Quillian desarrolla las redes semánticas como modelo de representación del conocimiento.
- 1964: Bertrand Raphael construye el sistema SIR (Semantic Information Retrieval) el cual era capaz de inferir conocimiento basado en información que se le suministra. En este mismo año, Daniel Bobrow desarrolla el programa STUDENT.
- 1965: Aparecen los sistemas expertos, que predicen la probabilidad de una solución bajo un conjunto de condiciones.
- 1968-1970: Terry Winograd desarrolla el sistema SHRDLU, que permitía interrogar y dar órdenes a un robot que se movía dentro de un mundo de bloques.

- 1968: Marvin Minsky publica *Semantic Information Processing*.
- 1968: Seymour Papert, Danny Bobrow y Wally Feurzeig desarrollan el lenguaje de programación LOGO.
- 1969: Alan Kay desarrolla el lenguaje Smalltalk en Xerox PARC; se publica en 1980.
- 1973: Alain Colmenauer y su equipo de investigación en la Universidad de Aix-Marseille crean PROLOG (del francés PROgrammation en LOGique) un lenguaje de programación ampliamente utilizado en IA.
- 1973: Shank y Abelson desarrollan los guiones, o scripts, pilares de muchas técnicas actuales en inteligencia artificial y la informática en general.
- 1974: Edward Shortliffe escribe su tesis con MYCIN, uno de los sistemas expertos más conocidos, que asistió a médicos en el diagnóstico y tratamiento de infecciones en la sangre.
- 1970-1980: crece el uso de sistemas expertos.
- 1981: Kazuhiro Fuchi anuncia el proyecto japonés de la quinta generación de computadoras.
- 1986: McClelland y Rumelhart publican *Parallel Distributed Processing* (redes neuronales).
- 1988: Se establecen los lenguajes de programación orientados a objetos.
- 1997: Gari Kaspárov, campeón mundial de ajedrez, pierde ante la computadora autónoma Deep Blue.
- 2006: se celebra el aniversario de la Conferencia de Dartmouth, con el congreso en español, 50 Años de Inteligencia Artificial - Campus Multidisciplinar en Percepción e Inteligencia 2006.
- 2009: se desarrollan sistemas inteligentes terapéuticos que permiten detectar emociones para poder interactuar con niños autistas.
- 2011: IBM desarrolla una supercomputadora llamada Watson, la cual gana una ronda de tres juegos seguidos de Jeopardy, y vence a sus dos máximos campeones.
- 2016: Un programa informático gana cinco a cero, al triple campeón de Europa de Go.

- 2018: Se lanza el primer televisor con inteligencia artificial por parte de LG Electronics, con una plataforma denominada ThinQ.

APLICACIONES COMERCIALES DE LA IA Y LA CD

La IA ha sido usada en un amplio número de campos como la robótica, la comprensión y traducción de lenguajes, aprendizaje de palabras, etc. Los principales y más destacados campos donde puede encontrarse una notoria evolución de la IA son en las Ciencias de la Computación, finanzas, hospitales y medicina, industria pesada, servicio de atención al cliente, transportación y juegos. Un estudio llevado a cabo por la empresa Gartner en 2016 muestra que para el 2020, al menos 30% de las todas las compañías usarán IA y CD (Ciencia de Datos), en al menos un fragmento de sus procesos de venta. Algunas áreas de oportunidad de la IA y la CD:

| | | |
|------------------------------------|-------------------------------|-----------------------------------------------|
| Asistentes personales. | Automóviles autónomos. | Banca y finanzas. |
| Edificios inteligentes. | Chatbots. | Ciberseguridad. |
| Comercio electrónico (e-commerce). | Cuidado de la salud. | Diseño de ropa, estilos de zapatos, etcétera. |
| Electrodomésticos. | Entretenimiento. | Logística y cadenas de suministro. |
| Manufactura optimizada. | Servicio al cliente en línea. | Sistemas de recomendación. |
| Telefonía celular. | Turismo. | |

EJEMPLOS DE APLICACIÓN DE LA IA EN EL MUNDO

Los siguientes son ejemplos específicos de la aplicación de la IA en nuestro mundo:

1. **Desplazamiento.** De acuerdo a un estudio en 2015 por parte del Instituto de Transporte de Texas y la Universidad de Texas A&M, los tiempos de desplazamiento en EU han aumentado

- drásticamente, año con año. Esto ocasionó una pérdida estimada de unos 160 billones en la productividad en 2014. Los sistemas de predicción basados en GPS, los sistemas de transporte compartido y los sistemas de autopiloto en los aviones ayudan a aminorar el problema un poco. Soluciones futuras: 1) autos autónomos, 2) transporte compartido optimizado, y 3) sistemas de semáforos inteligentes, relacionados con la IOT.
2. **Correo electrónico.** Pareciera que nuestro “inbox de correo” es lugar no apto para la IA, sin embargo, lo es. Los filtros de spam, los clasificadores inteligentes de correos son solo dos ejemplos de los muchos motores inteligentes que pueden ser usados para el manejo optimizado de nuestro correo electrónico. Soluciones futuras: los respondedores automatizados de correo.
 3. **Banca y finanzas personales.** Hoy por hoy, muchas de nuestras transacciones bancarias se llevan a cabo a través de motores basados en IA: 1) depósito de cheques móviles, 2) prevención de fraudes, y 3) otorgamiento de créditos. Soluciones futuras: sistemas aconsejadores inteligentes para recomendar dónde y cuándo invertir nuestro dinero.
 4. **Redes sociales.** En este caso se reportan ejemplos de sistemas: 1) reciben fotos, las analizan y si encuentran personas hacen recomendaciones de etiquetado; 2) reciben como entrada fotografías, las analizan reconociendo los objetos ahí presentes y hacen recomendaciones de objetos similares; 3) son capaces de identificar significado contextual de emojis y pueden, entonces, sugerir en forma automática emojis de respuesta; y 4) son capaces de seguir movimientos faciales y adicionar efectos animados o máscaras digitales cuando dichos rostros se mueven, estos sistemas son usados por doquier en el mundo. Soluciones futuras: conversación con chatbots inteligentes. Estos chatbots deberán incorporar poderosos: procesadores de lenguaje natural, reconocedores de rostros, analizadores de expresiones faciales, etcétera.
 5. **Compras en línea.** Muchas de las compras que uno hace hoy por hoy son a través de internet. Incluye buscadores, recomendadores, detectores de fraudes, etcétera. Soluciones futuras: sistemas de compra completamente personalizados.

6. **Telefonía móvil.** Prácticamente tres de cada cuatro personas en el mundo usan teléfono celular. Muchos de estos móviles actuales incluyen convertidores de voz a texto, asistentes personales inteligentes, capaces de sostener conversaciones sencillas. Ejemplo, Cortana de Microsoft. Soluciones futuras: asistentes personalizados inteligentes que reduzcan la brecha entre el ser humano y las cosas inteligentes. Deberán ser capaces de responder eficiente y eficazmente a comandos verbales para tocar música, responder preguntas en lenguaje natural, enviar reportes, noticias, actualizaciones financieras, llamar al taxi, hacer citas y recordatorios, entre otros.

ÚLTIMAS NOTICIAS SOBRE EL USO DE LA IA EN LA SOLUCIÓN DE PROBLEMAS DE ACTUALIDAD

Las siguientes son algunas de las últimas noticias, entre el 6 de diciembre de 2018 y el 4 de marzo de 2019, del uso de la IA en la solución de diversos problemas a nivel mundial:

1. El Chef que es capaz de preparar hamburguesas gourmet cada 30 segundos [Hambling, 2018].
2. Sistemas basados en IA para predecir erupciones volcánicas [Voosen, 2018].
3. Auto capaz de repartir mercancías [Metz, 2018].
4. El robot repartidor de Amazon [BBC, 2018].
5. Algoritmo capaz de romper texto tipo Captcha [Xinhua, 2018b].
6. Toyota quiere posicionar un robot amigable en cada hogar [Buckland, 2018].
7. Los algoritmos que son capaces de aprender y ganar a los campeones o profesionales en juegos complejos como el ajedrez, el Go, el Shogi, el póker, etcétera [Strogatz, 2018].
8. Los robots capaces de tomar decisiones por sí mismos en ambientes agrestes y desconocidos [Archer, 2019].
9. Los robots repartidores de helados en entornos semicontrolados [Mims, 2019].

10. Los robots que son capaces de aprender por sí mismos a caminar [Simos, 2019].
11. Los robots que realizan tareas antes atribuibles sólo a los humanos en supermercados y restaurantes [O'Brien, 2019].
12. Las prótesis inteligentes que ayudan a las personas con amputaciones a caminar [Hsu, 2019].
13. Los algoritmos que encuentran vulnerabilidades en programas de cómputo y las resuelven [Brumley, 2019].
14. Los *software* que son capaces de escribir noticias falsas y, a la vez, detectarlas [Kahn, 2019; Wiesemann, 2019].
15. Los que son capaces de hacer las funciones de un reportero [Peiser, 2019].
16. Algoritmos capaces de extraer conceptos a partir de dibujos básicos [Hutson, 2019a_].
17. Las máquinas que pueden diagnosticar casos de demencia [Blanchard, 2019].
18. Los robots que pueden cuidar a niños en hospitales [Dingfelder, 2019].
19. Los asistentes que pueden estacionar un auto y, a la vez, optimizar el lugar de estacionamiento [Choi, 2019].
20. Los que inventan lenguajes no verbales para ganar a las cartas [Hutson, 2019b_].
21. Los sistemas asistenciales que ayudan a reducir las fatalidades relacionadas con los peatones [Eisenstein, 2019].
22. Los robots de Fedex que pronto repartirán pizzas [Shaban, 2019].
23. El sistema basado en DeepMind usado para predecir la potencia de salida del viento [Frangour, 2019].
24. En la conservación de especies [Knok, 2019].

CÓMO SE ESTÁN ADOPTANDO LA IA Y LA CD EN MÉXICO

Nuestro país no se ha quedado atrás, ha hecho esfuerzos importantes para asimilar y utilizar la IA y la CD en la solución de problemas de diversa índole. Analicemos algunos ejemplos:

1. **Chatbots. Prospera Digital:** chatbot que permite darle seguimiento a todo el proceso materno infantil gratuito y para el cual se realizó una alianza con la Secretaría de Desarrollo Social [GDM, 2018].
2. **Cobots (robots colaborativos).** Interactúan con el humano de forma directa y están cada vez más presentes en distintos sectores industriales de nuestro país, incorporan tecnología de software desarrollada en México o integran componentes nacionales de la industria de la robótica. Un caso de éxito es YuMi, un *cobot* de dos brazos para la automatización de procesos, actualmente ocupado en áreas de ensamble en la industria de la electrónica de consumo, para la unión y conexión de piezas pequeñas [Tejeda, 2018].
3. **Educación.** Se han cubierto varios puntos con éxito [Universia, 2020]: la realidad virtual que permite a los alumnos interactuar con entornos creados expresamente para los contenidos a enseñar; la robótica educativa que ofrece a los alumnos la posibilidad de aprender a programar robots y, a través de ella, entrar en contacto con la tecnología; los sistemas de tutoría inteligente que guían a los estudiantes en su aprendizaje; los sistemas de aprendizaje en línea que admiten evaluaciones más sencillas; y la analítica del aprendizaje que detecta errores y puntos débiles de los alumnos de forma temprana.
4. **Robótica de servicio.** Diversos tipos de robots ya operan en el hogar, oficinas, fábricas, hoteles, hospitales, museos, restaurantes, aeropuertos, etcétera [Velazco, 2017].

Con el título *En Miras Hacia una Estrategia de Inteligencia Artificial (IA) en México*, la consultoría británica Oxford Insights, con el apoyo financiero de la Embajada de Reino Unido, desarrolló el primer diagnóstico rumbo a una Estrategia Nacional de Inteligencia Artificial. El estudio apareció en septiembre de 2018. Contiene un análisis sobre las ventajas, oportunidades y desafíos de México en la materia, así como recomendaciones en el corto, mediano y largo plazo para todos los sectores de la sociedad [Zapata, 2018a; Zapata, 2018b].

RETOS ACTUALES (2018) PARA IMPLEMENTAR SISTEMAS BASADOS EN IA

Toda técnica, conforme se va estableciendo, experimenta diversos retos. La IA no es la excepción. Aunque ya se cuenta con diversas mejoras en el poder de cómputo (músculo); en las técnicas para el manejo de grandes cantidades de datos; y en las técnicas para el aprendizaje de máquinas (aprendizaje profundo); todavía hay mucho camino que recorrer en materia de datos, inteligencia artificial multitarea y *hardware*.

En cuanto a los datos, la IA necesita miles de veces más datos que los requeridos por el cerebro humano para poder comprender conceptos y características. La capacidad de las máquinas para ver, entender e interactuar con el mundo está creciendo a un ritmo acelerado, apoyado en el volumen de datos que les ayuda a aprender y entender aún más rápido. Cada año la cantidad de datos que producimos se duplica. En la próxima década habrá unos 150 mil millones de sensores conectados a la red, equivalente a más de 20 veces la población de la Tierra. En este sentido, el *big data* ha sido y será un gran aliado de la IA para procesar esta cantidad cada vez más grande de información y volverla útil.

En cuanto a la IA multitarea, una vez que una máquina de IA es entrenada, es muy efectiva para tareas como el reconocimiento facial o de voz. Sólo son capaces de realizar tareas específicas. Actualmente, no hay máquinas inteligentes capaces de cambiar de una tarea a la otra. Este es uno de los retos actuales de la IA. Una propuesta de solución consiste en usar las llamadas “redes neuronales progresivas”, aún en fase de desarrollo y prueba [Rusu, 2016; Zhan, 2018].

Finalmente, en cuanto al *hardware*, a pesar de la gran capacidad de procesamiento que han alcanzado las máquinas y toda la información que se encuentra disponible, esta tecnología aún se encuentra limitada por el *hardware*. La infraestructura necesaria para experimentar con la IA es escasa y costosa. Se necesitan equipos más potentes para tener mejores resultados. Posibles soluciones a este problema son: el cómputo masivamente paralelo, el cómputo cuántico, el cómputo DNA, el cómputo óptico, el cómputo atómico, el cómputo bacteriano, etcétera.

CONSIDERACIONES PARA LAS EMPRESAS

Las consideraciones más importantes de la IA descansan en la formación de recursos humanos calificados y especializados. Se considera que un experto en IA, por ejemplo, aparte de los conocimientos básicos, debería también incorporar a su currícula conocimientos de aprendizaje para máquinas (aprendizaje profundo y análisis predictivo), procesamiento de lenguaje natural (traducción, clasificación y agrupamiento, extracción de información), habla (de habla a texto y de texto a habla), sistemas expertos, planeación y optimización, robótica, visión (análisis de imágenes y reconocimiento de patrones).

Para el desarrollo de la IA, en nuestro país se requiere de personal altamente calificado. Las siguientes instituciones ya cuentan, al menos, con un programa de estudios:

1. Ingeniería en IA, Universidad Panamericana [UP-IIA].
2. Maestría en IA, UNAM [UNAM-MIA].
3. Maestría en Ciencias en IA, UAQ [UAQ-MCIA].
4. Maestría en IA, Universidad Veracruzana [UV-MIA].
5. Maestría en Ciencias con Especialidad en Sistemas Inteligentes, Tecnológico de Monterrey [TM-MCESI].
6. Maestría en Ciencia de Datos, ITAM [ITAM-MCD].
7. Maestría en Ciencia de Datos, UVM [UVM-MCD].
8. Maestría en Ciencia de los Datos y Procesamiento de Datos Masivos (Big Data), Universidad Cuauhtémoc [UC-MCDPDM].
9. Doctorado en IA, Universidad Veracruzana [UV-DIA].
10. Ingeniero en IA. Instituto Politécnico Nacional [IPN-IIA].
11. Licenciado en Ciencia de Datos. Instituto Politécnico Nacional [IPN-LCD].

Otras instituciones que también tienen planeado abrir al público carreras con IA son el Instituto Politécnico Nacional, el Tecnológico Nacional de México, la Benémérita Universidad Autónoma de Puebla y la Universidad del Valle de Atemajac.

En el siguiente sitio se dan detalles de porqué crear carreras relacionadas particularmente con Ciencia de Datos. De acuerdo con

un estudio reciente, realizado por la empresa Gassdoor, la carrera de científico de datos está entre las 50 mejores pagadas en Estados Unidos en 2019. Durante cuatro años consecutivos, el salario medio promedio anual para esta carrera fue de 108,000 dólares, con un nivel de satisfacción de 4.3/5.0 y unas 6,500 posiciones disponibles de trabajo el 2 de enero de 2019 [Dolon, 2019].

Una carrera relacionada con IA, como por ejemplo, un ingeniero en aprendizaje para máquinas o un ingeniero en aprendizaje profundo, puede llegar a obtener un salario promedio de 144,000 dólares [Indeed, 2020]. En ambos casos, esto indica que dichas carreras son, actualmente, de las más solicitadas y de las mejores pagadas a nivel mundial.

ÉTICA EN EL USO DE LA IA

Como toda tecnología, la IA debería ser usada con cautela y con ética, por ejemplo:

1. Debería ser utilizada para construir máquinas para el servicio del ser humano, no para su perjuicio.
2. No debería ser usada para construir máquinas de guerra.
3. No debería ser usada para influenciar las elecciones presidenciales en países.
4. No debería ser usada para no contratar personas para un trabajo determinado.

Para esto es necesario entender los riesgos asociados a la IA, expertos explican las diferencias entre la IA débil y la fuerte, sus usos e implicaciones [Chawla, 2019]. En el artículo “¿Debería regularse la Inteligencia Artificial?” [Etzioni, 2017], se describe con claridad el por qué tecnologías como la IA debieran ser reguladas. En artículos, como los reportados en “Empresas estadounidenses pusieron a trabajar a un número récord de robots en 2018” [Integra Nexus, 2020; Tegmark, 2016], se habla de los beneficios, de los riesgos, así como de lo bueno y lo malo de construir sistemas basados en IA.

POR QUÉ AHORA ES EL TIEMPO DE LA IA

Es indudable que hoy por hoy es el tiempo de la IA, ya que cuenta con más y mejores capacidades de cómputo, de datos, de técnicas de procesamiento y de análisis. De acuerdo con [Logicalis, 2018], 2018 es considerado como el año definitivo en el despegue de las soluciones basadas en la IA.

Consultoras de prestigio han pronosticado un mercado de más de 31 mil millones de dólares en 2019, un 55% más que en 2014. La empresa Gartner ha detectado que una de cada cuatro empresas ya usa IA o tiene planes de hacerlo en el a corto plazo. El 73% de los desarrolladores que no usan IA planifican aprender a usarla en 2018 [Rayome, 2018].

PAÍSES LÍDERES EN EL DESARROLLO DE LA IA

Entre los países que más aportan en el desarrollo y uso de la IA y la CD se puede citar a los siguientes:

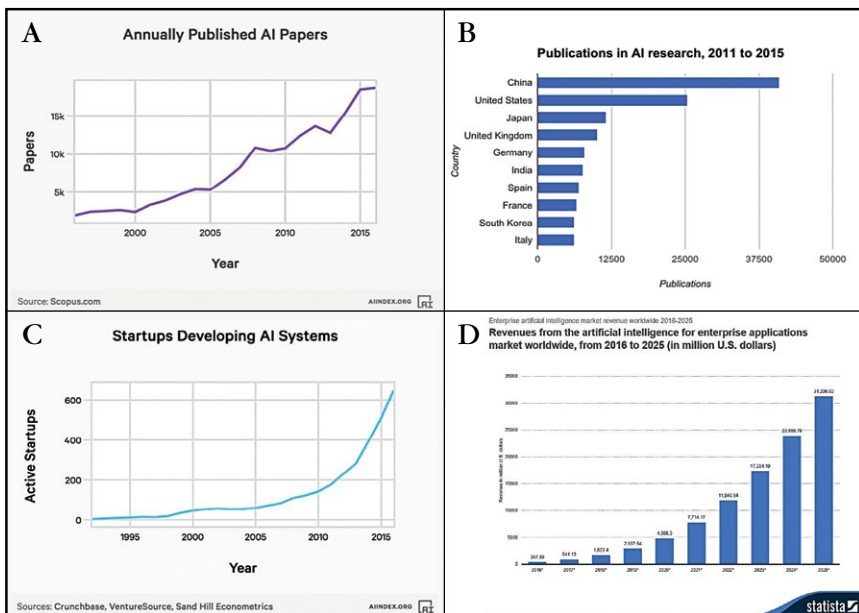
| País número o grupo de países | Acción |
|--------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Estados Unidos | Google anuncia la apertura de un laboratorio de IA en la Universidad de Princeton. Diciembre 18,2018. El presidente Trump anuncia su plan para mantener a EU en primer lugar en IA. Febrero 11, 2019 [74,75]. |
| China | Lanzar su plan nacional para el desarrollo de la IA, marzo de 2017. Un trillón de dólares en IA para 2030. |
| El grupo de 24 países europeos | Firman convenio en IA para competir con EU y China. Abril, 2018. |
| Francia | Anuncia dedicar 1.85 billones de dólares para el desarrollo de IA para competir con EU y China. Marzo, 2018. |
| La Comisión Europea | Anuncia 1.5 billones de euros para el desarrollo de la IA. Abril, 2018. |
| Canadá | Informa sobre una inversión de 125 millones de dólares para el desarrollo de IA. Marzo, 2017. Montreal quiere ser el centro mundial de desarrollo en IA. Mayo 20, 2017. |
| Inglaterra, Alemania, Suiza y Suecia | Proponen establecer un IA Hub para competir con EU y China. Abril, 2018. |

| | |
|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Noruega | Anuncia la apertura de un Centro de Investigación en IA, Universidad de Adger. Febrero, 2017. |
| Inglaterra | Anuncia que la Universidad de Cambridge recibirá apoyo de 10 millones de libras para una supercomputadora para IA. Anuncia que Samsung instalará un Laboratorio de Investigación en IA en la Universidad de Cambrige. Mayo, 2018. |

ALGUNAS ESTADÍSTICAS

En las figuras 2A a la 2D se muestran algunas estadísticas sobre el crecimiento acelerado en las publicaciones relacionadas con IA, los países con más publicaciones a nivel mundial, empresas con aplicaciones basadas en IA, así como las ganancias de dichas empresas. Como se puede ver, en materia de publicaciones, China ya encabeza las estadísticas.

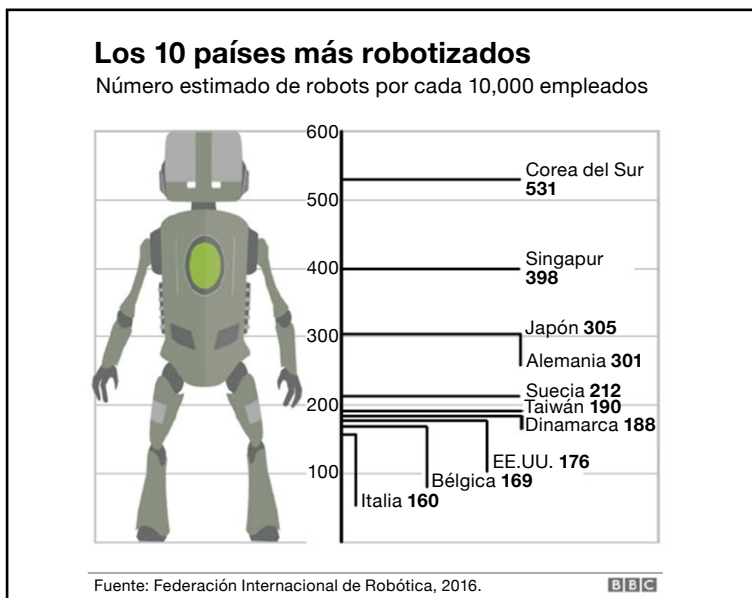
Figura 2. (A) Crecimiento en el número de publicaciones anualizado. (B) Publicaciones por país en IA entre 2011 y 2015. (C) Crecimiento en el establecimiento de empresas de base tecnológica en IA. (D) Crecimiento de las ganancias en millones de dólares para las empresas que invierten en IA.



En materia de patentes, la World Intellectual Property Organization (WIPO) estima que en los últimos cinco años se ha producido un *boom* en IA, con más de 50% de todas las patentes desde 2013. De 2013 a 2017, el crecimiento ha pasado de 18,995 a 55,660 patentes. IBM con 8,290 y Microsoft 5,930, respectivamente. Sin embargo, como se reporta en [Xinhua, 2018a], China ostenta el 22% de patentes a nivel mundial. La proporción de artículos/invenciones ha bajado de 8:1 en 2010 a 3:1 en 2016. Finalmente, el aprendizaje para máquinas se ha convertido en la técnica más usada en patentes, mientras que la visión por computadora es la aplicación más popular.

Periódicos como el Reforma informan a la población sobre el rezago en IA que las universidades, institutos y centros de investigación tienen con respecto a otros países en materia de producción de patentes [González, 2019]. Mientras que instituciones de educación en China tienen unas 2600 solicitudes de patentes, instituciones mexicanas tienen sólo 8 (7 el Tecnológico de Monterrey y una la UNAM).

Figura 3. Países más robotizados.



Las aplicaciones representan un mercado que equivale hoy a 21 mil 400 millones de dólares, pero llegará a 190 mil millones de dólares en 2025, según la firma Markets and Markets.

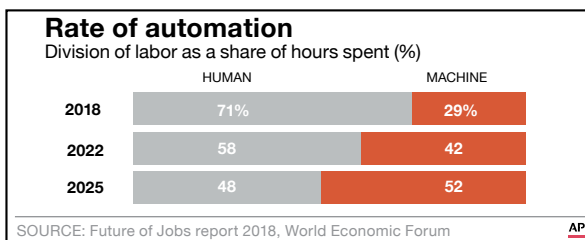
Los países más robotizados por cada 10,000 empleados a nivel mundial son Corea, Singapur, Japón, Alemania, Suecia, Taiwán, Dinamarca, EU, Bélgica e Italia (Figura 3). México ocupa el trigésimo tercer lugar con 33 robots por cada 10,000 empleados, por encima de Argentina que ocupa el lugar trigésimo sexto con 16 robots por cada 10,000 empleados y Brasil, lugar vigésimo octavo, con 11 robots por cada 10,000 empleados. En [Reuters, 2019] se dan otras estadísticas interesantes.

PARA REFLEXIONAR

Con los avances mostrados en el desarrollo de la IA, los seres humanos deberíamos hacernos ya preguntas como las siguientes [Bui, 2015; Ordman, 2017]: 1) ¿Tarde o temprano, un algoritmo nos reemplazará? 2) ¿Mi trabajo será realizado por una máquina?

Se dice que 2025-2035 será un periodo crítico para los trabajos tradicionales. Muchos empleos van a desaparecer, otros van a aparecer. Más del 50% de los trabajos serán realizados por máquinas en 2025. Muchos trabajos desaparecerán y muchos otros aparecerán. Un gran reto será el entrenamiento y reentrenamiento de empleados para el nuevo mundo de trabajos (Figura 4). Este tipo de estudios especializados, por ejemplo, han sido llevados a cabo por la Universidad de Yale y la Universidad de Oxford, instituciones que realizaron una encuesta a 252 investigadores prominentes en IA.

Figura 4. Tasa de automatización de los trabajos actuales.



Información relevante sobre la importancia de la IA en nuestros días, la pertinencia de patentar e innovar, los jugadores más destacados del sector industrial y la academia, así como, la distribución geográfica y de trabajos científicos se pueden encontrar en el reciente reporte de la World Intellectual Property Organization (WIPO) Technology Trend 2019 - Artificial Intelligence [Gurry, 2019].

CONCLUSIONES

De todo lo visto hasta el momento podemos concluir lo siguiente: aunque todavía no se descubren las leyes que rigen el proceso de inteligencia en los seres vivos, ya es posible construir máquinas inteligentes, no hay leyes de la física que se opongan a esto.

Se ha “rascado” la superficie, a través de múltiples ejemplos de la aplicación de la IA en la cotidianidad. El cómo la IA afectará nuestras vidas a gran escala en el futuro cercano, lo expresa Kevin Kelly, quien predice que la IA conforme se va integrando más profundamente en nuestras vidas, vendrá a ser, sin duda, la nueva infraestructura que impulse la siguiente revolución industrial [Kelly, 2016].

Muchas otras interesantes referencias al tema pueden consultarse en [Adidas 2018; Bostrom 2016; Ford 2016; Heaven 2017; Tegmark 2018; West 2018].

Agradecimientos. Juan Humberto Sossa Azuela agradece a la Academia de Ingeniería y al Instituto Politécnico Nacional por el apoyo brindado para la realización de esta investigación. Esta investigación fue financiada con apoyos económicos bajo los proyectos SIP 20190007 y 20200886 y 65 en el marco de la convocatoria: Fronteras de la Ciencia del CONACyT.

REFERENCIAS

- Adidas, W. and J. Ledezma Millán. *El Ascenso de la Automatización: La Tecnología y los Robots Reemplazarán a los humanos*. Adidas Wilson. 2018.
- Archer, J. *Martian robot will explore the Red Planet with mind of its own*. *The Telegraph*. Enero 2 de 2019. Disponible en: <https://www.telegraph.co.uk/technology/2019/01/02/uk-tests-autonomous-martian-robot/>.
- BBC News. *Amazon Scout robots take to pavements in Washington State*. Diciembre 17 de 2018. Disponible en: <https://www.bbc.com/news/amp/technology-46987779>.
- Blanchard, S. *'Robbie the Robot' can spot worsening dementia after watching 13 episodes of Emmerdale (and next the scientists want to make it view Friends...)*, Mail Online. Febrero 6 de 2019. Disponible en: <https://www.dailymail.co.uk/health/article-6673115/Robbie-Robot-spot-worsening-dementia-watching-13-episodes-Emmerdale.html>.
- Bostrom, N. *Superinteligencia: Caminos, peligros y estrategias*. TEELL. Febrero 25 de 2016. Tercera Edición.
- Brooks, R. *A Robot that Walks; Emergent Behaviors from a Carefully Evolved Network, Neural Computation*, 1989, 1(2):253-262, DOI:10.1162/neco.1989.1.2.253.
- Brown, N. and T. Sandholm. Libratus: *The Superhuman AI for No-Limit Poker*. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. Demos*. Pp. 5226-5228. Disponible en: <https://doi.org/10.24963/ijcai.2017/772>. 2017.

Brumley, D. *Mayhem, the Machine That Finds Software Vulnerabilities, Then Patches Them*. *IEEE Spectrum*. Enero 29 de 2019. Disponible en: https://spectrum.ieee.org/computing/software/mayhem-the-machine-that-finds-software-vulnerabilities-then-patches-them?utm_source=circuitsandsensors&utm_campaign=circuitsandsensors-02-05-19&utm_medium=email.

Bui, Q., *Will Your Job Be Done By A Machine? Planet Money*. Mayo 21 de 2015. Disponible en: <https://www.npr.org/sections/money/2015/05/21/408234543/will-your-job-be-done-by-a-machine>.

Buckland, K. *Toyota Wants to Put a Robot Friend in Every Home*. Bloomberg. Diciembre 26 de 2018. Disponible en: <https://www.bloomberg.com/news/articles/2018-12-26/toyota-wants-to-put-a-robot-in-every-home-and-make-it-your-pal?srnd=premium-asia>.

Chawla, S. *Understanding the Risks Associated With AI*. *IEEE Spectrum*. Febrero 22, 2019. Disponible en: <https://spectrum.ieee.org/news-from-around-ieee/the-institute/ieee-news/understanding-the-risks-associated-with-ai>.

Choi, Ch. Q. *How Self-Driving Cars Might Transform City Parking*. *IEEE Spectrum*. Febrero 20 de 2019. Disponible en: https://spectrum.ieee.org/cars-that-think/transportation/self-driving/autonomous-parking?utm_source=robotics-news&utm_campaign=roboticsnews-02-26-19&utm_medium=email.

Copeland, B. J. *CYC computer science*. Britanica. Disponible en: <https://www.britannica.com/topic/CYC>.

Coskun, M., A. Uçar, Ö. Yildirim and Y. Demir, "Face recognition based on convolutional neural network," *2017 International Conference on Modern Electrical and Energy Systems (MEES)*, Kremenchuk, 2017, pp. 376-379, doi: 10.1109/MEES.2017.8248937.

Davies, A. *The WIRED Guide to Self-Driving Cars*. Wired. December 13, 2018. Disponible en: <https://www.wired.com/story/guide-self-driving-cars/>.

Dingfelder, S. *Is that a robot? Dr. Bear Bot helps care for kids at local hospital*. The Washington Post. Febrero 14 de 2019. Disponible en: https://www.washingtonpost.com/express/2019/02/15/is-that-robot-dr-bear-bot-helps-care-kids-local-hospital/?utm_term=.4742f92ba94f.

Dolon, M., *Data Scientist Best Job in US for 2019*. IEEE. Enero 31 de 2019. Disponible en: <https://jobs.ieee.org/jobs/content/-Data-Scientist-Best-Job-in-US-for-2019-2019-01-31>.

Eisenstein, P. A. *New technology shows promise reducing skyrocketing pedestrian fatalities*. CNBC. Febrero 24 de 2019. Disponible en: <https://www.cnbc.com/2019/02/23/new-technology-shows-promise-reducing-pedestrian-fatalities.html>.

Etzioni, A. and O. Etzioni. *Should Artificial Intelligence Be Regulated?* Issues. Summer 2017. Disponible en: <https://issues.org/perspective-should-artificial-intelligence-be-regulated/>.

Ford, M. *El ascenso de los robots*. Paidós. Septiembre 16 de 2016.

Frangour, A. *Alphabet's DeepMind uses machine learning to predict wind power output*. CNBC. Febrero 27 de 2019. <https://www.cnbc.com/2019/02/27/alphabets-deepmind-uses-machine-learning-to-predict-wind-power-output.html>.

Gadam, S. *Artificial Intelligence and Autonomous Vehicles*. Data Driven Investor. Abril 19, 2018. Disponible en: <https://medium.com/datadriveninvestor/artificial-intelligence-and-autonomous-vehicles-ae877feb6cd2>.

Galvan, M. *(Pre) Historia de los videojuegos* — Parte 1. Level Up. Diciembre 7 de 2019. Disponible en: <https://www.levelup.com/articulos/553174/Pre-Historia-de-los-videojuegos-Parte-1>.

Garner, H. *Los 9 tipos de inteligencia según Howard Gardner*. Noviembre 19 de 2018. Disponible en: <http://noticias.universia.com.ar/cultura/noticia/2015/04/01/1122560/9-tipos-inteligencia-segun-howard-gardner.html>.

GDM. *Prospera Digital: Plataforma digital para el desarrollo social y la inclusión financiera*. Gobierno de México. 2018. Disponible en: <https://www.google.com/search?client=firefox-b-d&q=Prospera+Digital>.

Godoy, M. *De continuar con el proceso evolutivo, una máquina podría alcanzar niveles de inteligencia inaccesibles para los seres humanos en poco tiempo*. Nic México. Disponible en: <https://www.nicmexico.mx/2020/05/10/inteligencia-artificial-parte-iii-mas-alla-de-nuestro-control/>.

González, A., *Se rezagan universidades en inteligencia artificial*. *Reforma*. Marzo 4 de 2019. <https://www.reforma.com/aplicacioneslibre/preacceso/articulo/default.aspx?id=1622677&urlredirect=https://www.reforma.com/aplicaciones/articulo/default.aspx?id=1622677>.

Gurry, F. *Wipo Technology Trends 2019: Artificial Intelligence*. World Intellectual Property Organization. Disponible en: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf.

Hambling, D. *The chef that can make a gourmet burger every 30 seconds*. *BBC*. Diciembre 5 de 2018. Disponible en: <http://www.bbc.com/future/story/20181204-the-chef-making-fast-food-even-faster>.

Heaven, D. *New Scientist: Machines that Think: Everything you need to know about the coming age of artificial intelligence (Instant Expert)*. John Murray Learning. Octubre 19 de 2017.

Hsu, J. *AI Helps Amputees Walk With a Robotic Knee*. IEEE Spectrum. Enero 25 de 2019. Disponible en: https://spectrum.ieee.org/the-human-os/biomedical/bionics/ai-helps-humans-walk-on-robot-prosthetic-knee?utm_source=techalert&utm_campaign=techalert-01-31-19&utm_medium=email.

Hutson, M. *Pictionary-playing computer connects to humans' 'deep thoughts'*. Science. Febrero 5 de 2019. Disponible en: https://www.sciencemag.org/news/2019/02/pictionary-playing-computer-connects-humans-deep-thoughts?r3f_986=https://www.google.com/.

Hutson, M. *Artificially Intelligent Players Invent Non-verbal "Languages" to Win Card Games*. IEEE Spectrum. Febrero 22 de 2019. https://spectrum.ieee.org/tech-talk/computing/software/artificially-intelligent-players-invent-nonverbal-languages-to-win-card-games?utm_source=techalert&utm_campaign=techalert-02-28-19&utm_medium=email.

Indeed. *Artificial Intelligence Engineer Salaries in the United States*. Indeed. Septiembre 10 de 2020. Disponible en: <https://www.indeed.com/salaries/Artificial-Intelligence-Engineer-Salaries>.

Integra Nexus. Ventajas y Desventajas de la Inteligencia Artificial en Empresas. Integra Nexus. Enero 17 de 2020. Disponible en: <https://nexusintegra.io/es/ventajas-y-desventajas-de-la-inteligencia-artificial/>.

Instituto Politécnico Nacional. *Ingeniería en Inteligencia Artificial*. Disponible en: <https://www.escom.ipn.mx/htmls/oferta/iia2020.php>.

_____. *Licenciatura en Ciencia de Datos*. Disponible en: <https://www.escom.ipn.mx/htmls/oferta/lcd2020.php>.

Instituto Tecnológico Autónomo de México. Maestría en Ciencia de Datos. Disponible en: <https://mcienciadatos.itam.mx/>.

Jodra, Y. *Primer programa de ajedrez*. Julio 16 de 2015. Disponible en: <http://informaikta.blogspot.com/2015/07/primer-programa-de-ajedrez.html>.

Jones, W. D. *DeepMind's AI Shows Itself to Be a World-Beating World Builder*. IEEE Spectrum. February 1, 2019. Disponible en: https://spectrum.ieee.org/tech-talk/robotics/artificial-intelligence/deepminds-ai-shows-itself-to-be-a-worldbeating-world-builder?utm_source=techalert&utm_campaign=techalert-02-07-19&utm_medium=email.

Kahn, J. *The AI That Can Write a Fake News Story From a Handful of Words*. Bloomberg. Febrero 14 de 2019. Disponible en: <https://www.bloomberg.com/news/articles/2019-02-14/the-ai-that-can-write-a-fake-news-story-from-a-handful-of-words>.

Kelly, K. *Lo inevitable: Entender las 12 fuerzas tecnológicas que configurarán nuestro futuro*. Penguin Random House LLC. Junio 7 de 2016.

Knok, R. *AI empowers conservation biology*. Nature. Marzo 4 de 2019. <https://www.nature.com/articles/d41586-019-00746-1>.

Lenat, D. and R. V. Guha. CYC: *A Midterm Report*. *AI Magazine*. 11(3): Fall 1990. DOI: <https://doi.org/10.1609/aimag.v11i3.842>.

Logicalis. *Inteligencia artificial, el tiempo es ahora*. Logicalis. Febrero 26 de 2018. Disponible en: <https://blog.es.logicalis.com/analytics/inteligencia-artificial-el-tiempo-es-ahora>.

López de Mantaras, M. *El futuro de la IA: hacia inteligencias artificiales realmente inteligentes*. OpenMind BBVA. *¿Hacia una nueva ilustración? Una década trascendente*. Open Mind, BBVA, 2019. Disponible en: <https://www.bbvaopenmind.com/articulos/el-futuro-de-la-ia-hacia-inteligencias-artificiales-realmente-inteligentes/>.

Marti, A. “Estas obras de arte han sido creadas por una doble red neuronal y logran gustar más que las pintadas por humanos.” *Xataka*. Julio 1 de 2017. Disponible en: <https://www.xataka.com/robotica-e-ia/estas-obras-de-arte-han-sido-creadas-por-una-doble-red-neuronal-y-lo-gran-gustar-mas-que-las-pintadas-por-humanos>.

Metz, C. *A Toaster on Wheels to Deliver Groceries? Self-Driving Tech Tests Practical Uses*. *The New York Times*. Diciembre 18 de 2018. Disponible en: <https://www.nytimes.com/2018/12/18/technology/driverless-mini-car-deliver-groceries.html>.

Mims, Ch. *Why Your Ice Cream Will Ride in a Self-Driving Car Before You Do*. *The Wall Street Journal*. Enero 5 de 2019. Disponible en: https://www.wsj.com/articles/why-your-ice-cream-will-ride-in-a-self-driving-car-before-you-do-11546664589?mod=itp_wsj&ru=yahoo.

Musk, E. "Hay que regular la inteligencia artificial antes de que se convierta en un peligro". *El Mundo. Tecnología*. Julio 18 de 2017. Disponible en: https://amp-el-mundo-es.cdn.ampproject.org/v/s/amp.elmundo.es/tecnologia/2017/07/18/596dc3acca4741ea3b8b45a0.html?usqp=mq331AQCCAE%3D&_js_v=a2&_gsa=1#referrer=https%3A%2F%2Fwww.google.com&share=https%3A%2F%2Fwww.elmundo.es%2Ftecnologia%2F2017%2F07%2F18%2F596dc3acca4741ea3b8b45a0.html.

Muente, G. *Conoce los 12 tipos de inteligencia existentes, sus características y cómo influyen en el comportamiento humano*, Marzo 17 de 2020. Rockcontent. Disponible en: <https://rockcontent.com/es/blog/tipos-de-inteligencia/>.

National Geographic. *Breve historia visual de la inteligencia artificial*. Agosto 5 de 2019. Disponible en: https://www.nationalgeographic.com.es/ciencia/breve-historia-visual-inteligencia-artificial_14419/1.

O'Brien, M. *Will robots take your job? Quarter of US workers at risk*. *AP News*. Enero 24 de 2019. Disponible en: <https://www.apnews.com/6034c9ce1af347ec8da996c39b29c51b>.

Ordman, N., *Will an Algorithm Replace Me?* IEEE. Febrero 11 de 2017. Disponible en: <http://jobs.ieee.org/jobs/content/Will-an-Algorithm-Replace-Me-2017-11-02>.

Peiser, J. *The rise of the robot reporter*. *The New York Times*. Febrero 5 de 2019. Disponible en: <https://www.nytimes.com/2019/02/05/business/media/artificial-intelligence-journalism-robots.html>.

Pérez Barreira, A. *Tipos de Inteligencia Artificial* | Débil, general y súper-inteligencia. Futuro Eléctrico. Disponible en: <https://futuroelectrico.com/tipos-de-inteligencia-artificial/>.

Rayome, A. D. *73% of developers who don't use AI plan to learn how in 2018*. *The Republic*. Diciembre 20 de 2017. Disponible en: <https://www.techrepublic.com/article/73-of-developers-who-dont-use-ai-plan-to-learn-how-in-2018/>.

Reuters Staff. "U.S. companies put record number of robots to work in 2018." *Reuters*. Febrero 28 de 2019. Disponible en: <https://www.reuters.com/article/us-usa-economy-robots/u-s-companies-put-record-number-of-robots-to-work-in-2018-idUSKCN1QH0K0>.

Ross, P. E. *DeepMind Achieves Holy Grail: An AI That Can Master Games Like Chess and Go Without Human Help*. *IEEE Spectrum*. December 6, 2018. Disponible en: <https://spectrum.ieee.org/tech-talk/robotics/artificial-intelligence/mb>.

Rusu, A. A. *et al. Progressive Neural Networks*. arXiv: 1606.04671 [cs.LG]. Septiembre 7 de 2016. Disponible en: <https://arxiv.org/abs/1606.04671>.

Simonite, T. *The WIRED Guide to Artificial Intelligence*. *Wired*. Marzo 13 de 2018. Disponible en: <https://www.wired.com/story/guide-artificial-intelligence/>.

Shaban, H. "Your next Fedex delivery could be a pizza." *The Washington Post*. Febrero 27 de 2019. Disponible en: https://www.washingtonpost.com/technology/2019/02/27/your-next-fedex-delivery-could-be-pizza/?noredirect=on&utm_term=.10f2225621d2.

Simos, M. *The Clever Clumsiness of a Robot Teaching Itself to Walk*. *Wired*. Enero 8 de 2019. Disponible en: <https://www.wired.com/story/the-clever-clumsiness-of-a-robot-teaching-itself-to-walk/>.

Simonite, T. *Trump's Plan to Keep America First in AI*. *Wired*. Febrero 11 de 2019. Disponible en: <https://www.wired.com/story/trumps-plan-keep-america-first-ai/>.

Strogatz, S. "One Giant Step for a Chess-Playing Machine." *The New York Times*. Diciembre 26 de 2018. Disponible en: <https://www.nytimes.com/2018/12/26/science/chess-artificial-intelligence.html>.

Syafeeza, A. R. *et al.* "Convolutional Neural Network for Face Recognition with Pose and Illumination Variation,". *International Journal of Engineering and Technology* 6(1):44-57, 2014.

Tegmark, M. Benefits & Risks of Artificial Intelligence. *Future of Life*. 2016. Disponible en: <https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/?cn-reloaded=1>.

Tegmark, M. *Vida 3.0. Ser humano en la era de la Inteligencia Artificial*. Taurus. Septiembre 25 de 2018.

Tejeda León, C. "El uso de la inteligencia artificial en los cobots." *El Economista*. Marzo 26 de 2018. Disponible en: <https://www.eleconomista.com.mx/opinion/El-uso-de-la-inteligencia-artificial-en-los-cobots-20180325-0095.html>.

Tecnológico de Monterrey. *Maestría en Ciencias con Especialidad en Sistemas Inteligentes*. Disponible en: <https://maestriasydiplomados.tec.mx/posgrados/maestria-en-ciencias-con-especialidad-en-sistemas-inteligentes>.

Trump. D. J. “Executive Order on Maintaining American Leadership in Artificial Intelligence”. The White House. Febrero 11 de 2019. Disponible en: <https://www.white-house.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/>.

Universia. *Cómo aplicar la inteligencia artificial en la educación*. UNIVERSIA.MX. Abril 20 de 2020 Disponible en: <http://noticias.universia.net.mx/educacion/noticia/2018/02/26/1157929/5-usos-inteligencia-artificial-educacion.html>.

Universidad Autónoma de Querétaro. Maestría en Ciencias en Inteligencia Artificial. Disponible en: <https://www.uaq.mx/index.php/nivel-posgrados/maestrias/fi/maestria-en-ciencias-en-inteligencia-artificial>.

Universidad Cuauhtémoc. Maestría en Ciencia de los Datos y Procesamiento de Datos Masivos. Disponible en: https://www.ucuauhtemoc.edu.mx/educacionadistancia/index.php/ead/posgrados/mtria_en_ciencia_de_los_datos_y_procesamiento_de_datos_masivos_big_data.

Universidad Autónoma de México. Maestría en Inteligencia Artificial. Disponible en: <https://www.educaedu.com.mx/maestria-en-inteligencia-artificial-master-10666.html>.

Universidad Panamericana. Ingeniería en Inteligencia Artificial, Campus Aguascalientes. Disponible en: <https://www.up.edu.mx/es/licenciatura/ags/ingenieria-en-inteligencia-artificial>.

Universidad del Valle de México. Maestría en Ciencia de Datos. Disponible en: <https://uvm.mx/maestria-ciencia-de-datos-en-linea>.

Universidad Veracruzana. Doctorado en Inteligencia Artificial. Disponible en: <https://www.uv.mx/dia/>.

_____. *Maestría en Inteligencia Artificial*. Disponible en: <https://www.uv.mx/mia/>.

Velazco, J. *México debe acelerar el uso de los robots*. *Milenio*. 2020. Septiembre 11 de 2017. Disponible en: <http://www.milenio.com/negocios/mexico-debe-acelerar-el-uso-de-los-robots>.

Voosen, P. *Artificial intelligence helps predict volcanic eruptions*. *Science*. Diciembre 11 de 2018. Disponible en: <http://www.sciencemag.org/news/2018/12/artificial-intelligence-helps-predict-volcanic-eruptions>.

Wang, M and D. Weihong. *Deep Face Recognition: A Survey*. arXiv:1804.06655v9 [cs.CV] 1 Agosto 2020.

West, D. M. *The Future of Work: Robots, AI, and Automation*. Brookings Institution Press. Mayo 15 de 2018.

Wiesemann, S. *Software that can automatically detect fake news*. *Fraunhofer*. Febrero 1 de 2019. Disponible en: <https://www.fraunhofer.de/en/press/research-news/2019/february/software-that-can-automatically-detect-fake-news.html>.

Xinhua. "China concentra el 22% de las patentes de inteligencia artificial a nivel mundial". *AE Tecno*. Mayo 11 de 2018. Disponible en: <https://tecno.americaeconomia.com/articulos/china-concentra-el-22-de-las-patentes-de-inteligencia-artificial-nivel-mundial>.

Xinhua Net. "AI can easily break text CAPTCHA: new research". Diciembre 24 de 2018. Disponible en: http://www.xinhuanet.com/english/2018-12/24/c_137695066.htm.

Zapata, E. (a) *Estrategia de Inteligencia Artificial MX 2018*. Gobierno de México. Marzo 22 de 2018. Disponible en: <https://www.gob.mx/mexicodigital/articulos/estrategia-de-inteligencia-artificial-mx-2018>.

_____. (b) *Inteligencia Artificial en México*. Gobierno de México. Marzo 29 de 2018. Disponible en; <https://datos.gob.mx/blog/inteligencia-artificial-en-mexico?category=proyectos&tag=nula>.

Zhan, Z. *et al. Progressive Neural Networks for Image Classification. Ground IA*. Abril 25 de 2018. Disponible en: <https://www.groundai.com/project/progressive-neural-networks-for-image-classification/1>.

Ontología y narrativa en los sistemas de inteligencia artificial, una mirada antropológica

LILY DÍAZ-KOMMONEN

INTRODUCCIÓN

El objeto del conocimiento surge y se elabora en el contexto de la actividad que se realiza. Cada campo de investigación tiene sus áreas y perímetros de acción, en los cuales se desarrollan sus distintas tareas. La historia, la práctica y la teoría, por ejemplo, se pueden considerar como tres facetas, a través de las cuales, ejercer una investigación procedente del arte, que se hace con el arte y por la que se puede saber más sobre el arte.

Es posible observar en el cine, por ejemplo, los cambios que han sufrido nuestras ciudades en los últimos cien años. También entendemos que no es lo mismo la Gestalt visual que recibimos al observar una proyección en 35 mm, en un video digital en alta resolución, o en un modelo de simulación hecho con una computadora. En nuestra era de la cultura digital en la cual se gestionan temas como los de la realidad virtual, la investigación histórica con el arte y sobre el arte, nos puede ayudar a instar formaciones de archivística que quizás nos parecerían como esotéricas. Estas nos pueden proveer historiografías sensuales sobre la cultura, objetos de arte que informen nuestra comprensión de la realidad.

EXAMINAR LA CULTURA

La base primordial de una mirada antropológica hacia la actividad

artística debe partir de un entendimiento sobre la cultura material. Todo comienza con el artefacto. No importa si ha sido realizado con métodos artesanales y materiales tradicionales o si surge como resultado de una tarea transdisciplinaria que envuelve varias disciplinas, inclusive las artes, el diseño o la ciencia. La razón de ser de un artefacto, sus interfaces y cómo estos se extienden en su entorno, es un tópico de investigación continua en nuestro grupo.

El artefacto es algo conceptual o material creado por un ser humano. Todos los artefactos se pueden describir usando un número infinito de propiedades y atributos. Estas propiedades, que se pueden representar como vectores, también definen la posición del artefacto en un espacio multidimensional (Díaz y Kaipainen).

Utilizamos el análisis del artefacto basado en el método Artefact Analysis, desarrollado por la doctora Susan Pearce en la Universidad Leicester del Reino Unido. Con este método, a través de la descripción de sus propiedades, hemos podido desgajar en elementos y facetas los distintos componentes del objeto de arte, o de diseño y su forma representativa. Este último detalle también nos provee una plataforma inicial para la elaboración de las ontologías en una aplicación.

ANTECEDENTES CON EL USO DE REDES NEURONALES

Este entendimiento nos ha facilitado el poder trabajar, usando materiales y tópicos arraigados en el arte y el diseño, en una de las áreas de investigación relacionadas con la inteligencia artificial: las redes neuronales.

En el proyecto sobre las “Comunidades de interés para promover la herencia cultural regional en Europa” (Communities of Interest to Promote the Heritage of European Regions/CIPHER, 2002-2004) desarrollamos actividades con el Mapa auto organizado, (Self-Organizing Map. SOM), basado en el algoritmo inventado por el científico finlandés, Teuvo Kohonen. Las actividades que se llevaron a cabo incluyeron experimentos con el SOM, como herramienta para organizar y visualizar materiales culturales que se muestran en archivos

digitales en la red. Distintas comunidades, entre ellas, un grupo de estudiantes de presecundaria (de la edad entre 12-14 años de edad) participaron en el estudio (Díaz y Parttanen).

SOM es un tipo de red artificial neuronal (ANN) que no necesita supervisión y que se puede utilizar para agrupar datos, sin clasificación predefinida del input data. Honkela describe al SOM como una matriz, cuyos nódulos (células) responden a señales de input de una manera ordenada. Al usar una metáfora del ámbito de la biología, el SOM es un algoritmo cuya actividad mimetiza la capacidad del cerebro de formar mapas topológicos. El profesor Marín Diazaraque (Departamento de Estadística y Economía, Universidad Carlos III de Madrid) lo define de la siguiente manera:

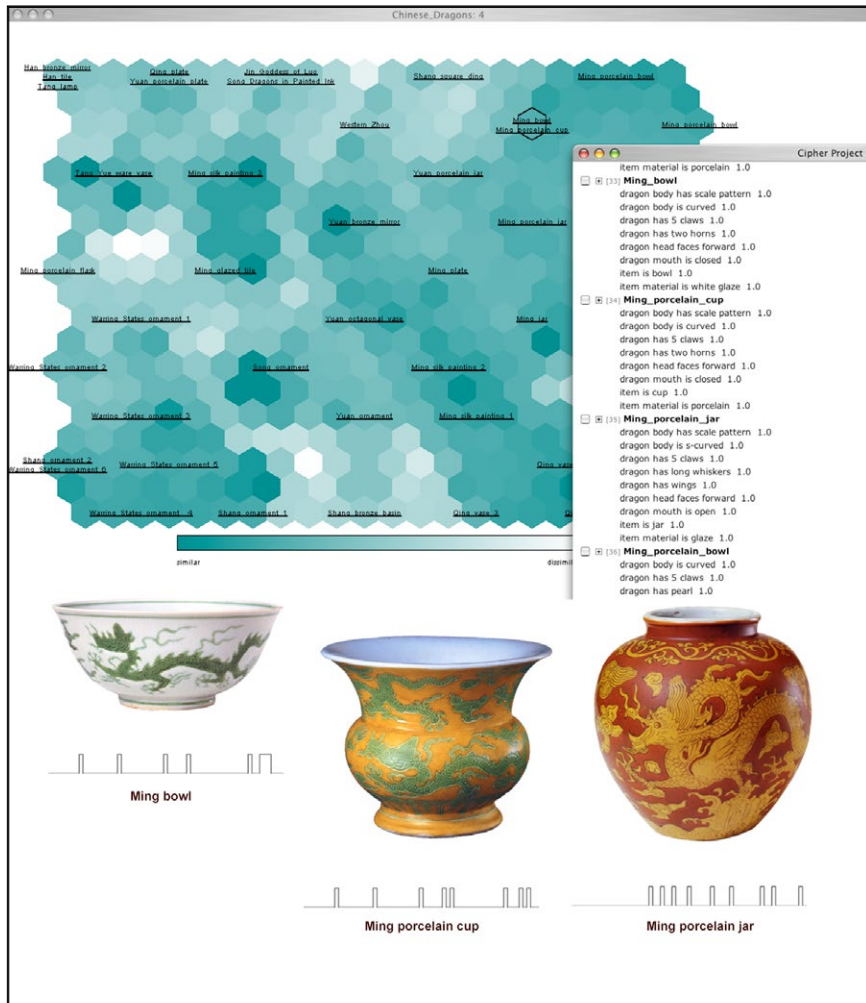
“Este tipo de red posee un aprendizaje no supervisado competitivo. El objetivo de este aprendizaje es categorizar los datos que se introducen en la red. Se clasifican valores similares en la misma categoría y, por tanto, deben activar la misma neurona de salida. Las clases o categorías deben ser creadas por la propia red, puesto que se trata de un aprendizaje no supervisado, a través de las correlaciones entre los datos de entrada”.

En la psicología Gestalt, la ley de semejanza define, como en la percepción humana, aquellos elementos que tienen características parecidas o iguales, y son percibidos como más relacionados que aquellos que no. Pero en una visualización científica esto no es suficiente para llegar a un análisis conclusivo. Hay que establecer también la identidad de los vectores señalados en cada artefacto, que estos portan las características (propiedades) que el usuario les atribuye como parte de la descripción, y finalmente que los artefactos que se aglutinan en un núcleo sí comparten las propiedades indicadas.

Para esta tarea, se prosiguió con un experimento en forma de una visualización, en el cual histogramas de los distintos artefactos se usan como elementos de identificación. En la Figura 1 señala un núcleo en un mapa auto-organizado, la lista de las propiedades correspondientes a los tres artefactos, una ilustración de los artefactos

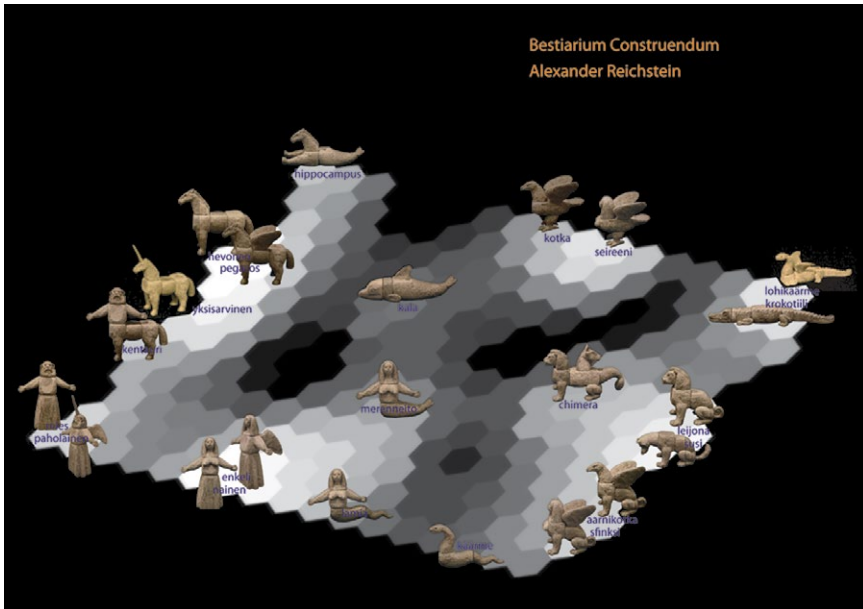
con histogramas, y finalmente una comparación de los histogramas de cada artefacto.

Figura 1. Mapa auto-organizado (SOM) que muestra un experimento de como la representación del dragón aparece en artefactos de la cultura china, a través de distintas épocas. Métodos como el análisis de artefactos se utilizaron para crear las descripciones.



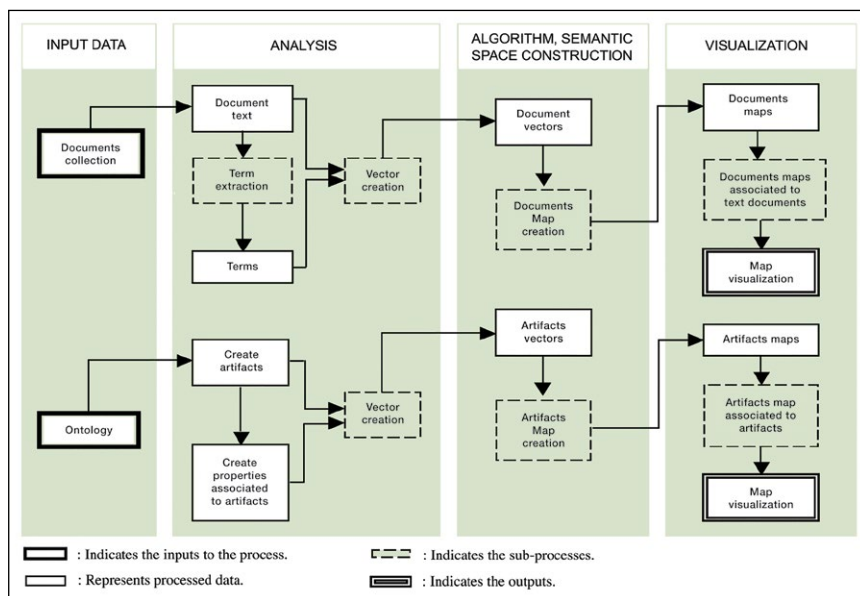
La Figura 2 muestra otro ejemplo de un experimento que hace uso de técnicas de visualización. En este caso se comienza por crear una ontología de propiedades basada en las distintas partes de los cuerpos de los monstruos de la *Carta marina*, por Olaus Magnus (1539). Este mapa, que se considera por muchos como uno de los primeros trabajos cartográficos que utilizan métodos empíricos, está repleto de figuras monstruosas, las cuales Magnus utiliza para recalcar las maravillas de esta región (Rossby). Esta ontología de propiedades se utiliza para crear un mapa auto-organizado SOM de las criaturas. Las criaturas se agrupan a sí mismas en núcleos (*clusters*) de acuerdo a las propiedades que comparten. A manera de ilustración, la visualización también muestra las esculturas creadas por el artista Alexander Reichstein para su pieza de arte interactiva *Bestiarium Construendum* (Reichstein).

Figura 2. Mapa auto-organizado (SOM) que muestra un experimento de como una lista de propiedades se usa para describir los monstruos en el mapa, la *Carta marina* de Olaus Magnus, 1539.



En la Figura 3 se muestran dos métodos de análisis que se desarrollaron durante el proyecto. Con estos métodos se pretendía crear índices archivísticos en los cuales las categorías fueran el resultado de las descripciones de los artefactos en el archivo. Estas dos metodologías parten de filosofías muy distintas. En el primero se hace uso del Latent Semantic Indexing (LSI) para extraer de los documentos una lista de términos que se utilizan como categorías para organizar la data. En el segundo método, el experto (o investigador) usa el programa para crear un espacio vector en donde se insertan las categorías que se han de utilizar para describir cada artefacto. Subsiguientemente se define - de manera binaria - si cada propiedad se aplica o no a cada artefacto.

Figura 3. Las figuras ilustran los métodos de análisis utilizados para organizar los datos en la Figura 1 y la Figura 2.



Ambos experimentos presentan una situación en la cual se introduce a un sistema artificial conocimientos (en forma de ontologías) que ya son implícitos para el ser humano. Los monstruos de la *Carta marina*

son tal porque como criaturas híbridas combinan entre sí las características de distintas especies: centauro es mitad caballo, mitad hombre, la sirena no tiene dos piernas, pero sí tiene una cola.

En ambos métodos, el resultado de procesar datos es una representación visual (llamado mapa) de data organizada, de acuerdo con el principio de semejanza (Avilés Collao et al.). La importancia de este tipo de representaciones gráficas de información yace en la capacidad que se vislumbra de poder ir más allá de las meras apariencias visuales. El poder utilizar estos mapas como instrumentos interpretativos para manejar grandes bases de datos de manera que haya correspondencias equivalentes entre los datos y las formas es uno de los objetivos. Más allá de esta tarea, en el horizonte se puede vislumbrar el desarrollo de una actividad parecida a la lectura, ya sea en la pantalla, en la imagen impresa, o hasta en una proyección holográfica. Todavía más adelante, está una posible interactividad en tiempo real. Esta interactividad, pudiese permitir al usuario una investigación abierta y natural tal y como sucede cuando pasamos las páginas de un libro a las que se accede con facilidad.

LA NUEVA ERA DEL ZETTABYTE

Terminado el proyecto en el 2004 continuamos trabajando con otros proyectos en el área de la archivística digital. (Referimos al lector a las páginas de Systems of Representation, grupo de investigación en la red). Realmente parece no haber sido hasta finales de la primera década del segundo milenio que, posiblemente, a través de la incrementación vertiginosa de datos que amenazan inútilmente nuestra capacidad de entendimiento, surge de nuevo un interés en materias relacionadas con las redes neuronales como el SOM.

El significado de la palabra era indica un punto de partida en una cronología o un período en la historia de una civilización que se caracteriza por un nuevo orden y que generalmente comienza con un suceso importante o notable. En el año 1986 el Internet albergaba 2.6 *exabytes* (EB) de información. Para el año 2000, son 54 *exabytes* una suma que se supera cuatro veces con los 295 *exabytes* en el 2007. En el año 2012 el tráfico global excede la marca del *zettabyte* (1021

bytes). Claramente la cantidad de data que se recoge sobrepasa la capacidad de análisis del ser humano. La plataforma DOMO, por ejemplo, ofrece estadísticas actualizadas sobre la inmensa cantidad de data que se mueve en internet. Sus visualizaciones, Data that Never Sleeps, se han convertido en un género de por sí.

Dentro de este contexto, Boden explica como el modelo conexionista para la inteligencia artificial que se basa en el uso de redes neuronales para el análisis de grandes repositorios de datos adquiere nueva relevancia. Cabe resaltar los experimentos en áreas relacionadas con el reconocimiento de imágenes. Por ejemplo, en el Image Net Large Scale Visual Recognition Challenge (ILSVRC), una competencia auspiciada por el proyecto ImageNet que, desde el 2012, se celebra anualmente, para el año 2017 reportaba resultados en los cuales, las redes neuronales podían, en el 97.7 por ciento de los casos, acertar en la categoría correcta de un objeto mostrado a través de una imagen.

En una publicación del 2018, *Artificial Intelligence, A European Perspective*, el Joint Research Centre de la Unión Europea en Ispra declara:

Aunque la inteligencia artificial lleva una larga trayectoria, el progreso reciente se debe a la convergencia de un incremento en la capacidad de computación, la accesibilidad a grandes bases de datos y la existencia de nuevos algoritmos para el procesamiento de datos. (Joint Research Centre 2018) (Traducción de la autora.)

Entonces, se percibe en estas tecnologías una posible respuesta a la creciente y continua masa de datos. Se anticipa que la máquina con su inmensa capacidad para manejar, organizar y estructurar los datos dispares pueda percibir y hacer inteligibles patrones de significado inherentes en las corrientes de información, que fluyen a través del planeta, día a día. Lejos de ser un sueño, existe una necesidad pujante, de una interfaz que permita un acceso fácil, intuitivo e interactivo a esa *vorágine de datos*. ¿Cuán cerca estamos de que ese objetivo se vuelva realidad?

NARRATIVAS PARA UNA MITOLOGÍA

Podemos estudiar el *software* para entender que cambios nos sobrevienen a raíz de su uso en la vida real. Pero si se queremos entender su impacto social y cultural, debemos también prestar atención a las diferentes narrativas que circulan en su entorno. Entre las historias populares referentes a la inteligencia artificial están aquellas que coinciden con la existencia de distintos órdenes del ser (y los derechos que esto conlleva) en distintos sistemas políticos y económicos. En la República Popular de China, por ejemplo, la falta de privacidad individual no se considera tan importante y es común que la genta acceda a que se utilicen programas de reconocimiento facial en lugares públicos y en tareas tan banales como ordenar la comida en un establecimiento comercial. Se estima que, contrario a la sociedad occidental, la china se orienta hacia los grupos, por tanto, la identidad personal no se valora tanto (O'Meara 2019). Sin embargo, el alcance de las tecnologías digitales en la vida privada de un ciudadano de Europa asume otras dimensiones. La valoración del individuo como entidad y con derechos, como el de ser olvidado, es uno de los factores que ha llevado a la Unión Europea a la formulación de leyes para la regulación y protección del uso de datos. Esta disyuntiva va al corazón de muchas de las técnicas de inteligencia artificial y la opacidad que existe en el uso de algoritmos como parte del proceso de análisis de datos.

Otra de las narrativas que tiene mucha resonancia, es la que atribuye un cambio en el rol que juega la tecnología cuando pasa de ser un asistente para convertirse en sustituto y maestro del ser humano. Los recientes desarrollos, cuya escala (como es el caso de los grandes repositorios de datos) va más allá del alcance humano y requiere del uso de dispositivos mecánicos para su organización y procesamiento, a los cuales se tildan de 'inteligentes', se leen como manifestación y prueba de la realidad de estas narrativas. En su artículo "La inteligencia artificial y el problema de ontoteología digital, consideraciones sobre la ciencia ficción y la teoría", Crombez y Dahms resaltan la existencia de fenómenos como la *apofenia* o la "...experiencia consistente en ver patrones, conexiones o ambos en sucesos aleatorios o en datos sin sentido". En su análisis de estas narrativas señalan la importancia de

un examen crítico acerca de las bases ontológicas de la inteligencia artificial y proponen a su vez una nueva lectura sobre relación de la teoría científica con la ciencia ficción: “La teoría no solamente debe guiar la investigación por el ser humano sino también la investigación que se conduce por la máquina para entender mejor nuestra relación con los objetos que creamos”. Dicho de otra manera, cómo se debe caracterizar nuestra reflexión ante ideas como las de que el análisis de la máquina puede facilitar la creación de nuevas pinturas de Rembrandt, como lo propone el proyecto del Next Rembrandt o que una reconstrucción digital ha facilitado el habla a una momia del periodo del faraón Ramses XI (c.1099-1069 a. C.). Como si el cuerpo humano fuese un simple recipiente o una probeta de laboratorio. Como si los quehaceres del día a día no tuviesen relación con esa masa de carne y espíritu que se encoge y se moldea para ajustarse a su entorno, y reflejar a su vez los accidentes y torpezas que marcan sus vivencias.

EL USO CREATIVO DE LAS REDES NEURONALES

Como gráfica, el SOM es una representación visual que, con base en el principio de semejanza, destaca la similitud (ya sea conceptual o material) entre los diversos elementos de un grupo en un espacio multidimensional. El SOM mantiene las relaciones de cercanía entre vectores que son contiguos en el espacio multidimensional y hace posible el desplegar esta información en forma visual. Marín Díaz-araque describe el SOM como “en realidad, un tipo de algoritmo para clasificar observaciones... [y] conceptualmente similar al MDS [escalamiento multidimensional] que transforma observaciones similares en puntos cercanos del espacio bidimensional.”

En esta sección detallamos brevemente otros proyectos que hemos realizado con el mapa auto organizado (SOM), como instrumento de diseño. Estos proyectos se llevaron a cabo durante los años 2012 al 2018. Incluyen una serie de talleres sobre visualización y la curación de una exhibición en la red para SIGGRAPH. Se presentan en este ensayo con la idea de resaltar el hecho de que, aparte de las más reconocidas, hay muchas maneras en que pueden utilizarse las formas tecnológicas relacionadas con la inteligencia artificial.

TALLERES SOBRE VISUALIZACIÓN

Durante los años 2000 al 2016 tuve la oportunidad de desarrollar e implementar un área de estudios sobre la visualización en el Media Lab Helsinki, lugar donde cursé mis estudios de doctorado en diseño y en donde he sido investigadora y profesora durante los últimos veinte y cinco años. En este programa, que eventualmente pasó a ser un programa de maestría en el área de comunicación visual del Departamento de Media impartí clases en tópicos como diseño de información, visualización y seminario sobre diseño de información. A partir del año 2012, comenzamos a utilizar las herramientas que habíamos desarrollado en el proyecto CIPHER. Esta iniciativa coincide con la presencia de un nuevo miembro experto en computación lingüística, quien brinda apoyo técnico y pone al alcance de nuestros estudiantes de maestría estas herramientas, consideradas instrumentos para el procesamiento avanzado de datos (*advanced information processing tools*). La forma particular en que se han usado es para enseñar principios básicos sobre la organización de la información.

Este objetivo fue implementado a través del taller de visualización, en el cual los participantes desarrollan nuevos conceptos acerca del uso de la información y crean presentaciones descriptivas con gráficas creadas con el SOM, como parte de una narrativa.

El tópico de la información como género y materia es demasiado amplio para tratar en un ensayo. Suficiente cabe mencionar la discusión elaborada por la profesora emerita Elaine Svenonius, la cual presenta una definición que con base en la explicación del diccionario y en los trabajos de S.R. Ranganathan y Seymour Lubetsky, realza el origen de la información en el acto de la descripción, a través de la palabra y el texto como algo que se escribe (ya sea un libro, un reporte, o una carta) y simultáneamente la contrasta con su representación física (una moneda, una piedra). Esta última puede también comunicar el pensamiento por medio del uso de las marcas o símbolos. De acuerdo a Svenonius, “en la organización de información para facilitar acceso, se requieren no solamente descripciones textuales sino también conocimiento sobre el entorno físico y los mecanismos a través de la cual esta se incorpora...” (Svenonius 2000, 8-9).

Las discusiones en este taller se enfocaron, tanto en los aspectos teóricos como en la práctica, a través de ejercicios. Entre los conocimientos teóricos consultados resalta el trabajo del experto en infografía Jacques Bertin, cuya obra, *La semiología de las gráficas* ha inspirado a numerosos peritos como el profesor Edward Tufte, quien a su vez es autor de diversos trabajos importantes. Conocimientos acerca de la diferencia entre los procesos para recoger datos y desplegarlos, ya sea en forma de una serie de tiempo o en una red, por ejemplo, y cómo se pueden usar estas representaciones gráficas en una narrativa, lo que también formó parte de los objetivos de aprendizaje para los participantes. En este contexto el uso de gráficas informáticas, basadas en redes neuronales como el SOM, significó la añadidura de un género más.

En los talleres se utilizaron también métodos como el trabajo independiente, en el cual se organiza a los participantes en grupos, con distintas destrezas. Esta asimetría de conocimientos puede fomentar la colaboración y la creatividad. Durante el año 2015, en uno de los equipos de trabajo, por ejemplo, había un estudiante en diseño gráfico, otro del programa de diseño de juegos, una programadora y una experta en estudios urbanos. Todos los participantes cursaban estudios de maestría. Este equipo optó por investigar cómo se percibe la ciudad de Helsinki a través de las películas. Utilizaron distintas fuentes de información, seleccionaron nueve películas y analizaron las representaciones de la ciudad, para crear un mapa en el que mostraban los sentimientos que la ciudad puede suscitar en un espectador.

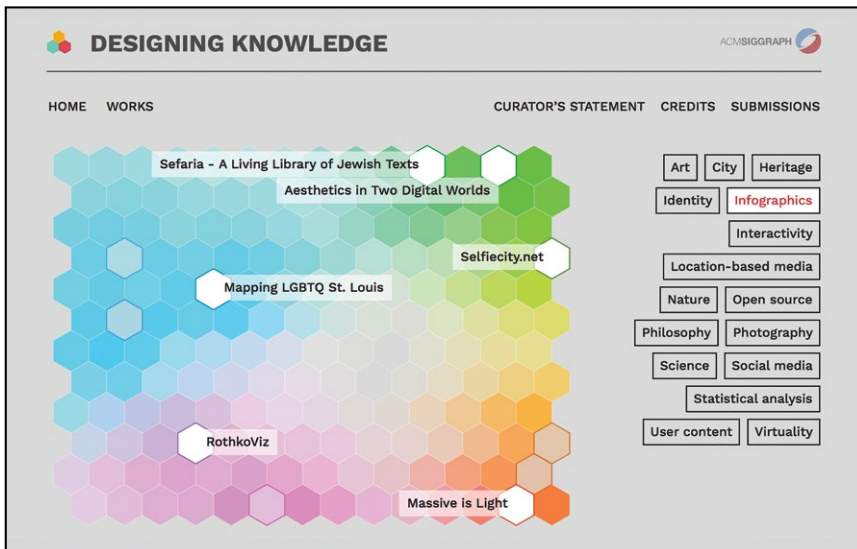
¿Qué tipo de datos se pueden utilizar en este formato gráfico? ¿Cómo hay que recoger la data para poder codificarla y desplegarla? ¿Qué posibilidades de comunicación puede ofrecer el SOM para los usuarios que no tienen experiencia con este tipo de instrumento? ¿Es posible utilizar la capacidad organizadora del SOM como herramienta de creatividad? Estos tópicos eran parte de las discusiones entre docentes y estudiantes durante el transcurso de estas actividades.

EXPOSICIÓN DISEÑANDO EL CONOCIMIENTO

En este apartado queremos hacer mención de nuestra exposición *Diseñando el conocimiento (Designing Knowledge)*, para el Association

of Computing Machinery (ACM) SIGGRAPH. Esta exhibición incluye una muestra internacional de once trabajos que de una manera artística (y técnica) nos muestran distintos aspectos y usos de la información para crear y diseminar el conocimiento. Ya sea la documentación y exposición de una historia escondida (Mapping LBG7Q Saint Louis), o para hacer asequible los textos que en su forma material no se pueden consultar de manera fácil, como lo son las distintas secciones del *Talmud: Sefaria* (a living library of Jewish texts), o para establecer una comparación entre los ciudadanos de una urbe y la otra (*Selfiecity.net*).

Figura 4. Diseñado el conocimiento (*Designing Knowledge*), portada de la exhibición hecha en 2019, para ACM SIGGRAPH.



En la introducción a la exhibición se muestran los trabajos a través de un pequeño mapa de SOM, el cual enseña las distintas categorías que se pueden apreciar en las obras seleccionadas. Al mover el cursor encima del mapa se revelan las distintas características de cada pieza, lo que permite al usuario establecer una comparación global del contenido de la exhibición. Por ejemplo, *Tracking Transcience*, una

pieza realizada por el artista Hasan Elahi (USA) despliega las categorías de arte, identidad, medios basados en la ubicación, *social media* y contenido hecho por el usuario, que en este caso es también el autor del trabajo. Además se puede observar como las distintas piezas comparten las propiedades asignadas.

En el apartado de cada pieza se exhibe un video que muestra como funciona el repositorio y un diagrama que explica la arquitectura de información (*information architecture*) del trabajo. De esta manera, se pretenden destacar los aspectos de la artesanía digital, del diseño y del entorno material, en el cual se desarrolló cada trabajo. En el archivo audiovisual *RothkoViz*, por ejemplo, el diagrama explica claramente las distintas tareas que se utilizaron para procesar las imágenes y crear visualizaciones que realzan las características de las pinturas del pintor norteamericano, Mark Rothko.

CONCLUSIONES

La inmensa cantidad de datos que se generan a través de la vida diaria es una de las características de la Sociedad de la Información (SI). Existe una gran necesidad de encontrar métodos que nos permitan trabajar con la data de una forma responsable e inclusiva, en la cual también se use la capacidad creativa del ser humano. Una mirada antropológica al uso de las tecnologías de inteligencia artificial debe de tomar en cuenta no sólo la eficacia y los resultados producidos, sino la manera en que estas se sitúan y contribuyen en el contexto de la actividad y la cultura humana.

BIBLIOGRAFÍA

Artificial Intelligence, A European Perspective. Reporte especial del Joint Research Centre (JRC), Luxembourg: Publications Office of the European Union, 2018.

Avilés Collao, J., Díaz, L., Kaipainen, M., and Pietarila, J. "Soft Ontologies and Similarity Cluster Tools to Facilitate Exploration and Discovery of Cultural Heritage Resources". *Proceedings of DEXA 2003, Fourteenth International Workshop on Database and Expert Systems Applications*, Los Alamitos, CA: IEEE Computer Society, 75-80.

Bertin J. *Semiology of Graphics, Diagrams, Networks, Maps*. Redlands, California: Esri Press, 1967, (2010).

Boden, M. *A Very Short Introduction to Artificial Intelligence*, Oxford, UK: Oxford University Press, 2016.

Crombez, J., Dahms Harry F. "Artificial Intelligence and the Problem of Digital Ontotheology: Toward a Critical Rethinking of Science Fiction as Theory", *Bulletin of Science, Technology & Society*, 2015, Vol. 35(3-4)110.

Domo, *Data that Never Sleeps*. <https://www.domo.com/learn/data-never-sleeps-7>. Disponible el 9 de enero de 2020.

Designing Knowledge, An online exhibition. Disponible el 9 de febrero de 2020. <https://designing-knowledge.siggraph.org/wp/>.

Díaz, L. & Kaipainen, M. "Designing Vector-Based Ontologies: Can Technology Empower Open Interpretation of Cultural Heritage Objects?", *Proceedings of DEXA 2002, 13th International Workshop on Database and Expert Systems Applications*, Los Alamitos, CA: IEEE Computer Society, 521-525.

Diaz, L. and Partanen, L. “Digital culture heritage to support novel activities in the classroom of the future”, K. Mäkitalo-Siegl, F. Kaplan, J. Zottmann, & F. Fischer (Eds.), *Classroom of the Future: Orchestrating collaborative spaces*. The Netherlands: Sense Publishers, 2010, pp. 181-198.

Diazaraque, Marín, JM. *Los mapas auto-organizados de Kobonen (SOM)*. Disponible el 1 de enero de 2020. <http://halweb.uc3m.es/esp/Personal/personas/jm-marin/esp/DM/tema5dm.pdf>.

Honkela, Timo. 1997. *Self Organizing Maps in Natural Language Processing, Dissertation*, Espoo, Finland: Helsinki University of Technology, Neural Networks Research Center, 1997. <http://www.cis.hut.fi/~tho/thesis/>, (19/01/2020).

Image Net Large Scale Visual Recognition Challenge, ILSRVC. Disponible el 25 de enero de 2020. <http://www.image-net.org/challenges/LSVRC/>.

Image Net Project, Disponible el 25 de enero de 2020. <http://image-net.org/about-overview>.

Joint Research Centre (2018). *Artificial Intelligence, A European Perspective*.

O'Meara, S., “Will China Overtake the U.S. in Artificial Intelligence Research by 2030?”, *Nature*, 21 August 2019.

The Next Rembrandt, Disponible el 28 de enero de 2020. <https://www.nextrembrandt.com>.

Reichstein, Alexander, *Bestiarium Construendum*, Disponible el 9 de septiembre de 2020. <http://www.reichstein.name/best.html>.

Rossby, T. H. "Ocean Eddies in the 1539 Carta Marina by Olaus Magnus", *Oceanography*, 14(4) 77-85, 2003. Disponible en: https://tos.org/oceanography/assets/docs/16-4_rossby.pdf, (09/02/2020).

Svenonius, E. *The Intellectual Foundation of Information Organization*, Cambridge, MA: The MIT Press, 2000.

Tufte, E. *The Visual Display of Quantitative Information*, Connecticut: Graphic Press, 2001.

Big Data between privacy and copyright

AMELIA BONGARZONE

INTRODUCTION

The possibility provided by Big Data to develop interpretative, analytical and even predictive models of human events and behaviours, difficult to imagine until a few years ago, draws the attention of private law and prompts reflection on a problem which, due to its not only patrimonial implications, is assuming exponential proportions, in many ways out of control. The game is played between needs certainly deserving of protection, which however, in their meeting and not infrequently collide, call into question an eurythmy not easy to reach. The availability of a large quantity of information, while proving to be a source of progress and innovation in terms of public and private benefits, also highlights the exigency to protect compromised legal situations. Among these, in addition to the freedom of economic initiative and all that this entails, the rights of personality stand out.

Opportunities for improvement in the public sector (e.g. in the areas of security, prevention, strengthening of services, promotion of life's quality, health, culture, etc.), as well as the indisputable competitive advantages for private entrepreneurship (of launching, consolidation *et similia*, in the reference markets or in those not yet explored) are counterbalanced by the risks of abuse perpetrated against individuals who let their data more or less unconsciously and in exchange

for only apparently free services (Resta and Zeno-Zencovich 2018, 422). The benefits are certainly as many as the critical profiles and the sense of their proportion is given back by the sometimes insidious methods of collection, by the often uncertain places of conservation, by the not always transparent criteria of selection and analysis as even the frequently unpredictable treatments, whose description is made even more evanescent by the possible regeneration of the acquired data and by the continuous refinement of the technologies able to use them.

Faced with a phenomenon that is undergoing such rapid transformation and of transversal importance, the jurist is called upon to question himself on the adequacy and efficiency of the current regulatory framework, without indulging in dogmatic positions, in the awareness of the relativity of the institutes and the precariousness of the experiences. Reason can be found in the idea of law as a structure, “conditioned by economic-social relations” and “in turn conditioning the wider and more complex reality of which it is historically an integral part” (Perlingieri 2006, 161).

ABSENCE OF A CONVENTIONALLY ACCEPTED DEFINITION OF BIG DATA AND ATTEMPT TO DESCRIBE THEM

The perspective of investigation opens the field to a complex panorama whose difficulties can be seen, at first glance, where we reflect on the absence of a conventionally accepted definition of Big Data, an obvious symptom of a “chaotic state of the art” (De Mauro, Greco and Grimaldi 2016, 128).

In the heterogeneity of the proposals, the reference can be to “the collection, analysis and the recurring accumulation of large amounts of data, including personal data, from a variety of sources, which are subject to automatic processing by computer algorithms and advanced data-processing techniques using both stored and streamed data in order to generate certain correlations, trends and patterns” (European Parliament resolution 2017, Recital A). It is, of course, a very composite process, whose phases (collection, storage, aggregation, analysis, comparison, use and reuse) are developed in a variety

of contexts and engage a plurality of stakeholders which, each for their own expertise, participate to the achievement of the final result, usually delivering an intermediate one.

The constant is constituted by the *data*, undisputed protagonists of a historical moment, threatened by the spread of the so-called *surveillance capitalism* (Zuboff 2019; Foster and McChesney 2014). The highly evocative syntagm alludes to a new economic order that, through the observation of the entirety of individual and collective behaviours, is capable of allowing anticipatory, probabilistic indications of the preferences of the subsidiaries and, even more so, of directing their interests and choices, in a perspective of reification of human experience, reduced to a precious commodity of exchange.

Upstream, there is a transfer of data by users, mainly consumers, who provide them immediately (e.g. by adhering to very common loyalty programs such as customer cards of the supermarket) or generate them, not infrequently passively, through the use of devices (Internet of things) or through online access, enticed in most cases by the use of services and benefits, otherwise unavailable, in the face of which the initial distrust is bound to succumb. Thus, with a mechanism as much underestimated as pervasive, often supported by a sort of coercion, not even particularly subtle, based on the *do ut des* (I give you the advantage if you give me the data), a considerable impact on the freedoms and fundamental rights of the person is consumed.

INCIDENCE OF BIG DATA ON PERSONALITY RIGHTS. OBSOLESCENCE OF THE DISTINCTION BETWEEN PERSONAL AND NON-PERSONAL DATA. RELEVANCE OF DATA

The European approach to the problem of Big Data, originally focused on promoting the transition process to a data-driven economy capable of competing with the world's giants, soon became aware of the danger looming over the existential sphere of the affiliates, highlighting the need for a new, adequate, empowerment able to contain otherwise devastating outcomes. In this direction is placed the European Regulation 2016/679, the GDPR, for the protection of personal data, issued in an attempt to combine the improvement possibilities,

on several fronts now indispensable, offered by the strategic asset of data, with the need to preserve human existence, in some of its essential values, from the invasive use of a certain technology.

The balancing is demonstrated by the definitive emancipation of the right to protection of personal data from proprietary paradigms through the confirmation that it is a fundamental right of natural persons (Rodotà 1991, 526, 531; Zeno-Zencovich and Giannone Codiglione 2016, 33; Directive EU 2019/770, Recital 24), according to the Article 1 of the GDPR, and by the simultaneous provision of free movement of data within the Union, in a prospect of cooperation and exchange, evidently functional to digital progress. In the same direction also leads Recital 4, stating that the “processing of personal data *should be* designed to serve mankind” and specifying that “the right to the protection of personal data is not an absolute right; it must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality”.

It does not seem superfluous to point out that the use of the conditional *should be* is significant, where, unless we want to consider its randomness, it reveals the dissatisfaction for a proposed objective but, evidently, not yet achieved. This issue has a primary importance, already clearly indicated in the choice of the legislative form adopted. The use of the Regulation, because of its general scope, the mandatory nature of its elements and its direct application in each Member State, clearly reflects the necessity to ensure uniform levels of protection in the EU.

In 2018, the European Regulation 2018/1807 on non-personal data joins the GDPR. Once again, the European legislature establishes the principle of free movement within the Union, unless public security reasons justify a restriction or prohibition, and points out that, in any event, “the legal framework on the protection of natural persons with regard to the processing of personal data, and on respect for private life and the protection of personal data in electronic communications and in particular Regulation (EU) 2016/679 [...] are not affected by this Regulation” (Regulation EU 2018/1807, Recital 8). A further demonstration of the ongoing balancing of interests.

Personal data and non-personal data, therefore, as the *summa divisio* from which there would appear to be no reason to deviate even for Big Data. However, what criteria for the distinction? The question, the importance of which is made evident by the fact that only the provision of personal data generally requires the consent of the person concerned, would seem not to generate perplexity and to find immediate confirmation in the combined reading of certain provisions of the cited Regulations.

The reference is to Article 4, in point (1), of the GDPR, according to which personal data “means any information relating to an identified or identifiable natural person (‘data subject’)”, with the specification that “an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person”. The notion of non-personal data, instead, is achieved by exclusion, since Article 3 of Regulation (EU) 2018/1807 defines such “data other than personal data as defined in point (1) of Article 4 of Regulation (EU) 2016/679”.

Anyway, if on the definition level the difference is rather easy, the same does not happen in the actual resolution of cases. An overlap, confusion or even a conversion of data is not at all remote, especially when one reflects on the exploratory capacity of technology, the constant development of which advises against the assumption of definitive positions. And indeed, if at present it cannot be radically excluded that anonymous or anonymized data - when involved in a complex process of analysis, perhaps crossed with additional data or simply processed with more refined devices - may lead to a result that allows the tracing or inferring of personal information, even more it is not unlikely that this will happen in the future. Leaving out the possibility that (apparently) neutral data, inasmuch as they are not immediately referable to an “identified or identifiable natural person”, reveal on the contrary their nature of personal data if analyzed from another subject or from another perspective. Consider, for example, of the image of a square that might interest some for the artistic pro-

files of the fountain that is in the center, others perhaps to identify the people present there (AGCOM 2018, 35), or even think of the information obtained from sensors installed on cars which, normally instrumental to an investigation relevant to the object, much can actually reveal about habits, inclinations, preferences of the driver, for example according to the places frequented.

The transversal nature of the contents, which is exponentially amplified for Big Data, therefore undermines the traditional distinction by type of data, voting for its obsolescence. Thus, the opportunity arises to reconsider the issue, starting from data *tout court*.

EXTENSIBILITY TO BIG DATA OF THE PRECAUTIONS FORESEEN IN THE GDPR: CRITIQUE

The proliferation of sources of (personal) data collection, and the intertwining of relations that develop around them, lead us to dwell on one of the possible presuppositions of their circulation: consent (Regulation EU 2016/679, Article 6; Nazzaro 2018, 1248).

Defined in Article 4, in point (11), of the GDPR, as “any freely given, specific, informed and unambiguous indication of the data subject’s wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her”, many other provisions and recitals of the same body of legislation deal with it. In this way, a regulatory plan is outlined to ensure requirements and conditions for a conscious transfer by settlor, based on a genuinely free choice, the recurrence of which should be doubtful if, for example, the person concerned is “unable to refuse or withdraw consent without detriment” (Ibid., Recital 42). The picture is completed by the attention paid to the presupposed act, *id est* the information, which must be expressed “in a concise, transparent, intelligible and easily accessible form, using clear and plain language” (Ibid., Article 12, paragraph 1) on the basis of the principle according to which “it should be transparent to natural persons that personal data concerning them are collected, used, consulted or otherwise processed and to what extent the personal data are or will be processed” (Ibid., Recital 39) and in the wake of which a series of precautions are envisaged.

The discipline certainly seems suitable for a binary model, in which there is a well-identified titular subject, who collects the data of a natural person and processes them, or has them processed by his auxiliaries, for purposes punctually established. A system, in short, in which the owner is immediately able to inform the person concerned of the destination of the data, who, in turn, is in the position of being able to fully accept or disagree. Otherwise, however, it happens in the Big Data scenario, where, to simplify matters extremely much, it can be said that the data rise up a long chain, only as a result of which they are aggregated and examined by an 'upstream' holder who may not coincide, and usually does not coincide, with the one who initially acquired them. Moreover, before the data are analyzed, it is not certain what correlations and results their processing will reveal, nor can it be reliably established how and for what purpose they will be used in the end. As a result, it is more or less impossible for the settlor to make a conscious choice, as he often lacks even a definite perception of the data he is giving up and although he is neither in a position to ascertain *ex ante* what the destination will be, since, as has been said, it may happen that not even the actors called upon to deal with the relative process of decomposition and composition of the accumulated material know this.

The hypothesis is accredited by the increasingly frequent use of automated decision-making processes, including profiling, in which case the intervention of the artificial intelligence and algorithms, able to handle potentially infinite dimensions, often makes it difficult to identify the interlocutors of the treatment process. In the absence of such essential coordinates, the interested party is effectively precluded from using the remedies predisposed by the GDPR (access, rectification, cancellation, oblivion, limitation, etc.), so that the rules on information and consent show all their inadequacy. What emerges is an asymmetrical knowledge, stigmatized from many sides, fed by the opacity of the processes used for data processing, in place of the much celebrated transparency, and further aggravated by the lack of adequate literacy of the data subject, devoid of the cognitive baggage necessary to understand the complex dynamics involved, certainly beyond the ordinary skills widespread among the population.

A remedy would appear to be provided by Article 35, paragraph 1, of the GDPR, which requires a “data protection impact assessment” to be carried out if it is assumed that “a high risk to the rights and freedoms of natural persons” may result from a type of processing. The aim is to prepare the most suitable measures to avoid the danger, with the intervention, in case of inability to contain it, of the supervisory authority (Ibid., Article 36). However, it should be noted that the application of the mentioned provision, as of the entire GDPR, is limited to personal data concerning natural persons, while very often the trap is consumed in the unpredictable transformation of anonymous or non-personal data that, for example, as a result of combinations with other data become specifically referable to individuals (or small groups, such as the family), sometimes even revealing sensitive aspects (political, religious or sexual orientation, etc.). Among the different risk profiles, this is perhaps the most felt and it is in such cases that the current protection systems show their inefficiency, with worrying repercussions not only on the level of self-determination of the individual - based on a consent that continues to preach, but only formally, as free, specific, informed and unequivocal - but also on the power of control and inhibitory of unauthorized exploitation.

On this point, one reflects, once again, on the various digital interactions where the usability of the contents, allegedly free, is instead radically conditioned to the authorization to the processing of data (Di Porto 2016, 13; Resta and Zeno-Zencovich 2018, 416, 422; Nazario 2018, 1250) which is usually done in a mechanical way and, often, without it being even clear what data is, in fact, available. It should also be borne in mind that in assessing “whether consent is freely given, utmost account shall be taken of whether, *inter alia*, the performance of a contract, including the provision of a service, is conditional on consent to the processing of personal data that is not necessary for the performance of that contract” (Regulation EU 2016/679, Article 7, paragraph 4). It is legitimate to ask oneself, in how many cases a thoughtful investigation would lead to the conclusion of a truly free and conscious transfer.

PROCESSING OF BIG DATA AND OWNERSHIP OF THE RESULT:
COPYRIGHT AND ACQUISITION BY SPECIFICATION

It is just the case to point out that the generic nature of the term *data*, and in particular its ability to tendentially describe every immaterial representation of events and/or conducts, pushes to circumscribe its reference to those elements that are subject to computational use, *id est* suitable to be object of automated or artificial elaboration. In this way, the data constitute an observable entity of which the free movement, provided for by European legislation, participates in the process of legal objectification. Since, in fact, a circulatory event can concern only persons, services or goods, tangible or intangible, it does not appear risky to note that guaranteeing the possibility with regard to the data presupposes their qualification of goods, precisely of intangible goods (Resta 2010, 20).

It is worth clarifying that the single data, considered in isolation, enjoy a value certainly different from the sum of the data of which they are a possible element; just as the sum of the data collected, as components of a whole, is not equal to the value of the whole in its complexity, since the whole is something more than the sum of its parts.

In this perspective, it should be pointed out that a huge amount of data stored is still only the raw material, destined to be processed within a complex process. It is here that the use of the algorithmic instrumentation allows the generation from the inputs received of an *aliquid novi*, different from the single goods which have originated it, with respect to which manifests the problem of ownership and the individuation of the protection mechanisms. In particular, the impossibility of immediately bringing the result obtained back to the original expression of an idea, and therefore to human creativity, raises perplexity as to its inclusion in the scope of works of intellect. It is argued that while the structuring of a database, by virtue of the selection and organization of the content, can lead to the recognition of a copyright of the founder, not so for unstructured databases, typical of Big Data, which would lack the 'creative touch'.

Nevertheless, beyond the reservations aroused by radically negatory positions, it must be considered that the composition of a databa-

se involves substantial investments, not even neglected by the Italian law on copyright (l. 633 of 1941). From an antiparasitic point of view, Article 102 *bis* grants the maker of the database the right to prohibit, for a certain period, “the operations of extraction or re-utilization of all or a substantial part of the same”, “regardless of whether the database can be protected under copyright or other rights and without prejudice to the rights to the content or parts thereof”.

It is questionable whether, and in what terms, such cautions can be extended to Big Data, without affecting the superior observation according to which one thing is the database, what set of data, another thing is the product that from the processing of those data is obtained. The transformation of the data, constituting the original material, into a *res nova* having a distinct socio-economic individuality, and therefore the transformation of goods into goods with added value, evokes in some way the rule contained in Article 940 of Italian Civil Code, whose foundation is found in the principle of ownership of the goods to those who produce them, even if the material used belongs to others, unless the value of the latter greatly exceed that of the processing. The point would deserve to be examined in more detail, nevertheless, the boundaries entrusted to the present reflections only allow to mention it.

ESSENTIAL BIBLIOGRAPHY

AGCOM, “Big data. Interim report nell’ambito dell’indagine conoscitiva di cui alla delibera n. 217/17/CONS”, June 2018. <http://www.agcom.it>.

De Mauro, Andrea, Marco Greco, and Michele Grimaldi. “A formal definition of Big Data based on its essential features”. *Library Review*, 65 (March 2016): 122-135. doi: 10.1108/LR-06-2015-0061.

Di Porto, Fabiana, “La rivoluzione big data. Un’introduzione”. In *Big data e concorrenza* edited by Di Porto Fabiana, *Concorrenza e mercato*, no. 23 (2016): 5-14.

Directive (EU) 2019/770 of the European Parliament and of the Council of 20 May 2019 on certain aspects concerning contracts for the supply of digital content and digital services. *Official Journal of the European Union*, May 22, 2019: L 136/1-L 136/27.

European Parliament resolution of 14 March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI)). <http://www.europarl.europa.eu>.

Foster, John Bellamy, and Robert W. McChesney, "Surveillance Capitalism. Monopoly-Finance Capital, the Military-Industrial Complex, and the Digital Age". *Monthly Review. An independent socialist magazine*, Jul 01, 2014. <http://www.monthlyreview.org>.

Nazzaro, Anna Carla, "L'utilizzo dei *Big data* e i problemi di tutela della persona". *Rassegna di diritto civile*, no. 4 (2018): 1239-1260.

Perlingieri, Pietro. *Il diritto civile nella legalità costituzionale secondo il sistema italo-comunitario delle fonti*. 3th ed., Napoli: Edizioni Scientifiche Italiane, 2006.

Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union. *Official Journal of the European Union*, November 28, 2018: L 303/59-L 303/68.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, May 4, 2016: L 119/1-L 119/88.

Resta, Giorgio and Vincenzo Zeno-Zencovich, "Volontà e consenso nella fruizione dei servizi in rete". *Rivista trimestrale di diritto e procedura civile*, no. 2 (2018): 411- 440.

_____. "Nuovi beni immateriali e *numerus clausus* dei diritti esclusivi". In *Diritti esclusivi e nuovi beni immateriali* edited by Resta Giorgio, Milanofiori Assago (MI): Wolters Kluwer Italia S.r.l, 2010: 3-72.

Rodotà, Stefano, "Privacy e costruzione della sfera privata. Ipotesi e prospettive". *Politica del diritto*, no. 4 (December 1991): 521-546.

Zeno-Zencovich, Vincenzo and Giorgio Giannone Codiglione, "Ten legal perspectives on the "big data revolution"". In *Big data e concorrenza* edited by Di Porto Fabiana, Concorrenza e mercato, no. 23 (2016): 29-57.

Zuboff, Shoshana. *The age of surveillance capitalism. The Fight for the Future at the New Frontier of Power*. London: Profile Books Ltd, 2019.

Una estrategia para gestionar los big data en archivos sonoros y audiovisuales

ANTONELLA BONGARZONE
Universidad Magna Graecia de Catanzaro

INTRODUCCIÓN

Los documentos de audio y video son más vulnerables a sufrir daños, en comparación con los que tienen, como soporte, el papel. Por lo que, a menudo están sujetos a pérdidas, a causa de la alta densidad de información.

Muchos de estos materiales están reunidos en colecciones y archivos que representan una fuente indispensable para la memoria personal y colectiva y que constituyen una parte fundamental del patrimonio cultural del siglo XX.

Los documentos multimedia pueden incluirse en dos categorías generales de referencia: los documentos estáticos, que son independientes de la variabilidad temporal; y los documentos dinámicos que, por el contrario, dependen de la variabilidad temporal y, por tanto, necesitan mecanismos de sincronización para ser utilizados correctamente.

Son documentos estáticos las imágenes fotográficas analógicas, los documentos que contienen texto y los documentos gráficos; en cambio, son documentos dinámicos, los documentos audiovisuales y las grabaciones sonoras.

En la web, el principio que rige no es el de la jerarquía y la autoridad, sino el de la conexión. En efecto, no existe una separación en la web entre libros, archivos e imágenes almacenadas en diferentes instituciones, pero es posible identificar las líneas temáticas en torno

a las cuales se pueden reconstruir contextos de objetos e informaciones de diferente naturaleza.

EL ALMACENAMIENTO DE CONTENIDOS EN SOPORTES DIGITALES

Como se sabe, la revolución informática ha experimentado varias fases, en particular, en lo que se refiere a los archivos. La primera ha sido la acumulación de datos en bases de datos, luego la digitalización que funcionó de manera aleatoria, sin hipótesis precisas de diseño, lo que no siempre permitió la visibilidad de los documentos en contextos digitales diferentes.

La preservación digital es una tarea global, continua y compleja, donde factores físicos y lógicos de la información (Romero 2006), permanencia y acceso del contenido (Voutssás 2009), aseguran su mantenimiento y uso a largo plazo, por medio de políticas de conservación y seguridad informática (Termens 2013). La preservación digital de archivos sonoros y visuales puede ser entendida como el “método sustentable a través del cual se conserva, administra, gestiona y se proporciona acceso, difusión y reaprovechamiento permanente” (Rodríguez 2016a, 216).

Desde la creación de sistemas de almacenamiento masivo digital (DMSS -Digital Mass Storage Systems) ha sido necesario establecer los estándares de metadatos para la descripción de contenidos multimedia y luego proteger las informaciones a través de la integridad de los medios y los datos contenidos, que en este último caso deben garantizar el acceso; no siempre libre, pero donde sea necesario, limitado a la consulta y a la reproducción gratuita.

El documento producido debe ser confiable, auténtico, accesible e íntegro.

Ello tiene que poseer los requisitos objetivos de calidad, seguridad, integridad e inmutabilidad.

La inmutabilidad, en particular, es la característica que hace que el documento no se altere en la forma y en el contenido a lo largo de todo el ciclo de gestión y garantiza su retención estática. (Bongarzone 2019, 79-80).

Como ha dicho Perla Olivia Rodríguez Reséndiz:

Ante la ausencia de un modelo de repositorio digital que haga frente a los factores de riesgo de la preservación digital y que a su vez sea confiable y sustentable, en la última década se ha incrementado el interés por contar con un marco de referencia para la preservación digital a largo plazo. Por ello, se han comenzado a utilizar modelos y estándares internacionales como el OAIS (Open Archival Information System), diseñado como un modelo para la creación de un sistema de información de archivo abierto (Rodríguez 2016c, 200).

En el modelo de referencia OAIS está previsto que, cada objeto digital conservado, sea asociado con un grupo de metadatos, que permita en el futuro su fácil localización. En el interior de los metadatos se introducen las informaciones sobre la representación (sintáctica y semántica) que proveen todos los datos necesarios para leer e interpretar correctamente la secuencia de *bit* de los objetos archivados (Bongarzzone 2019, 81-82).

Durante mucho tiempo, los archivos sonoros y audiovisuales tuvieron que negociar el espacio en las bóvedas de almacenamiento, porque conforme pasaba el tiempo el acervo se incrementaba. La transferencia de contenidos analógicos a plataformas digitales, a través de la digitalización, cambió el uso de bóvedas de almacenamiento de soportes analógicos por bóvedas de herencia digital (Rodríguez 2016c, 206).

Las soluciones compartidas para almacenar, administrar, mantener, distribuir y preservar un complejo conjunto de objetos digitales heredados, junto con los metadatos relacionados para administrar los archivos, podrían ser el uso de tecnologías innovadoras, como e-infraestructuras, e-servicios, la nube y los micros servicios, así como de tecnologías que no han sido adoptadas en la preservación digital como el *big data* (Rodríguez 2018; Fresa, Justrell y Prandoni 2015).

BIG DATA Y DOCUMENTOS DINÁMICOS

Lo que se entiende generalmente por *big data* es un conjunto heterogéneo de técnicas para procesar grandes cantidades de datos digitales. Según Robert Kitchin (2013, 1-2) lo que caracteriza el *big data*

es la capacidad de procesar para el análisis un enorme volumen de datos (en *petabytes*), y que estos datos, además, puedan ser generados y tratados a gran velocidad –prácticamente en tiempo real– y estar estructurados de diversas formas o incluso no estar estructurados, y que pueden ser exhaustivos, es decir, que pueden llegar a representar prácticamente la totalidad de las poblaciones o sistemas que se consideren, sin que se pierda su trazabilidad con la fuente original. La característica más relevante de la analítica *big data* no es la cantidad de datos, sino la complejidad de los datos que se pueden manejar y la posibilidad de relacionar distintos tipos de datos de forma agregada y en relación con otras bases de datos. (Ardévol 2017, 17-18).

El objetivo final de la conservación es asegurar no sólo la custodia, la gestión y difusión del documento o colecciones de documentos, sino todas las informaciones de contexto de los metadatos y, sobre todo, de las relaciones entre los documentos que sirven para la reconstrucción del vínculo archivístico (necesario, originario, determinado y, hoy también, virtual) (Bongarzone 2019, 78).

Por lo tanto, el proceso de conservación parece estar bien definido; no se puede decir lo mismo sobre el tratamiento descriptivo de los documentos dinámicos.

Por tratamiento de documentos dinámicos se entienden las actividades de identificación, descripción, almacenamiento, conservación, digitalización, restauración, valorización y gestión de estos materiales.

Un buen sistema de catalogación de los documentos dinámicos debería ser capaz no sólo de describirlos, sino también de documentar, correlacionar, hacer explícitas todas las operaciones relativas a su tratamiento y a su posible uso y reutilización. También debería ser capaz de contextualizar el documento dentro del proceso de producción del que procede.

En el pasado, en la mayoría de los casos, la catalogación de estos documentos se realizaba por medio de una ficha técnica de papel y se limitaba al registro de datos esenciales, como el título y la fecha y, sobre todo, de los datos técnicos o de gestión y almacenamiento.

La entrada de los documentos dinámicos a las bibliotecas y los archivos, ha determinado la atención de disciplinas, como la biblioteconomía

y la archivística a los problemas de describir en particular estos documentos.

Sobre el tratamiento descriptivo de estos documentos, cabe preguntarse:

¿Cómo debe ser la representación del contenido de un documento visual y sonoro? y ¿cuál es la relación, o el vínculo, que tiene un documento visual y/o sonoro con su contexto productivo?

Por lo que se refiere a la catalogación de los documentos dinámicos, las normas de descripción son las elaboradas por FIAF (International Federation of Film Archives) en lo que se refiere a los documentos visuales y las de la IASA (Internacional Association of Sound and Audiovisual Archives) en cuanto a los documentos sonoros y audiovisuales.

Sin embargo, estas normas no resuelven el problema de la descripción semántica del documento y el vínculo entre el documento dinámico con otros documentos pertenecientes al mismo proceso productivo, pero seguramente ofrecen una cierta uniformidad en cuanto a los criterios de descripción general y descripción física de este tipo de documentación.

Los documentos dinámicos son portadores de dos tipos de informaciones, el primero consiste en el contenido sonoro o visual y, el segundo refleja el contexto histórico (no archivístico) en el que se sitúa dicho documento (el entorno histórico, sociológico, antropológico y cultural en el que nace).

Una posición común es que la analítica *big data* considera que los datos que maneja son “descontextualizados” y que la función de las ciencias sociales es “contextualizarlos”. Así, el contexto objetivo, o sea, las informaciones relacionadas de tipo social, histórico y cultural, pueden ampliar las posibilidades de análisis. Solamente estos “datos profundos” (Manovich 2007, 17; Ardévol 2017, 24.) pueden penetrar en las descripciones superficiales, instantáneas e interesantes ofrecidas por los grandes datos.

Al lado del contexto parece siempre necesario el uso de dos fuentes de información para efectuar una gestión integral: las fuentes primarias o documentos en sí mismos y las fuentes secundarias o fichas técnicas, anotaciones y fuentes de consulta (Rodríguez 2012, 117-118).

Por lo tanto, es particularmente útil construir, a través de un lenguaje controlado, un *authority file* o una herramienta constituida por listas de términos, descriptores/palabras clave, además de controladas y normalizadas, con relación entre sí, que identifiquen temas, eventos, conceptos, instituciones, sociedades, lugares y personas.

Gracias a los sistemas de *information retrieval* y a las posibilidades de realizar búsquedas *full text* en todos los campos del formulario desde la intersección de los dos tipos de información, se pueden obtener respuestas exhaustivas a las preguntas de un usuario diversificado y más o menos especializado.

También sería aconsejable tener la posibilidad de una catalogación siempre abierta y actualizable a lo largo del tiempo.

UNA POSIBLE ESTRATEGIA

¿Cómo tratar un documento visual o sonoro?, ¿cómo obra individual, producto acabado, un libro a gestionar, para la conservación, según criterios biblioteconómicos?, ¿o debería considerarse, en muchos casos, un documento de archivo vinculado a otros documentos?, ¿o debe tratarse en el marco de un proceso de producción, dentro de un contexto?

No siempre los catalogadores tienen en cuenta la información relacionada con los documentos visuales y sonoros, vinculada a sus respectivos procesos de producción.

Es indiscutible y parece necesario que exista el vínculo entre los documentos estáticos y dinámicos, puesto que originalmente estaban vinculados al mismo contexto histórico-productivo, donde la descripción multinivel ofrecida por el estándar ISAD(G), permitía uniformar el criterio de almacenamiento.

El contexto de los documentos dinámicos consiste esencialmente en las fases de un proceso de producción específico y de la historia de quienes iniciaron, participaron y realizaron ese proceso (por ejemplo, empresas productoras, organismos públicos, compradores, empresas de distribución, autores, actores, artistas, intérpretes o ejecutantes).

Las reglas de catalogación desarrolladas hasta ahora para la descripción de documentos visuales y sonoros se basan principalmente

en los estándares utilizados en las bibliotecas para la catalogación de materiales que no son libros ISBD(NBM) - International Standard bibliographic description for non book materials.

En este caso, el material visual o sonoro se considera, como un libro, la expresión de una obra concebida para su publicación o, en cualquier caso, para su presentación a un público.

En realidad, el enfoque del documento visual/sonoro debe incluirse en un lugar más amplio o como parte de un conjunto de otros documentos (no necesariamente estáticos, sino también otros documentos dinámicos) con los que está relacionado. Siempre es necesario trazar la historia del documento antes y después de su adquisición por parte del archivo como la descripción del fondo al que pertenece el documento.

En la mayoría de los casos, la redacción del formulario de información de la obra sonora y video se basa en las fuentes secundarias que se consideran autorizadas (tales como registros, publicaciones, escritos y documentos en papel) y se agregan los datos de la descripción física del material.

Para la correcta gestión de los documentos dinámicos es imprescindible también un enfoque museológico, ya que se trata de un bien cultural o bien identitario de una realidad específica.

La necesidad de concebir una única norma internacional descriptiva para los documentos dinámicos podría ser una estrategia para gestionar los archivos visuales y sonoros.

En principio, la catalogación debe estructurarse en cinco macro áreas descriptivas sin parcelación porque creo que demasiados campos descriptivos confunden, no sólo al usuario final, sino también al especialista llamado a la catalogación. ¡Querer describir todo termina con describir nada!

Las macro áreas consideradas podrían ser las siguientes:

1. **Identificación** (o datos generales) o la denominación de la colección y el título de identificación del documento dinámico y su ubicación física.
2. **Estructura**, la descripción física del material, los aspectos relacionados, las técnicas, los lugares y las fechas, los personajes, los intérpretes, el director, etc.

3. **Documentación** y el área de la información relativa a la documentación conexas y complementarias, que permite conectar los archivos descriptivos de los diferentes documentos y fondos de archivo, indicar el material de acompañamiento disponible en el instituto, identificar todas las publicaciones (libros, artículos, sitios web, etc.) utilizadas para la reconstrucción histórica del material descrito.

4. **Acceso** o área de información relacionada con el contexto de origen y las condiciones de acceso. La difusión de documentos dinámicos requiere que conciliemos los principios de privacidad, protección y propiedad intelectual, en relación con el tratamiento general de toda la información que acompaña a los datos alfanuméricos y que se transmite mediante un rango documental visual y sonoro.

5. **Derechos de autor** (*copyright*), en el ámbito de la propiedad intelectual existen restricciones legislativas especiales sobre los sonidos y las imágenes en movimiento. Además de los derechos de autor, de los autores o compositores, se agregan los derechos de los productores e intérpretes.

CONCLUSIONES

Si se introducen elementos especializados de almacenamiento y descripción, y si el dato tiene pleno significado porque se inserta en un contexto, será evidente que las operaciones de almacenamiento digital, ya no concentradas en agregados documentales, sino como complejos homogéneos, predispuestos a las funciones de extracción e interpretación, no se convertirán en un mar de datos incontrolados sino que ofrecerán un vehículo para la circulación del conocimiento.

BIBLIOGRAFÍA

- Ardévol, E. (2017). "Big data y descripción densa". *Virtualis*, 7 (14) <https://www.revistavirtualis.mx/index.php/virtualis/article/view/186>.

- Bachi V.; Fresa A.; Pierotti C.; Prandoni C. (2014). "The Digitization Age: Mass Culture Is Quality Culture. Challenges for Cultural Heritage and Society". En: Ioannides M.; Magnenat-Thalmann N., Fink E., Žarni R., Yen AY., Quak E. (eds). *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection. EuroMed 2014. Lecture Notes in Computer Science*. Cham: Springer, 786-801.
- Bongarzone, A. (2019), "La custodia dei documenti informatici in Italia: il manuale di conservazione". En Rodríguez, P.; Fernández, M.T. (eds). *Conectandi los saberes de bibliotecas, achivos y museos (BAM) en torno a la preservación de documentos analógicos y de origen digital*. México: Instituto de Investigaciones Bibliotecológicas y de la Información, UNAM, 77-91.
- Fresa, A.; Justrell, B.; Prandoni, C. (2015). "Digital curation and quality standards for memory institutions: PREFORMA research Project". *Archival Science*.
- Kitchin, R. (2014). "Big Data, new epistemologies and paradigm shifts". *Big Data & Society*. <https://journals.sagepub.com/doi/full/10.1177/2053951714528481>.
- Manovich, L. (2007). "Cultural analytics: analysis and visualization of large cultural data set". <http://lab.softwarestudies.com/p/cultural-analytics.html>.
- _____. (2011) "Trending: The promises and the challenges of big social data". http://www.manovich.net/DOCS/Manovich_trending_paper.pdf.
- Rodríguez, P. (2012). *El Archivo Sonoro. Fundamentos para la creación de una Fonoteca Nacional*. Mexico: Escuela Nacional de Biblioteconomía y Archivonomía.

- _____. (2016a). *La preservación digital sustentable de archivos sonoros*. México: Instituto de Investigaciones Bibliotecológicas y de la Información, UNAM.
- _____. (2016b). “La preservación digital sonora”. *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*, 68.
- _____. (2016c). “El OAIS en la preservación digital de Archivos sonoros”. *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*, 70.
- _____. (2018). “Factores para la preservación digital sustentable de archivos sonoros”. *Bibliotecas*, 36 (2).
- Romero, J. (2006). “El papel de los metadatos en la preservación digital”. *El profesional de la información*, 15 (2).
- Térmens, M. (2013). *Preservación digital*. Barcelona: Editorial UOC.
- Voutsas, J. (2009). *Preservación del patrimonio documental digital en México*. México: Centro Universitario de Investigaciones Bibliotecológicas, UNAM.

EXPERIENCIAS DE USO Y APLICACIÓN

Using Computational Tools and Experts to Improve Access to Digital Media Archives

KAREN CARIANI
DAVID O IVES

*Executive Director WGBH Media Library
and Archives Boston, MA*

INTRODUCTION

This paper describes projects involving machine learning (ML) and artificial intelligence (AI) to create descriptive metadata for the American Archive of Public Broadcasting (AAPB), and the importance of archives and libraries collaborating with experts in the AI community. There is a challenge to improve access to audio visual collections, and especially the ever-growing digital collections. These collections represent our cultural heritage and need to be seen, used, and available. However, most cultural heritage archives and libraries are not acquiring more funding to manage the new digital collections, just more files and items, and audiovisual (A/V) collections in particular pose special challenges. Digital A/V collections offer great opportunities to use the new technologies for search and access. How can archivists continue to make these important A/V collections available through all the wonderful platforms the new digital technology has to offer? It is necessary to have good descriptive metadata to enhance computer search discoverability or better search tools that can use sound and images. An opportunity exists to collaborate with other professions and disciplines that are expert in computational tools, machine learning and artificial intelligence (AI). There is enough benefit to both areas of expertise, that working together to improve the tools, more can be accomplished. The AI and

machine learning community need good data sets. The archive community has datasets that need to be improved. By working together, the datasets can be improved by improving the quality of the meta-data and thus making the data set more viable for analysis. There is a saying that goes: If you want to go fast, go alone, if you want to go far, go together. This paper will outline AAPB past projects, and AAPB's current work with computational linguistics at Brandeis University through an Andrew W. Mellon Foundation funded project.

BACKGROUND

WGBH is Boston's public television and radio station. WGBH produces fully one third of the content broadcast on the PBS network nationally. WGBH has been broadcasting since 1951 with radio and 1955 with television. The WGBH archive holds about 750,000 items, mostly audio, video or film materials. This does not include the born digital media files which are now created daily. The programming produced includes public affairs, science programming historical documentaries, children's programming and "how to's" like cooking, woodwork, home improvement. This paper focuses on the American Archive of Public Broadcasting (AAPB) collection which is a collaboration between the Library of Congress and WGBH.

The AAPB goal is to coordinate a national effort to preserve and make accessible, as much as possible, rights permitting, historically significant public tv and radio programming. The AAPB is a digital archive with a website at <http://americanarchive.org>. Users anywhere in the US can access a wide range of historical public television and radio programs from the late 1940s to the present. The AAPB supports current stewards of the materials and facilitates the use of historical public broadcasting by researchers, educators, students, and others. The Library is primarily responsible for the long-term preservation of the digital files. WGBH spearheads access and outreach, and together share overall governance, policy, collection development, ingest and access and rights decisions. As an aggregator of content, AAPB hopes to provide a centralized web portal of discovery for public media materials. The collection is growing. Access is for research, educational,

and informational purposes only. Due to rights restrictions, half is available through the On-line Reading Room anywhere in the US. The entire collection of over 100,000 items is available for viewing on location at the Library of Congress and WGBH.

Given the collection, there are, of course, problems and challenges. The material comes from over 100 sources around the US. It has great variety, which is a treasure and a challenge. The content of the collection varies from a single speaker, like a news announcer at a desk with a single microphone, to a man on the street with a heavy accent and background noise, to a musical performance, to foreign language, and potentially all mixed into 1 program. It comes from all regions of the country with people speaking in different accents, dialects, speech patterns, and speed of talking.

The metadata is also variable. Often coming from local public media stations with no archivists or librarians on staff, much of the content has limited metadata. And the AAPB is growing annually with each new collection adding thousands of files, increasing the problem, a common challenge with many audiovisual archives. For example, just looking at a digital audio file tells a researcher nothing about the content. The audio file needs to be played and listen to for possibly an hour or more to determine the content and add appropriate descriptive metadata.

Some items have a transcript which can provide rich data. The transcript can be indexed, and the terms searched. It also allows for an easy read or scan of content, and roughly a time location within the file to reach topics of interest. Obviously, in order to find things in the collection it needs to be described and catalogued, but human cataloguing is too slow, given the volume of material constantly being added. Can machines can do it faster and describe everything quicker with the same accuracy?

Most current search engines (machines) use words to locate an item. To add those appropriate texts and terms, knowledge of the program is necessary. There are thousands if not millions of footage, frames, and audio tracks. It would take a human many years to fully catalogue and describe the AAPB collection. There is a great opportunity to utilize machine learning and computational tools to help

create metadata to improve discoverability. Creating text from the audio, using Natural Language Processing (NLP) tools is an easy start, in addition to identifying people and sounds by matching audio sound waves, and to use crowd sourcing to help correct machine originated errors and verify the machine results.

PROJECT RESULTS

The initial AAPB project collaborated with Pop Up Archive and University of Texas (UT), Austin on an IMLS project to 1) create transcripts using a speech to text tool (Kaldi) for 68,000 items in the AAPB collection, 2) to use other NLP tools to pull out named entities and locations from transcripts and 3) to test use of audio identification tools to identify key speakers in the collection working with UT Austin, High Performance Sound Technologies for Access and Scholarship (HiPSTAS). Crowdsourcing games were created to help correct or fix the computer-generated transcripts.

The results of the audio fingerprinting project with HiPSTAS were mixed. The goal was to use identified speaker soundwaves to find identical sound waves in the AAPB collection to identify speakers, music, applause, laughter, etc. HiPSTAS had a sample size of 4,000 hours (from the AAPB 68,000 item collection). WGBH identified 10 speakers: Hillary Clinton, Bill Clinton, Ronald Reagan, Julia Child, James Baldwin, Malcolm X, Martin Luther King Jr., Lyndon Johnson, Richard Nixon, and Gloria Steinem. These key people were probably in the AAPB collection, and nationally identifiable. The end product was to create a reference database of sound waves, or sound fingerprints, for speakers that could be available for others to use.

The process was the following. WGBH gave HiPSTAS a set of files with the 10 identified speakers. Humans identified the time code where these named people existed in the audio files for a set of 103 hours of content. Supervised learning was used to train the machine to identify those people soundwaves. A specific machine learning model needed to be built for each speaker. There was another set of files that might possibly have those speakers, to see if the machine

tool could get a sound wave match and identify the speakers. A random batch of about 1000 hours was provided to see if the machine could identify the speakers unsupervised.

After 3 years, only 1 speaker model was created for Bill Clinton. HiPSTAS tools were indeed capable of identifying Bill Clinton for a key set of files. However, the compute power and resources for the machine learning alone, in addition to a human tagging a speaker in a sampling of items, building an algorithm specifically for 1 speaker, and then to search for the speaker across a larger set, is not yet a feasible way of identifying speakers in 100,000 items of local tv and radio programs.

However, useful tools were created as a result of this project. (Figure 1) There is a workflow chart, an audio tagging toolkit, a dockerized Jupyter notebook for machine learning tools, 2 notebooks to show how to use the tools using the Bill Clinton model, 2 notebooks to detect music, and other models created. Each effort forward creates more opportunities even if the original goals are not completely fulfilled.

Figure 1.

Tools created for audio fingerprinting

- Workflow charts are available here: https://github.com/hipstas/aapb-data/tree/master/Workflow_flowcharts.
- The team developed tools (available and described at <https://github.com/hipstas>) including an Audio Tagging Toolkit, Audio Labeler, and the Kaldi + Pop Up Archive Docker Image.
- HiPSTAS created a Dockerized Jupyter notebook environment with pre-installed audio machine learning tools, a package we were calling the Audio ML Lab : <https://hub.docker.com/r/hipstas/audio-ml-lab> and <https://github.com/hipstas/audio-ml-lab>.
- Two notebooks to show how to use these tools using a Bill Clinton model here: https://github.com/hipstas/sida/tree/master/Bill_Clinton_template.
- They also created two notebooks for detecting music: https://github.com/hipstas/sida/tree/master/Music_template; and for detecting applause: https://github.com/hipstas/sida/tree/master/Applause_template.
- Other models are here, including those that HiPSTAS produced specifically from AAPB and those they used to create the Kaldi instance: <http://xtra.arloproject.com/datasets/>.

AAPB also worked with Pop Up Archive to create speech to text transcripts using the open source tool Kaldi, that was developed by the BBC. Having a transcript for a digital media file allows search within the transcript and the ability to highlight the found words. The audio/video can be synced to a timecode in the transcript, allowing users to navigate directly to content of interest. The key words can be indexed for search engines to find related items. The goal is to create transcripts at scale, pushing a large number of files through the tool without first categorizing the sounds in the program.

Many of the videos in the collection start with bars and tone, or have music, or other sounds that Kaldi tries to transcribe into text and because of this, the transcript is not accurate. Kaldi tries to turn the sounds incorrectly into words. In addition, accurate transcripts are dependent on audio quality, speaker accents, background noise, etc. Given that the AAPB collection is from 100 different local tv and radio stations across the country, the variety of audio and audio quality varies widely. WGBH used 3 tools to allow the public to help fix and correct the English transcripts and add additional metadata.

The AAPB used Zooniverse, a crowd sourcing platform, to utilize the public to transcribe the credit information on the screen that a speech to text tool could not transcribe because there is no audio. This is valuable metadata which includes copyright notice, date of broadcast, validating program title, producer names, etc. The tool was called Roll the Credits and within 3 months, because of the large user base of Zooniverse, we had completed a data set of 917 program credits which included 5 verification passes. However, to set the tool up, we needed a screen grab of rolling credits that became 29, 206 frames, which took significant human time.

WGBH also built a game to fix the transcripts, called FIXIT. There are actually 3 games in one – identify errors, suggest fixes, and validate fixes. The pipeline to output a finished corrected transcript was too slow. It took too long, and too many players for 1 transcript to be completed, corrected, and validated. The most successful tool was FIX-IT+ which uses a transcript editor tool that New York Public Library

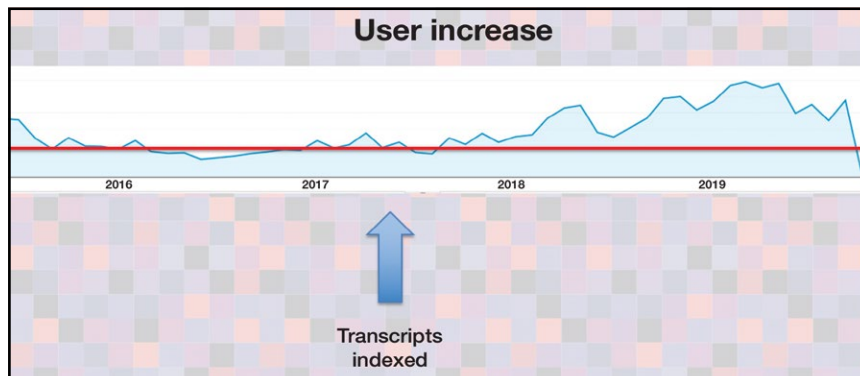
developed. The tool is a straight editor rather than a game and users are able to more effectively and efficiently correct transcripts. Two people have to agree on a fix for it to be validated.

A small percentage of transcripts have been corrected through this effort. The crowdsourcing campaign has engaged the general community and users intimately with the archive but has not corrected transcripts at scale. The most successful effort occurred when a vendor challenged AAPB contributors to push their communities to fix more transcripts. For every transcript corrected, the station would get a free tape digitized, up to 100, but they had to reach a 20 transcript threshold first. As a result, several stations engaged their volunteers with the challenge and there are now 530 completed transcripts. On the one hand, stations jumped on the challenge and people got involved, however, there are still thousands of transcripts to correct. Crowdsourcing corrections is not an efficient method to correct volumes of transcripts.

Once the transcripts have been verified, the JSON transcripts are stored in the AAPB's Amazon S3 account and indexed for keyword searching on the AAPB website. The transcripts will be made available alongside the media on the record page and can be played like captions within the video player. They will be able to be harvested via an API to be used as a data set for research. Researchers will be able to use the AAPB collection as a data set and start analyzing trends from programming over the last 60 years such as how language has been used in reporting, how it has changed over the years, and how it is different in different areas of the country. But to use this collection effectively as a data set, we need accurate good descriptive metadata.

The uncorrected transcripts were indexed in April 2017. There was an overall general increase in users coming to the site. (Figure 2) Of course, this is one of many factors that could have cause an increase in users, but it does show that search engines were driving traffic to the site even with inaccurate transcripts. The increased traffic is due to speech to text transcripts enhancing the metadata and thus discoverability and search. The question is whether users were actually finding what they were looking for.

Figure 2.



There is a lot of variety in the content and many of those indexed words are not accurate depiction of the content. For example, Kaldi does not recognize a foreign language. It tries to take the foreign language sounds and assign them English words that match the sounds. Understandable that there are tools for Spanish language transcription and translation, but many of the programs have mixed language and it is unknown which items or files have which languages. It would be helpful if a foreign language could be identified automatically, and then skipped over, instead of spending the effort to transcribe non-English into English. The same is true for music or other sounds. It is not very helpful and a waste of computational power and time to attempt to turn these sounds into English words. A helpful tool would characterize the language as not English and skip over it.

Our results from the Pop Up Archive version of Kaldi were for English transcripts, with an average of 81% accuracy, across the collection including all the mistakes around music and foreign languages. Other problems are punctuation errors and speakers with strong accents that were not well transcribed. Common errors are personal names. The project success is the creation of transcripts for the initial collection of 68,000 items which took 6-8 months to process. The quality of the transcripts varied. Specifically, the accuracy of programs from a single source single speaker, formal announcer was about 95% accurate (no accents, one speaker). However, transcripts

for a television program from Mississippi with a strong Southern U.S. accent was only 55% accurate. Data of named entities and locations from the Pop Up tools, was not useable. By sampling entities for 101 recordings, the conclusion was the number of tags, usefulness of those tags, and confidence in the accuracy of those tags would not significantly enhance the metadata records to warrant building the necessary new workflows and technologies. Many of the topic tags were not specific enough to be useful, such as “United States” “Entertainment and Culture” and “Human Interest.”

Near the end of the project, Pop Up Archive was bought by Apple, and shut down their Github account. The Kaldi tool being used was forked into the WGBH Github open account that anyone can access, but it is not the sophisticated trained tool being used for AAPB transcripts. WGBH has since dockerize the tools and is feeding programs through a Kaldi workflow on a regular basis. The accuracy these transcripts is about 56%, considerably less than the Pop Up trained tool of 81% average accuracy. WGBH continues to create transcripts using the lesser version of Kaldi for new content added to the AAPB. Even an inaccurate transcript gives some data about the program content. The crowd sourcing campaign to fix the output transcripts for accuracy is taking a long time. Another solution is needed to expedite the process, like perhaps trying to train the tool or using better machine learning tools.

Perhaps a commercial service such as Amazon speech to text is a bit more accurate, but 1) with this volume it gets expensive and 2) with this variety it is actually not that much more accurate overall. And this is not actually helping the open tool to learn and get better. Tools off the shelf, output the lowest common denominator and need programming to extract specific data beyond the common base. For example, Amazon web services facial recognition is probably not going to recognize many of the people in the AAPB collection. Better tools in the open source space, that are easy and affordable for archives to use is a much better long-term solution.

Now the AAPB is partnering with computational linguists at Brandeis University Computer Science department to develop workflows that utilize NLP and machine learning tools to extract key metadata for audio visual collections. Our goal is to begin communication

between the archives and computer science communities and develop open source tools that can be used by other cultural heritage organizations. Some key issues have been identified that would help with discoverability and access of the collection such as forced alignment of transcripts to time stamps, time stamping bars and tone at the beginning to improve the user experience, identify music and foreign languages to improve speech to text output, and identifying text on the screen and outputting it as data using Optical Character Recognition (OCR).

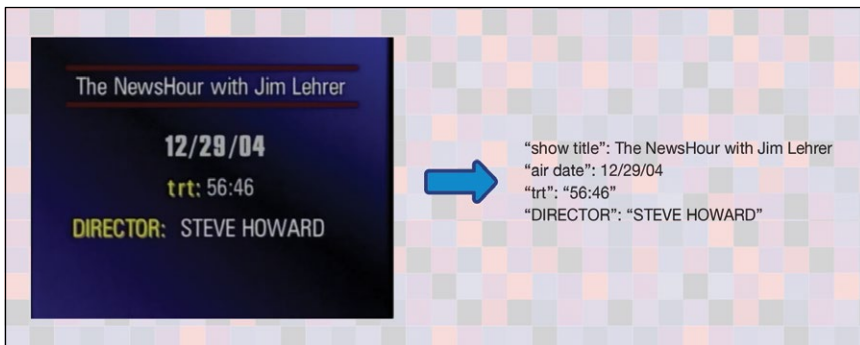
The data set focus is 30 years of NewsHour programs. The first step is to create key data that is verifiable such as program slates that can verify title, date aired, and producer; lower thirds identifying people on the screen; and credits at the end giving us production staff, participants, and copyright information. (Figure 3) That text can be used as metadata. If possible, the tools can also use the speaker lower third identification and verify against an announcer introducing the speaker, and then use facial recognition to find the same person again within the same video file. Eventually the tools can help with program type identification, better named entities, and better Kaldi output. For example, there is a single person at a desk therefore this might be a news broadcast.

Figure 3.



There is not one single tool that can create all this metadata, but rather a series of tools where the output of one tool becomes information for the next tool to refine the specific characterization or data extraction desired. Just as each speaker sound wave needed a specific algorithm or 'model' to teach the machine to identify a speaker, a separate tool is needed to characterize the item or attribute, decide which tool is needed to isolate the data, and pull out the informative text. Time based media complicates this process. What makes media different and harder is that the information you are trying to capture is moving across frames or tracks, and there is a huge matrix of pixels across each frame or image. The data on this frame is needed as text metadata. Text on a video frame, however is an image, not text. It can identify air date, director, running time and show title. But first the machine has to find the slate frame within the video file, get rid of everything that is not a slate, isolate the frame. Bounding boxes then find the text and character recognition for the text in the bounding box is performed, and it needs to recognize the words and what they are. For example, to pull the slate information off a video frame and make it useable metadata, the workflow is 1) a tool finds and isolates the slate video frame 2) the text location is identified on the video frame 3) the text information on the video frame is put through an OCR tool to create text output 4) the output is in a form that is understandable and correctly labelled. (Figure 4)

Figure 4.



The Brandeis team has created CLAMS. CLAMS, Computational Linguistics Applications for Multimedia Services, is a workflow tool that can plug in the appropriate tool for the appropriate attribute, once the file has been characterized. (Figure 5 and 6) This pipeline, or noodling tools together, will be used to create workflows of tools. One problem across all these tools is that the developers create their own interchange format and the tools from different sources don't talk to each other or exchange data. CLAMS uses a standard format (MMIF) to allow the tools to be used together. A common language that all the tools can understand.

Figure 5.

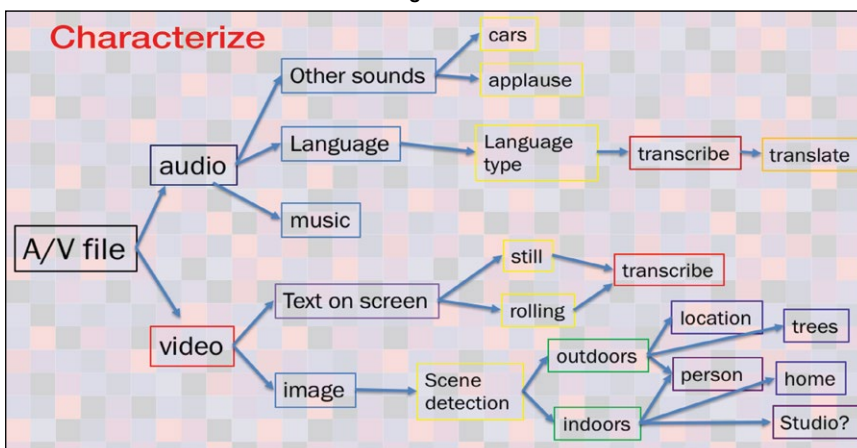
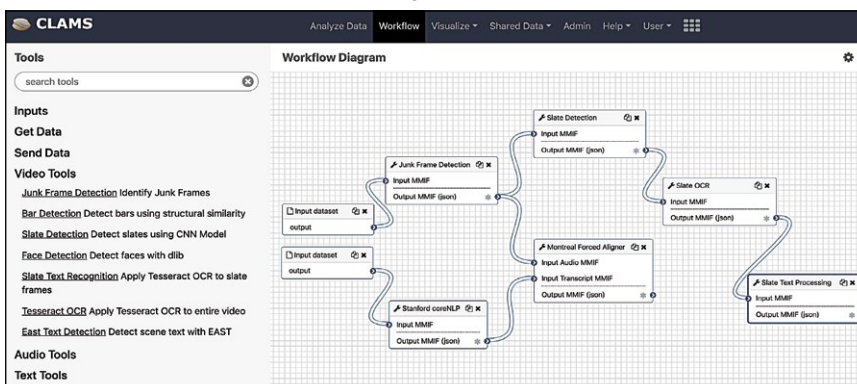


Figure 6.



The Brandeis team will write the algorithms, create the MMIF format, and have begun to build workflows and tools that have useful output for the AAPB. They are interested in developing an open source pipeline that will work with a variety of tools keyed to definitive specific tasks that A/V archivists might need - to enable an archivist to create a workflow by dragging and dropping tools from a tool shed to enable certain data outputs. This project is generously funded by the Andrew W. Mellon Foundation.

CONCLUSIONS

With the volume, variety and complexity of digital audio-visual collections, using machine learning and AI tools effectively, will take better knowledge of the tools than most archivists currently have to pipeline the tools into workflows. Libraries and archives have great data sets and computational scientists need large datasets to test tools, build tools and analyze trends. Libraries and archives need tools to improve the metadata thus improving the datasets. A collaboration between computational scientists and archives to improve computational tools, AI, and machine learning for better data in archives and libraries, would benefit both communities.

Computational linguists and AI experts have been working with and building tools for years that are just now being used by libraries and archives. AI experts can help build workflows and algorithms that will improve the tools archives and libraries need and make sure the output is useful to improve collections. Libraries and archives need to be able to talk to computer scientists using their definitions and terminology to understand the tools being built. The elasticity of the human brain to recognize variety and perform many tasks at once, is not yet there for machines. There is actually quite a lot of human effort that goes into machine learning or building the training for the machines. In order to use these tools for complicated collections they need to be adapted and trained. Archivists should work collaboratively with computer scientists toward better affordable, and open tools for archives and libraries.

BIBLIOGRAPHY

Cariani, Karen, & Kaufman, Casey Davis. "Crowd Sourcing Metadata for Time Based Media in the American Archive of Public Media." In *Participatory Archives: Theory and Practice*, edited by Edward Benoit III and Alexandra Eveleigh, 95-101. London: Facet Publishing, 2019.

_____. "Improving Access to Time-Based Media through Crowdsourcing and CL Tools: WGBH Educational Foundation and the American Archive of Public Broadcasting." In *Proceedings of the CLARIN Annual Conference 2018* (2018): 66-71.

Ide, Nancy, Pustejovsky, James, Suderman, Keith, & Verhagen, Marc. "Enhancing Access to Media Collections and Archives Using Computational Linguistic Tools." In *Proceedings of the Workshop on Corpora in the Digital Humanities* (2017): 19-28.

Rim, K., Lynch, K., & Pustejovsky, J. (2019, June). *Computational Linguistics Applications for Multimedia Services*. In *Proceedings of the 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (pp. 91-97).

URL REFERENCES:

WGBH Kaldi Github: <https://github.com/WGBH/kaldi-pop-up-archive>.

AAPB website: <http://americanarchive.org>.

FixIT: http://fixit.americanarchive.org/?utm_source=aapb_help-us_promo1&utm_medium=website&utm_campaign=help-us_from_website.

FixIT+: <http://fixitplus.americanarchive.org>.

Zooniverse: https://www.zooniverse.org/projects/sroosa/roll-the-credits/?utm_source=aapb_help-us_promo3&utm_medium=website&utm_campaign=help-us_from_website.

Reflexiones en torno a la inteligencia artificial: el caso de los archivos de RTVE

VIRGINIA BAZÁN-GIL

Responsable de proyectos Fondo Documental RTVE

INTRODUCCIÓN

El objetivo de mi comunicación no es otro que el de dar a conocer la travesía de los archivos de RTVE hacia la generación automática de metadatos. Es un viaje sin duda repleto de quimeras, espejismos y peligros, al final del cual, la tierra prometida parece esperarnos. Porque sin duda, una vez alcanzada la digitalización de los archivos, la inteligencia artificial es lo primero en lo que pensamos cuando alguien nos dice “pide un deseo”.

En primer lugar, presentaré el Fondo Documental de RTVE, a continuación, y de forma muy breve, expondré las tecnologías implicadas en la generación automática de metadatos, después señalaré su utilidad en la cadena de producción, difusión y archivo de los contenidos, identificaré a los proveedores tecnológicos y presentaré algunos casos de uso de la tecnología en las que hemos trabajado. Finalmente, les plantearé algunas conclusiones, poco definitivas, sobre el largo viaje del archivo de una televisión pública hacia la generación automática de metadatos.

EL FONDO DOCUMENTAL DE RADIO TELEVISIÓN ESPAÑOLA

El archivo de Radio Televisión Española es un archivo dividido en dos: el de radio y el de televisión. El archivo de Radio Nacional de España

(RNE) fue establecido en 1957 y conserva los programas emitidos por nuestras seis estaciones radiofónicas. El primer archivo sonoro de producción propia que conservamos data de 1947, se trata de los discursos oficiales de Franco y Eva Perón durante el viaje oficial de la primera dama argentina a nuestro país. El archivo de televisión ha cumplido ya 63 años. La grabación más antigua de producción propia se remonta al año 1959 y corresponde a un rodaje en las instalaciones de Miramar en Barcelona. Este archivo conserva, además, otras colecciones como los noticiarios y documentales de NO-DO producidos entre 1943 y 1981.

La principal misión del fondo documental de RTVE es la preservación del archivo, con el fin de garantizar su disponibilidad para la producción, reutilización y la comercialización de los contenidos. Además, como servicio público estamos comprometidos con la accesibilidad de nuestras colecciones, ya que preservamos contenidos de un indudable valor histórico. Pero, ¿cómo preservamos ese patrimonio? El archivo de RNE se digitalizó de forma temprana en 2002. La digitalización del archivo de TVE se realizó entre 2008 y 2014, con la colaboración de Telefónica. Este proyecto fue sin duda el mayor de su tipo en la digitalización de un archivo de televisión, implementado en España (Lopez-De-Quintana-Saenz 2014). En la actualidad esta tarea se hace con medios propios y queda únicamente por digitalizar el material de los centros territoriales y nuestro archivo de cine, una verdadera joya que requiere de procesos artesanales de recuperación de forma previa a la digitalización.

Como parte de nuestro plan de preservación digital hemos llevado a cabo la migración de nuestras cintas de datos de LTO4 y LTO5 a LTO7 y se ha multiplicado por cuatro la capacidad de nuestro almacenamiento digital. Asimismo, hemos logrado la unificación de las librerías de programas e informativos, de forma que podremos gestionar de manera más eficiente nuestros recursos.

La gestión de las colecciones (texto, audio, video y fotografías) se realiza a través del gestor documental ARCA, un desarrollo propio implantado en 2011, que permitió la unificación de las distintas bases documentales y la adopción de un modelo de metadatos común, tanto para el archivo de radio como para el de televisión y las distintas

unidades implicadas. ARCA está integrado con otras aplicaciones de negocios, como emisiones, comercial, subtítulo, unidades de producción y RNE, así como con el CMS de medios interactivos, lo que permite a nuestros compañeros de la web publicar contenidos del archivo.

Como servicio público estamos obligados a abrir nuestras colecciones a los usuarios pero al mismo tiempo tenemos que cubrir nuestras necesidades de producción, lo que implica un equilibrio inestable de recursos, necesidades y demandas. Estamos realizando un gran esfuerzo para publicar el archivo de RTVE en la web (RTVE 2019) y responder a las peticiones de los investigadores. En este sentido, estamos explorando nuevas vías, la Biblioteca Nacional Española (BNE), con quien además estamos trabajando en una política de preservación de nuestros contenidos en la web y de quienes esperamos obtener, en el futuro, la calificación de repositorio seguro. Tenemos convenios de colaboración con distintas instituciones patrimoniales, entre otras, con la Filmoteca Española. Somos miembros activos de la Federación Internacional de Archivos de Televisión (FIAT/IFTA 2019) y colaboramos con otras iniciativas europeas como EUScreen (EUScreen 2019).

TECNOLOGÍAS DEL HABLA, PROCESAMIENTO DEL LENGUAJE NATURAL Y VISIÓN ARTIFICIAL

Los procesos de generación automática de metadatos en los archivos se fundamentan en tres tecnologías complementarias, visión artificial, tecnologías del habla y procesamiento del lenguaje natural (Lleida 2018):

- La visión artificial se ocupa del reconocimiento de imágenes, la agrupación y segmentación de escenas y el seguimiento de objetos y personas.
- Las tecnologías del habla, permiten tanto la conversión voz a texto como un tratamiento de los hablantes y sus emociones, facilita su identificación y la segmentación del discurso.

- El procesamiento del lenguaje natural permite el reconocimiento de entidades y temas tratados en un discurso y facilita la creación de resúmenes y descripciones.

En otras palabras, la visión artificial proporciona la descripción de la capa de imagen, bien en forma de etiquetas, bien en forma de descripciones, en lenguaje natural (Bazán-Gil y Guerrero Gómez-Olmedo 2018). Las tecnologías del habla permiten convertir la capa de audio en un texto y puede aplicarse tanto en la fase de edición para la transcripción automática de entrevistas, por ejemplo, como en la emisión y difusión de contenidos través de la web en forma de subtitulado; y finalmente, el procesamiento del lenguaje natural contribuye a enriquecer con metadatos el material en la publicación web y en el archivo, al generar puntos de acceso en forma de entidades para la recuperación o al facilitar la clasificación automática de los contenidos con el uso de ontologías. (Bazán-Gil et al. 2019)

PERO, ¿QUIÉN NOS PROPORCIONA LA TECNOLOGÍA?

Podemos identificar tres grupos de proveedores de inteligencia artificial:

- Grandes compañías internacionales Microsoft Azure (Microsoft 2019), IBM Watson (IBM 2019) y Google AI (Google 2019), mismas que gestionan una gran cantidad de datos procedentes de *social media*. Tecnología propietaria en constante proceso de mejora.
- Compañías especializadas en el sector *media*, que integran soluciones de terceros o tienen sus propios desarrollos basados en tecnologías *open source*. Pueden ser MAM Vendors como: AVID (AVID 2019), Dalet (DALET 2019), VSN (VSN 2019), o Tedial (TEDIAL 2019), o bien empresas dedicadas como Etiquedia (ETIQMEDIA 2019) y Limecraft (Limecraft 2019), que se ocupan únicamente del procesamiento de los contenidos en flujos secundarios.

- Grupos de investigación nacionales e internacionales. Esta es una fórmula de colaboración que en el caso de RTVE ha dado buenos resultados, ya que si algo demanda la IA es el procesamiento de grandes cantidades de datos, algo en lo que somos ricos los archivos. Identificar grupos de investigación en tecnologías del habla, procesamiento del lenguaje natural e inteligencia artificial y crear cátedras o solicitar proyectos de investigación conjuntos, puede ser una fórmula de éxito.

PERO, ¿ES LA GENERACIÓN DE METADATOS PATRIMONIO EXCLUSIVO DEL ARCHIVO?

La generación de metadatos no ha sido nunca patrimonio exclusivo del archivo, tradicionalmente otras áreas como las de producción, emisión o comercialización han generado metadatos para los mismos contenidos audiovisuales en función de sus propias necesidades. Y en el caso de RTVE, de forma reciente, estamos experimentando la gran demanda de datos procedente del área de medios interactivos. Los sitios web de las radiodifusoras demandan metadatos para identificar, clasificar, enriquecer, hacer más accesibles y buscables los contenidos y esta demanda se produce a un ritmo mayor del que los propios archivos podemos asumir. Más datos y más rápido. Y a ninguno de nosotros como profesionales se nos escapa que a menudo estas áreas cuentan con la capacidad de acción y los recursos necesarios para emprender nuevos proyectos tecnológicos que aun siendo necesarios son impensables para el archivo. Y esto, que puede considerarse como una amenaza, no es sino una oportunidad para encontrar en estas áreas aliados estratégicos en la adopción de las nuevas tecnologías. Porque como repetía el slogan “*Solo no puedes, con amigos sí*”, del mítico programa de TVE *La bola de cristal*, de los años 80.

Aún más, comprender el ciclo de vida de los datos en el proceso de producción, difusión y archivado, no solo nos permite identificar estas alianzas, sino que puede contribuir a la eficiencia y sostenibilidad de nuestros datos como organización (Gomez Zotano y Bazán-Gil 2018) Llamémoslo si quieren la ecología del dato ¿por qué producir constantemente nuevos metadatos cuando podemos usar los que otros ya han generado?

Esta es una de las lecciones más valiosas que hemos aprendido, prueba de concepto tras prueba de concepto, en los últimos tres años.

CASOS DE USO DE LA TECNOLOGÍA EN RTVE

Una vez identificadas las tecnologías y su utilidad, los *partners* tecnológicos y los aliados estratégicos, repasaremos algunas de las experiencias en el uso de la AI en Radio Televisión Española. En concreto, las más recientes que han tenido como protagonistas las tecnologías del habla: Iberspeech 2018 (Lleida et al. 2019) y el proyecto “Crónicas”.

Dentro de las tecnologías del habla, las más relevantes para nosotros son: transcripción de voz a texto (Speech to Text) y segmentación y agrupación por hablantes, también conocida como diarización.

Los sistemas de voz a texto nos permiten conocer qué se ha dicho y son fundamentales para lograr:

- Garantizar la accesibilidad de los contenidos mediante la generación de subtítulos en tiempo real y la posibilidad de traducción automática a otros idiomas.
- Mejorar los tiempos de producción a través de la transcripción automática de entrevistas, ruedas de prensa, etc.
- Automatizar parcialmente los procesos de análisis, ya que en las transcripciones se basa la clasificación y etiquetado automático contenidos.

La diarización nos ayudará a identificar quién habla, cuándo habla y qué aspecto tiene si aplicamos, además, técnicas de reconocimiento facial:

- Indexación de hablantes y discursos.
- Estructuración de contenidos.
- Reconocimiento de hablantes como apoyo al Speech to Text.

IBERSPEECH

En este contexto, se impulsó la investigación en tecnologías del habla en español, en mayo de 2018, con la Catedra RTVE Universidad de

Zaragoza (Cátedra RTVE Universidad de Zaragoza 2017), instituida un año antes con el objetivo de investigar en la creación automática de metadatos. De esta manera, RTVE Iberspeech Challenge 2018 (RTVE 2018), lanzó un reto tecnológico que puso a disposición de la comunidad científica más de 500 horas de contenidos emitidos, junto con sus correspondientes subtítulos.

El conjunto de datos liberado incluye cerca de una veintena de programas de distintos géneros y temática, producidos y emitidos por RTVE, entre 2015 y 2018. Estos programas presentan distintas dificultades desde el punto de vista de las tecnologías del habla como son: la diversidad de acentos del español, la superposición de diálogos, el habla espontánea, la variabilidad acústica, el ruido de fondo o el vocabulario específico, factores todos ellos que inciden en el rendimiento de los sistemas de transcripción.

Sobre este conjunto de programas, los 22 grupos participantes debían detectar automáticamente, etiquetar y transcribir segmentos del habla. Para ello se definieron tres tareas distintas: voz a texto, diarización y diarización multimodal.

RETO 1: TRANSCRIPCIÓN VOZ A TEXTO

El objetivo de este reto era transcribir de forma automática 39 horas de emisión. La métrica aplicada, para medir el rendimiento de los sistemas en la WER o tasa de error por palabras, que tiene en cuenta el número de inserciones, sustituciones y borrados en relación con el número total de palabras (Lleida et al. 2019). Los resultados de este reto ponen de manifiesto que la calidad de la descripción depende del tipo de programa:

- El sistema mejor evaluado obtiene una tasa de error del 16.45% y el peor evaluado un 35.80%
- Programas con entrevistas o conversaciones espontáneas y con audio en exteriores como “Dicho y Hecho”, “La Mañana” o “Arranca en verde” son los programas que suponen el mayor reto con tasas de error muy por encima del 20% para la mayoría de los sistemas.

- Muchos programas mantiene unas tasas de error por debajo del 20% e incluso para un programa como “Latinoamérica en 24H”, con distintos acentos del español, la tasa de error se sitúa por debajo del 10%.

RETO 2: DIARIZACIÓN

El objetivo de este reto era segmentar y agrupar por hablantes 22 horas de emisión. La tasa de error que se aplicó para medir el rendimiento de los sistemas es el DER o tasa de error en la diarización que tiene en cuenta el número de segmentos que no se atribuyen correctamente a un hablante (Lleida et al. 2019). Los resultados de este reto demostraron:

- La variabilidad en las prestaciones en función del tipo de contenido (tasas de error, entre el 17.27% y el 39.09%) .
- La asignación incorrecta de hablantes como error más común.

RETO 3: DIARIZACIÓN MULTIMODAL

El objetivo de este reto era segmentar y agrupar por hablantes y cara cuatro horas de emisión. La tasa de error aplicada es la correspondiente en la diarización por hablantes y caras. Los resultados demuestran la variabilidad en función del tipo de programa y un mejor rendimiento de la diarización de hablantes frente a la de caras (Lleida et al. 2019).

Estos resultados deben abordarse además desde la perspectiva del usuario, es decir desde la utilidad de las tecnologías para su integración en los flujos de trabajo de una organización como RTVE.

En los últimos cinco años los reconocedores de voz han mejorado de forma espectacular. Aun así, todavía existe una gran variabilidad en la tasa de error, que limita su utilidad en todos los escenarios posibles: desde el material original hasta los programas emitidos.

EL PROYECTO “CRÓNICAS”

Este proyecto, también en el contexto de las tecnologías del habla pretendía dar respuesta a un largo anhelo de los periodistas, la trans-

cripción automática de las entrevistas que realizan para sus reportajes. El objetivo fundamental de esta prueba piloto era determinar la utilidad de los reconocedores automáticos en la producción de nuevos contenidos. Para ello, quisimos probar distintas herramientas, su posible integración con nuestra herramienta de producción y determinar que tasa de error era asumible para nuestros usuarios. Este proyecto se realizó en dos fases. En la primera fase se definieron los requisitos del sistema, en colaboración con el equipo del programa, y se identificaron como fundamentales los siguientes:

- Sistema intuitivo y fácil de usar.
- Múltiples usuarios conectados al mismo tiempo
- Debe conservar los códigos de tiempo originales de la grabación
- La segmentación de la entrevista debe realizarse por hablantes y diferenciar entre preguntas y respuestas.
- Las transcripciones deben incluir capitalización y signos de puntuación.
- La edición y corrección de la transcripción generada debe ser sencilla.
- Tasas de error no superiores al 15%.

En esta primera fase, también, se definió el alcance de la muestra para realizar las pruebas. Se seleccionaron 25 rodajes originales con base en factores como la variedad de modelos del lenguaje, la calidad de audio, la presencia de habla espontánea, la existencia de diálogos superpuestos y de diferentes acentos del español.

Durante la segunda fase de este proyecto se llevaron a cabo pruebas con dos soluciones tecnológicas: MAM VSN y Limecraft.

El MAM de VSN integra varias soluciones de IA de Microsoft Azure y Google Cloud, para esta prueba se usaron ambas, con el objetivo de establecer una comparación entre el rendimiento de cada una. Los ficheros se subieron al MAM de VSN, mediante un sistema de *drag and drop* que permiten arrastrar directamente los ficheros desde el equipo del redactor a la solución de VSN. Los tiempos de subida y procesamiento de los ficheros, uno de los puntos críticos del proyecto, fueron satisfactorios -menos de la mitad de duración del fichero,

en el procesamiento, no así las tasas de error devueltas por ambos sistemas. Tanto Google como Microsoft, en todos los rodajes, devolvieron tasas de error por encima del 50%, además, destaca el hecho de que los ficheros con audio en alta calidad procesados por Google presentaron tasas de error más altas que los ficheros en baja calidad.

Tanto Azure como Google generan segmentos en torno al minuto de duración, es decir, no segmentaron por hablantes o al menos por fragmentos del discurso con significado semántico. Google no devuelve signos de puntuación ni capitalización y Azure sí lo hace, pero tiende a finalizar casi todas las frases con un signo de interrogación.

Una segunda iteración se llevó a cabo con Limecraft (Limecraft 2019), una solución cloud basada en Speechmatics (Speechmatics 2019), un reconocedor del habla con muy buenos resultados en distintos retos internacionales y que integran distintos canales de televisión internacionales. Es por tanto una herramienta específica de transcripción de voz a texto. Limecraft presentaba una interfaz sencilla de uso, proporcionaba segmentación por hablantes, capitalización y signos de puntuación y la posibilidad de editar los contenidos directamente en la interfaz de consulta. Las tasas de error para los distintos rodajes se situaron en torno del 20% al 30%, considerablemente mejores que las obtenidas por Google y Azure y, a diferencia de estos sistemas, Limecraft devuelve segmentación por hablantes, aunque como hemos visto en los resultados del reto de diarización en Iberspeech 2018, este no es un problema totalmente resuelto.

Una vez finalizadas las pruebas y, a pesar de que la percepción de la redacción respecto a la utilidad de la tecnología evolucionó positivamente a lo largo del proyecto, en ningún momento llegó a considerarse su integración de esta tecnología en los flujos de trabajo habituales de la redacción del programa. En opinión de los periodistas que participaron, el esfuerzo en la corrección de las transcripciones distraía su atención de la rememoración de la entrevista y dificultaba el proceso de edición. Nuestros colegas tampoco valoraron positivamente la integración de estas soluciones con herramientas de edición.

Desde el punto de vista del archivo, la integración de estas soluciones en la producción de contenidos es muy relevante, ya que los rodajes originales y, de forma más concreta, las entrevistas, no suelen

llegar al archivo con los metadatos suficientes para su adecuado tratamiento. La transcripción exacta de lo que se ha dicho, cobra especial relevancia en el contexto de la verificación de noticias y sigue siendo relevante para la recuperación en tanto que multiplica los puntos de acceso a los contenidos en la búsqueda por texto libre.

CONCLUSIONES

En los últimos 5 años los sistemas de reconocimiento del habla han mejorado considerablemente sus tasas de error. De acuerdo con las pruebas que hemos realizado en los dos últimos años, las tasas de error pueden situarse entre el 5% y el 50% y esto, demuestra que no hay soluciones universales. Existe una correlación entre tipo de programa y resultado de la transcripción. En este sentido, y desde el punto de vista del archivo, es necesario comprender el funcionamiento de la tecnología y bajo qué premisas tiene un mejor rendimiento. Este conocimiento debe ponerse además en relación con las necesidades reales de trabajo y con las expectativas de las distintas áreas en cuanto a la calidad de los resultados. La inteligencia artificial es enemiga de la calidad humana. No debemos olvidar tampoco que el grado de tolerancia respecto a los errores varía en función del escenario para el que se generan los metadatos: tasas de error por encima del 15% pueden ser tolerables para el archivo cuando se carece de otros datos para la recuperación, pero no son admisibles para la emisión de los contenidos.

Para los archivos audiovisuales la incorporación de IA a sus procesos diarios es más que una prioridad, pero la pregunta es ¿por dónde empezar? Para iniciar este viaje es necesario concienciarse sobre la importancia, no solo de generar metadatos en función de las necesidades de las distintas áreas, sino de establecer los mecanismos adecuados para garantizar la reutilización de estos datos en la cadena de producción - emisión - archivado. Si comprendemos el ciclo de vida de los datos en esta cadena seremos capaces de identificar a nuestros aliados estratégicos para este viaje.

Tras las experiencias y el conocimiento adquirido hasta el momento, podemos identificar algunas acciones como fundamentales antes

de embarcarse en el viaje hacia la implantación de la inteligencia artificial en un archivo de televisión:

- Analizar las necesidades como archivo y comprender las de otras áreas. ¿Dónde es prioritaria la generación de metadatos? Es necesario poner el foco en función de nuestros recursos.
- Identificar proveedores tecnológicos y grupos de investigación relevantes en las tecnologías que se quieran poner a prueba.
- Realizar una prospección de la tecnología, de sus capacidades y especialmente de sus limitaciones.
- Crear sinergias con áreas que puedan impulsar y hacer realidad la adopción de la tecnología: “Solo no puedes, con amigos sí”.
- Examinar el modelo de metadatos actual ¿es suficiente para asumir la avalancha de nuevos datos procedentes de la inteligencia artificial?
- Ajustar las expectativas de los usuarios al rendimiento real de la tecnología.
- Evitar el uso indiscriminado de la tecnología como la única forma posible de malgastar recursos económicos y humanos, sin la garantía de obtener resultados adecuados.

Por último, la adopción de soluciones de inteligencia artificial en distintas áreas de una cadena de radio y televisión es una magnífica oportunidad para hacer visible el archivo. El poder de los datos es propio de nosotros los y las archiveras. No duden en convertirse en el oráculo de otras áreas dentro de la organización.

REFERENCIAS

- AVID. (2019). “Avid - Tecnología y Herramientas Para Capacitar a Los Creadores de Contenido.” <https://www.avid.com/#Video>.

Bazán-Gil, Virginia, and Ricardo Guerrero Gómez-Olmedo. (2018). "Descripción Automática de Archivos Audiovisuales: NeuralTalk, Un Modelo de Video2text Aplicado Al Archivo de RTVE Cita Recomendada." *BiD: Textos Universitaris de Biblioteconomia i Documentació*, no. 41. <https://doi.org/10.1344/BiD2018.41.7>.

Bazán-Gil, Virginia, Eduardo Lleida, Carmen Pérez, Manuel Gómez, and Alberto De Prada. (2019). "Tecnologías Del Habla: Nuevas Oportunidades Para Los Archivos de Televisión." In *IV Congreso ISKO España-Portugal*. Barcelona: Universitat de Barcelona.

Cátedra RTVE Universidad de Zaragoza. (2017). "Cátedra RTVE de La Universidad de Zaragoza." <http://catedrartve.unizar.es/>.

DALET. (2019). "Orchestrated Media Solutions | Dalet." <https://www.dalet.com/>.

ETIQMEDIA. (2019). "Catalogación Supervisada de Contenido Audiovisual. ETIQMEDIA. Inicio. ETIQMEDIA." <http://www.etiqmedia.com/>.

EUscreen. (2019). "EUscreen." <http://www.euscreen.eu/>.

FIAT/IFTA. (2019). "FIAT/IFTA." <http://fiatifta.org/>.

Gomez Zotano, Manuel, and Virginia Bazán-Gil. (2018). "Interoperabilidad y Recomendación de Contenidos: Explotando El Archivo En La Web." *II Jornadas Fondo Documental RTVE: Los Archivos de Televisión Después de La Digitalización*. <http://www.rtve.es/rtve/20180202/jornadas-archivos-televisión/1672420.shtml>.

Google. (2019). "Productos de Inteligencia Artificial y Aprendizaje Automático." <https://cloud.google.com/products/ai/?hl=es>.

- IBM. (2019). "IBM Watson Products and Solutions | IBM." <https://www.ibm.com/watson/products-services>.
- Limecraft. (2019). "Limecraft - Your Video Production Workflow Set up in a Few Easy Steps : Limecraft". <https://www.limecraft.com/>.
- Lleida, Eduardo. (2018). "Tecnologías Para AI Análisis y Metadato de Contenidos Audiovisuales." Los Archivos de Televisión Después de La Digitalización: Gestión y Acceso Abierto. http://www.rtve.es/contenidos/documentos/instituto/4_Jornada_Archivos_tv.pdf.
- Lleida, Eduardo, Alfonso Ortega, Antonio Miguel, Virginia Bazán, Carmen Pérez, Manuel Gómez, and Alberto De Prada. (2019). "Albayzin 2018 Evaluation : The IberSpeech-RTVE Challenge on Speech Technologies for Spanish Broadcast Media." *Applied Sciences*, 1-23. <https://www.mdpi.com/journal/applsci>.
- Lopez-De-Quintana-Saenz, Eugenio. (2014). "Rasgos y Trayectorias de La Documentación Audiovisual: Logros, Retos y Quimeras." *Profesional de La Informacion*. <https://doi.org/10.3145/epi.2014.ene.01>.
- Microsoft. 2019. "Video Indexer: Información de Vídeos de IA | Microsoft Azure." 2019. <https://azure.microsoft.com/es-es/services/media-services/video-indexer/>.
- RTVE. (2018). "La Cátedra RTVE En La Universidad de Zaragoza Presenta Su Primer Reto Tecnológico a La Comunidad Científica - RTVE.Es." <http://www.rtve.es/rtve/20180521/catedra-rtve-universidad-zaragoza-presenta-su-primer-reto-tecnologico-comunidad-cientifica/1737360.shtml>.
- _____. (2019). "Archivo RTVE." Archivo RTVE. <http://www.rtve.es/television/archivo/>.

Speechmatics. (2019). "Speech to Text Recognition Powered by Machine Learning | Speechmatics." <https://www.speechmatics.com/>.

TEDIAL. (2019). "Tedral Media IT | Media Asset Management (MAM) Solutions." <https://www.tedral.com/>.

VSN. (2019). "VSN | Innovation and Media Solutions." <https://www.vsn-tv.com/es/>.

Catalogación automática: más cerca que lejos

ALEJANDRO ÁLVARO RAMÍREZ ACOSTA
JUAN CARLOS MURILLO PABÓN
*Laboratorio de Radio e Inteligencia Artificial RTVC
Sistema de Medios Públicos*

INTRODUCCIÓN

Automóviles que se conducen solos, computadoras que juegan mejor que nosotros, máquinas que diagnostican enfermedades, y motores de búsqueda de información que nos dicen lo que necesitamos, son algunos de los alcances de la Inteligencia Artificial (IA) hoy en día. Pero, ¿podríamos catalogar los archivos sonoros de manera automática?

La catalogación a través de la inteligencia artificial está más cerca que nunca y es el tema de esta ponencia.

El capítulo está dividido en tres partes, en esta se habla de las experiencias llevadas a cabo por el Archivo Sonoro de Señal Memoria en el ámbito de la catalogación, además, presenta los retos en este campo, a partir de las necesidades actuales; se comparten las experiencias del proyecto Mex-Culture en el campo de la inteligencia artificial, y se presenta la estrategia del Laboratorio de Radio e Inteligencia Artificial de RTVC para brindar soluciones innovadoras.

EXPERIENCIAS LLEVADAS A CABO POR EL ARCHIVO SONORO DE SEÑAL MEMORIA EN EL ÁMBITO DE LA CATALOGACIÓN

Durante el periodo de 2009 a 2015, se realizaron los procesos de catalogación sonora nivel 2 y la descripción del contenido sonoro con el *software de catalogación Mandarín*, el cual esta estandarizado bajo el formato MARC21, como se muestra en la *Figura 1*; la catalogación audiovisual se realizó en el aplicativo Winisis, donde se buscó identificar la cantidad de soportes y el título de marcación, *ver Figura 2*.

Figura 1. Primer sistema de catalogación sonora Mandarín (M3 catalogación).

Autoridades - M3 Catalogación

Archivo Ver Ayuda

LDR nz

003

005 200204171134220

008 020417nl acannaabn ln aab d

040 a c d

100 1 a c q d

400 1 a c q d

500 1 a c

Ejemplo

Guardar Cancelar Ayuda

Después de guardar... Cerrar Duplicar

Marcado para borrar

Anterior Siguiente Primero Ultimo

Figura 2. Primer sistema de catalogación audiovisual Winisis.

Adición global de campo

Rango MFN (de/a) 1/32000

Etiqueta de campo

☐ Add if not present

Contenido del campo

Cancelar Ok

De 2015 a 2018 se creó un desarrollo web para recibir la información de catalogación que estaba siendo producida en RTVC. Adicionalmente, se realizaron gestiones para la optimización del *software* **Mandarin** y se trató de mejorar la utilización del mismo en la recuperación de la información sonora. Los procesos de catalogación del material audiovisual se desarrollaron a través de una plataforma web que soportaba una base de datos, *ver Figura 3*.

Figura 3. Sistema actualizado de catalogación Mandarin en la web



El equipo de Señal Memoria durante el año 2016 diseñó dos plantillas de acuerdo con los estándares internacionales de catalogación audiovisual y sonora para una solución integral a los dos procesos de catalogación. Si bien la catalogación se estandarizó, **las plataformas no cumplían con los requerimientos de un catálogo**, el cual pueda ser accesible al público en general, así como tener una organización de las fichas catalográficas. Estas plataformas presentaron dificultades en la alimentación de las bases de datos de autoridades y vocabularios controlados, como también en la migración de las bases de datos a la página de Señal Memoria.

Para cumplir con los objetivos planteados para la vigencia 2016 del proyecto Señal Memoria, RTVC suscribió el Convenio interadministrativo no. 494 de 2016, con el Fondo de Tecnologías de la Información y las Comunicaciones, y cuyo objeto dice: “Aunar esfuerzos técnicos,

administrativos y financieros tendientes a recuperar la colección de archivos audiovisuales resultado de las producciones de televisión pública, que conforman el patrimonio digital de Radio Televisión Nacional de Colombia - RTVC, que se encuentra en mayor riesgo de pérdida por deterioro, haciendo uso de las tecnologías desarrolladas para garantizar su adecuada conservación”, y como uno de sus fines, continuar con los procesos de conservación, digitalización y catalogación del patrimonio audiovisual de RTVC.

La estrategia Señal Memoria requirió un soporte importante en la gestión de los contenidos a digitalizar, como en la solución de las dificultades presentadas en las plataformas, y otras actividades necesarias para el acceso a los contenidos audiovisuales y sonoros. Estos procesos se desarrollaron bajo orientación y con apoyo del área Técnica de RTVC; para lo cual se contrataron los servicios de un profesional que realizó el apoyo a la estructuración técnica y operativa de los catálogos digitales y aplicaciones del proyecto Señal Memoria. Este apoyo dio como resultado la adquisición de un *software* de catalogación que permitiera integrar las distintas bases de datos con los registros detallados de los documentos que componen el archivo audiovisual y sonoro de RTVC.

Para el año 2016, existían alternativas de programas de *software* con posibles soluciones a las limitaciones planteadas en los párrafos anteriores; por lo que, la unificación de la catalogación orientada al almacenamiento de la información relevante del archivo audiovisual y sonoro de Señal Memoria en la Nube, y la interacción de los metadatos de los activos digitales con el Sistema Gestor de Medios, dio como resultado, adquirir el programa de código abierto **KOHA**, el cual es un sistema de administración de bibliotecas escalable y completo. Este sistema constituye el archivo actual, como una biblioteca digital de acceso abierto al público, y permite direccionar distintos tipos de búsqueda de información. Desde el año 2016 a la fecha, los procesos de catalogación se realizan bajo KOHA y el estándar internacional MARC21, la *Figura 4* muestra al sistema KOHA y su interacción.

Figura 4. Nuevo sistema de catalogación KOHA, versión catalogador.

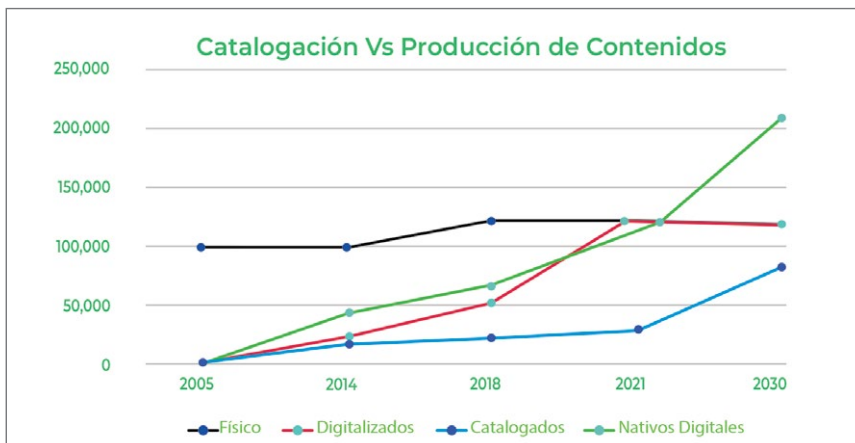


Figura 5. Nuevo sistema de catalogación KOHA, versión usuario (web).



Antes de iniciar el año 2018, se presenta un problema en la catalogación, el proceso creado en años anteriores enfocado en la gestión de la información y la catalogación se aleja del proceso inicial de recuperación (conservación, digitalización), y llega a quedar muy por debajo de las metas iniciales planteadas como línea de trabajo.

Figura 6. Catalogación vs. procesos técnicos.



En la *Figura 6* se puede ver la diferencia que toman los procesos técnicos, sin contar con la catalogación de los archivos nativos digitales.

RETOS EN LA CATALOGACIÓN A PARTIR DE LAS NECESIDADES ACTUALES

La catalogación en la actualidad cumple con estándares internacionales, pero la información crece día con día en RTVC, al ser la casa de los medios públicos de Colombia (Radio y Televisión) los contenidos nuevos superan cualquier proyección cuando se habla de recuperar la información de manera práctica y eficaz, esto lleva a replantear el trabajo y a buscar solución a los siguientes desafíos:

- Realizar la recuperación de la catalogación de los archivos de la radio y televisión en tiempo real.
- Alcanzar las metas del proceso de recuperación del archivo de Señal Memoria (analógico).

Otros retos son:

- Unificar criterios.
- Personal capacitado para realizar las tareas de descripción.
- Interoperabilidad con otras herramientas.

- Consulta y acceso a la información.
- Estandarizar el lenguaje.
- Agilizar el proceso.
- Procesos en tiempo real.

Adicional a la recuperación de la catalogación, la radio busca generar descriptores específicos a temas de interés nacional para medir procesos.

Es aquí cuando nace el Laboratorio de Radio e Inteligencia Artificial de **RTVC**, el cual busca resolver los problemas generales y específicos planteados anteriormente.

EXPERIENCIAS DEL PROYECTO MEX-CULTURE EN EL CAMPO DE LA IA EN LA CONSTRUCCIÓN DEL LABORATORIO DE RADIO E INTELIGENCIA ARTIFICIAL DE RTVC

En la actualidad se tienen diversas aplicaciones tecnológicas que están basadas en el análisis de la señal sonora, estas se centran en dos áreas, aquellas destinadas al procesamiento de señales sonoras de ambiente (De la Cruz 2012) o música (MaguaRed 2017) y aquellas destinadas al procesamiento de señales sonoras de voz (Morejón 2011).

Durante la construcción del proyecto Mex-Culture se desarrollaron algunas aplicaciones, con base en una segmentación robusta de la señal sonora, la cual debe brindar información de cada componente que conforma a esta señal, como por ejemplo, el componente de la voz, de la música y del silencio. La información recabada dependerá de la aplicación a la cual nos estemos enfocando y el propósito que esta tenga. Los métodos de segmentación robusta están basados en procesamiento de la señal sonora y métodos de inteligencia artificial, basados en redes neuronales artificiales de aprendizaje profundo.

Los métodos automáticos de segmentación de la señal sonora toman relevancia a nivel social cuando forman parte de la base de proyectos que generan aplicaciones de ayuda e impacto que se ven reflejados en distintas áreas sociales. En el área de procesamiento y clasificación de la música se puede ejemplificar con aquellos proyectos que buscan la preservación de arte sonoro a través de los años (Arratia *et al.* 2013), ya que esta cambia conforme avanza el tiempo y hay varios sonidos o

música que ya no se encuentran dentro de cierta cultura (Espejo Mojica 2017); en el área de procesamiento, clasificación y reconocimiento de la voz se puede ejemplificar con aquellos con mayor importancia que son los dedicados al tratamiento de enfermedades, a la identificación de características de personas (Romo Melo 2005), entre otros.

Dentro de las aplicaciones de los métodos de segmentación de la señal sonora se tiene como tema relevante a nivel internacional proyectos para preservar, almacenar y difundir los sonidos representativos del entorno natural y cultural (música, sonidos de la naturaleza, sonidos artificiales), para así identificar los componentes fundamentales y establecer una identidad cultural (UNESCO 2018); para realizar este tipo de proyectos, los principales retos que se tienen es tener un control de la información sonora, es decir, de los archivos digitales de contenido sonoro, los cuales forman parte de grandes bases de datos.

Para llevar a cabo lo anterior de manera automatizada, los métodos de segmentación establecen los fundamentos para el procesamiento de la información de la señal sonora y la relación con la gestión de las bases de datos. Esta relación de la gestión se basa primero en utilizar los criterios y protocolos que se fueron proponiendo a través de los años, en las recomendaciones y estándares que se deben de tener en cuenta para crear una base de datos para la preservación de archivos sonoros, después la base de esta recae en el medio y los metadatos [8], en donde el medio es el audio digital y los metadatos son la información descriptiva y fundamental de cada uno de los archivos sonoros, esta información debe provenir de una catalogación correcta, de una buena digitalización y de la administración del documento sonoro (Rodríguez Reséndiz 2014), es por ello que hacer una catalogación de manera automática es fundamental para estos proyectos.

ESTRATEGIA DEL LABORATORIO DE RADIO E INTELIGENCIA ARTIFICIAL DE RTVC PARA BRINDAR SOLUCIONES INNOVADORAS

Hoy en el mundo la gestión de contenidos digitales multimedia es un reto para todas las empresas de comunicación, en especial para las entidades de carácter público dado que éstos se construyen con recursos públicos y pasan a ser bienes intangibles de la sociedad que

deben ser de libre acceso para los ciudadanos lo que les proporciona un enlace con su identidad cultural.

También estos contenidos digitales multimedia son un recurso fundamental para que la entidad nutra sus plataformas digitales y preste un servicio a la ciudadanía de acceso y seguimiento a la información.

En la actualidad, el sistema de medios públicos de RTVC genera desde la subgerencia de radio más de 72 horas diarias de contenidos, además de información multimedia, que requieren ser analizados y procesados de inmediato para que se optimice el acceso desde las plataformas de memoria de RTVC, situación que en la actualidad no se puede satisfacer dada la cantidad de información mencionada, más la cantidad de documentos sonoros recuperados en el archivo histórico y del acervo de televisión, por lo que, es necesario y fundamental desarrollar herramientas multidisciplinarias con el objeto de presentar las últimas tecnologías de la radio con proyectos de impacto social y similares que permitan superar este rezago para darle una respuesta oportuna a los requerimientos de la ciudadanía.

La situación descrita anteriormente no solamente es de RTVC, en la actualidad grandes centros de información en el mundo reportan un rezago entre la cantidad de información generada y la cantidad de información procesada y puesta al servicio público, es por ello que diversas instituciones, en el área de la inteligencia artificial y la teoría de la información, vienen desarrollando herramientas que dan solución a este problema, y proporcionan recursos de última generación para la mejor gestión de la información digital multimedia en plataformas de enfoque digital al igual que su difusión.

Es por esto que en el año 2018 nace el Laboratorio de Radio e Inteligencia Artificial de RTVC, el cual brindará soluciones a las limitaciones enmarcadas en los párrafos anteriores, así como desarrollos de investigación e innovación en el campo de la catalogación, reconocimiento, indexación, consulta, recuperación y difusión de contenido multimedia.

Una de las actividades iniciales de investigación que se está desarrollando en este laboratorio es la conformación de la información, actividades, enfoques y gestión del esquema de la información, lo cual está siendo la base del marco de trabajo para enseñar, validar y probar las aplicaciones tecnológicas.

Las aplicaciones tecnológicas, basadas en el análisis de la señal sonora que se desarrollan en los proyectos para la construcción del Laboratorio de Radio e Inteligencia Artificial de RTVC, están enfocadas en dar a las diferentes plataformas la capacidad de ofrecer una mejor gestión de los documentos multimedia que RTVC alberga. Para realizar lo anterior, se está creando una fuente de información robusta basada en bases de datos que forman los modelos de referencia que representan los contenidos de RTVC, bajo los protocolos y las características requeridas para los sistemas basados en los modelos matemáticos y tecnológicos de la inteligencia artificial con redes neuronales profundas.

Para realizar la nueva catalogación basada en los protocolos internacionales especificados para la gestión, difusión e interoperabilidad de los sistemas con contenidos multimedia (video, imagen, audio, voz, texto) y mantener y utilizar la catalogación que se tiene con el actual sistema basado en Marc21, se desarrolló una base de relaciones entre los protocolos Marc21 y Dublín Core, la *Figura 7* muestra algunas de las relaciones entre estos protocolos.

A continuación, se presenta un ejemplo de lo mencionado en la sección IA en la construcción del Laboratorio de Radio e Inteligencia Artificial de RTVC. Este ejemplo muestra el uso de los métodos automáticos de segmentación de la señal sonora y una catalogación manual de segmentación de la misma señal sonora. Como se mencionó anteriormente:

Los métodos automáticos de segmentación de la señal sonora toman relevancia a nivel social cuando forman parte de la base de proyectos que generan aplicaciones de ayuda e impacto que se ven reflejados en distintas áreas sociales. En el área de procesamiento y clasificación de la música se puede ejemplificar con aquellos proyectos que buscan la preservación de arte sonoro a través de los años [4].

Para mostrar el ejemplo se utilizó un archivo con información sonora con diferentes calidades y tipos de contenido sonoro, entre ellos, registros sonoros de música, de música con voz de hombre, voz de hombre y mujer con fondos de música, voz de hombre y voz de mujer separados, además los registros sonoros presentan diferentes calidades y ruidos propios a la digitalización, ya que algunos registros sonoros son de 1948 y otros muy actuales.

Figura 7. Ejemplo de relación entre Marc21 y Dublin Core.

| Categoría | Marc | DublinCore | Campo |
|---------------------------------|---------------------|---------------------------|------------------------------------------------------|
| Contenido | 245\$a | dc:title | Título del archivo |
| | 245\$b | dc:alternative [title] | Subtítulo del archivo |
| | 600 | dc:subject | Indicadores |
| | /450/453 | | Palabras claves |
| | | | Autoridad |
| | | | Género |
| | | | Nombre de personaje |
| | | | Edad |
| | | | Rango de edades |
| | | | Región |
| | | | Grupo étnico |
| | | | Hora de emisión (Inicio) |
| | | | Hora de emisión (Final) |
| | | | Tiempo total |
| | | | Tema |
| Propiedad intelectual | 500/505 | dc:abstract [description] | Observaciones |
| | 583 \$1 | dc:type | Tipo de audio |
| | 500 | dc:period [coverage] | Problema de ruido y contenido |
| | 264 \$c | dc:temporal [coverage] | Parrilla de programación |
| | 264 \$a | dc:spatial [coverage] | Fecha |
| | 041 | | Lugar de producción |
| | 544 | dc:language | Idioma |
| | 008 \$07-10/ | dc:dateCopyrighted | Fecha de copyright |
| | 264 \$c | | |
| | 100 | dc:creator | Creador |
| | 508 | | |
| | 700 | dc:publisher | Productora |
| | 264 \$b | dc:accessRights | Nota de restricciones de acceso |
| | 506 | dc:rights | Derecho de autor |
| | 540 | | |
| Instanciación (características) | 561 | dc:provenance | Propiedad e historia de custodia/Nota de procedencia |
| | 264 \$c / 008-7-10 | dc:date | Fecha de creación del recurso |
| | 264 \$c / 008-11-14 | dc:issued [date] | Fecha de emisión/publicación |
| | 008 \$22 | dc:audience | Audiencia objeto |
| | 5/583 \$k | dc:created [date] | Fecha de creación o modificación de la ficha |
| | 300 \$a | dc:format | Nota de acción |
| | 306 | dc:extent [format] | Formato de audio |
| | 300 \$c | dc:extent [format] | Tiempo de reproducción-Duración total |
| | 337 | dc:medium | Tamaño digital |
| | 84 | dc:medium | Tipo de medio |
| | | dc:identifier | Código del archivo |
| | | dc:totaltime | Tiempo total |
| | | dc:seventypercent | Porcentaje 70% enseñanza |
| | | dc:thirtypercent | Porcentaje 30% enseñanza |
| | | dc:initialseventy | Tiempo inicial marcado 70% |
| Enseñanza LABRIA | 500 | dc:seventyrange | Rango de tiempo 70% |
| | | dc:initialthirty | Tiempo inicial marcado 30% |
| | | dc:thirtyrange | Rango de tiempo 30% |

Figura 8. Ejemplo de catalogación manual archivo de audio, anotador especializado.

| Fecha de creación de la ficha: 005 | Código del archivo: 084 | Parrilla de programación | Título del archivo: 245 \$a | Palabras claves: 653 | Subtítulo 245\$b | Fecha 245\$c |
|------------------------------------|---------------------------------|--------------------------|-----------------------------------------------------------------------------------------------------|----------------------|---------------------------|--------------|
| 2019-07-16 | CDPS-DGW-108570-01-SER001CPTD6A | T18:00:00 - T18:38:53 | "Inauguración Casa de Poesía José Asunción Silva" | Literatura, Poesía | Emisión 24 de mayo , | 1986 |
| 2019-07-16 | CDPS-DGW-108570-01-SER001CPTD6A | T18:00:00 - T18:38:53 | "Inauguración Casa de Poesía José Asunción Silva" | Literatura, Poesía | Emisión 24 de mayo , | 1986 |
| 2019-07-16 | CDPS-DGW-108570-01-SER001CPTD6A | T18:00:00 - T18:38:53 | "Inauguración Casa de Poesía José Asunción Silva" | Literatura, Poesía | Emisión 24 de mayo , | 1986 |
| 2019-07-16 | CDPS-DGW-108570-01-SER001CPTD6A | T18:00:00 - T18:38:53 | "Inauguración Casa de Poesía José Asunción Silva" | Literatura, Poesía | Emisión 24 de mayo , | 1986 |
| 2019-07-16 | CDPS-DGW-108570-01-SER001CPTD6A | T18:00:00 - T18:38:53 | "Inauguración Casa de Poesía José Asunción Silva" | Literatura, Poesía | Emisión 24 de mayo , | 1986 |
| 2019-07-16 | CDPS-DGW-108570-01-SER001CPTD6A | T18:00:00 - T18:38:53 | "Inauguración Casa de Poesía José Asunción Silva" | Literatura, Poesía | Emisión 24 de mayo , | 1986 |
| 2019-07-16 | CDPS-DGW-108571-01-SER001CPTD6W | T18:00 - T18:42:34 | Lanzamiento del libro Poemas de amor de Dario Jaramillo Agudelo. Presenta Hernando Valencia Goelkel | Literatura, Poesía | Emisión 5 de agosto , | 1986 |
| 2019-07-16 | CDPS-DGW-108572-01-SER001CPTD6W | T18:00 - T18:50:12 | Lanzamiento del libro Escribano del agua | Literatura, Poesía | Emisión 19 de agosto , | 1986 |
| 2019-07-17 | CDPS-DGW-108573-01-SER001CPTD6W | T18:00 - T18:19:47 | Presentación de los libros Pontuario de castaño y El mundo sobre la mesa | Literatura, Poesía | Emisión 12 de julio , | 1988 |
| 2019-07-18 | CDPS-DGW-108574-01-SER001CPTD6W | T18:00 - T18:46:18 | Presentación del libro Aborto escuchando el cercano de las sirenas | Literatura, Poesía | Emisión 4 de septiembre , | 1986 |
| 2019-07-18 | CDPS-DGW-108575-01-SER001CPTD6W | T18:00 - T18:53:43 | Lanzamiento del libro Muerte de Merlín | Literatura, Poesía | Emisión 9 de septiembre , | 1986 |

La *Figura 8* muestra la catalogación manual realizada por un anotador especializado, como se puede observar aquí se especifican los tiempos de inicio y final de los diferentes segmentos de voz, música y silencios que se tienen en el archivo digital sonoro.

La *Figura 9*, muestra la catalogación automática realizada por el algoritmo de inteligencia artificial, esto lo realiza en un segundo. Se puede observar que también se especifican los tiempos de inicio y final de los segmentos de voz, música y silencios que se tienen en un archivo digital sonoro.

Figura 9. Ejemplo de catalogación automática archivo de audio, con método basado en inteligencia artificial.

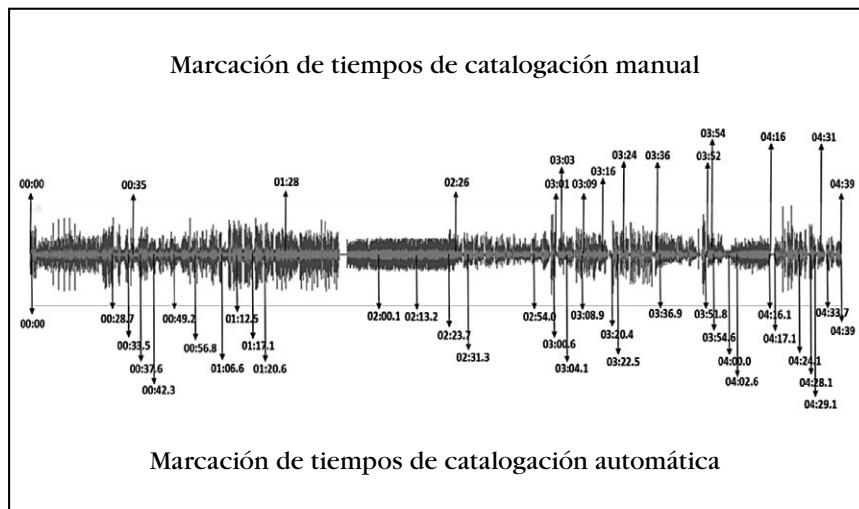
```
<format>.wav</format>
<duration>279.53</duration>
<frequency>48000</frequency>
<totalSamples>13437326</totalSamples>
<sampleRate>48000</sampleRate>
<nchannel>2</nchannel>
+<codec>
</media>
</content>
<body>

- <segments>
<segment id="1" type="Voice" name="AudioPresentMedellin2019_SV_000001" msfinish="4080" msduration="4080" msbegin="0"/>
<segment id="2" type="Voice" name="AudioPresentMedellin2019_SV_000002" msfinish="7488" msduration="3408" msbegin="4080"/>
<segment id="3" type="Voice" name="AudioPresentMedellin2019_SV_000003" msfinish="14832" msduration="7344" msbegin="7488"/>
<segment id="4" type="Voice" name="AudioPresentMedellin2019_SV_000004" msfinish="28704" msduration="13872" msbegin="14832"/>
<segment id="5" type="Voice" name="AudioPresentMedellin2019_SV_000005" msfinish="33528" msduration="4824" msbegin="28704"/>
<segment id="6" type="Voice" name="AudioPresentMedellin2019_SV_000006" msfinish="37656" msduration="4128" msbegin="33528"/>
<segment id="7" type="Voice" name="AudioPresentMedellin2019_SV_000007" msfinish="40608" msduration="2952" msbegin="37656"/>
<segment id="8" type="Voice" name="AudioPresentMedellin2019_SV_000008" msfinish="42336" msduration="1728" msbegin="40608"/>
<segment id="9" type="Voice" name="AudioPresentMedellin2019_SV_000009" msfinish="48096" msduration="5760" msbegin="42336"/>
<segment id="10" type="Voice" name="AudioPresentMedellin2019_SV_000010" msfinish="49248" msduration="1152" msbegin="48096"/>
<segment id="11" type="Voice" name="AudioPresentMedellin2019_SV_000011" msfinish="56832" msduration="7584" msbegin="49248"/>
<segment id="12" type="Voice" name="AudioPresentMedellin2019_SV_000012" msfinish="66624" msduration="9792" msbegin="56832"/>
<segment id="13" type="Voice" name="AudioPresentMedellin2019_SV_000013" msfinish="70440" msduration="3816" msbegin="66624"/>
<segment id="14" type="Voice" name="AudioPresentMedellin2019_SV_000014" msfinish="72552" msduration="2112" msbegin="70440"/>
<segment id="15" type="Voice" name="AudioPresentMedellin2019_SV_000015" msfinish="77160" msduration="4608" msbegin="72552"/>
<segment id="16" type="Voice" name="AudioPresentMedellin2019_SV_000016" msfinish="80688" msduration="3528" msbegin="77160"/>
<segment id="17" type="Voice" name="AudioPresentMedellin2019_SV_000017" msfinish="120096" msduration="39408" msbegin="80688"/>
<segment id="18" type="Music" name="AudioPresentMedellin2019_SM_000001" msfinish="133272" msduration="13176" msbegin="120096"/>
<segment id="19" type="Music" name="AudioPresentMedellin2019_SM_000002" msfinish="143736" msduration="10464" msbegin="133272"/>
<segment id="20" type="Voice" name="AudioPresentMedellin2019_SV_000018" msfinish="151392" msduration="7656" msbegin="143736"/>
<segment id="21" type="Voice" name="AudioPresentMedellin2019_SV_000019" msfinish="174000" msduration="22608" msbegin="151392"/>
<segment id="22" type="Voice" name="AudioPresentMedellin2019_SV_000020" msfinish="180648" msduration="6648" msbegin="174000"/>
<segment id="23" type="Voice" name="AudioPresentMedellin2019_SV_000021" msfinish="183000" msduration="2352" msbegin="180648"/>
<segment id="24" type="Voice" name="AudioPresentMedellin2019_SV_000022" msfinish="184152" msduration="1152" msbegin="183000"/>
<segment id="25" type="Voice" name="AudioPresentMedellin2019_SV_000023" msfinish="188976" msduration="4824" msbegin="184152"/>
<segment id="26" type="Voice" name="AudioPresentMedellin2019_SV_000024" msfinish="200400" msduration="11424" msbegin="188976"/>
<segment id="27" type="Voice" name="AudioPresentMedellin2019_SV_000025" msfinish="202536" msduration="2136" msbegin="200400"/>
<segment id="28" type="Voice" name="AudioPresentMedellin2019_SV_000026" msfinish="216912" msduration="14376" msbegin="202536"/>
<segment id="29" type="Voice" name="AudioPresentMedellin2019_SV_000027" msfinish="231840" msduration="14928" msbegin="216912"/>
<segment id="30" type="Voice" name="AudioPresentMedellin2019_SV_000028" msfinish="234648" msduration="2808" msbegin="231840"/>
<segment id="31" type="Voice" name="AudioPresentMedellin2019_SV_000029" msfinish="240000" msduration="5352" msbegin="234648"/>
<segment id="32" type="Voice" name="AudioPresentMedellin2019_SV_000030" msfinish="242616" msduration="2616" msbegin="240000"/>
<segment id="33" type="Voice" name="AudioPresentMedellin2019_SV_000031" msfinish="256080" msduration="13464" msbegin="242616"/>
<segment id="34" type="Voice" name="AudioPresentMedellin2019_SV_000032" msfinish="257088" msduration="1008" msbegin="256080"/>
<segment id="35" type="Voice" name="AudioPresentMedellin2019_SV_000033" msfinish="264168" msduration="7080" msbegin="257088"/>
<segment id="36" type="Voice" name="AudioPresentMedellin2019_SV_000034" msfinish="268176" msduration="4008" msbegin="264168"/>
<segment id="37" type="Voice" name="AudioPresentMedellin2019_SV_000035" msfinish="269184" msduration="1008" msbegin="268176"/>
<segment id="38" type="Voice" name="AudioPresentMedellin2019_SV_000036" msfinish="273768" msduration="4584" msbegin="269184"/>
<segment id="39" type="Voice" name="AudioPresentMedellin2019_SV_000037" msfinish="278736" msduration="4968" msbegin="273768"/>
```

En la *Figura 10*, se presenta la marcación de los tiempos realizada por el método de catalogación manual, por medio de un anotador especializado y la marcación de los tiempos, realizada por el método de catalogación automática, a través del método basado en inteligencia artificial.

Como se observa en la catalogación automática, se obtiene un mayor detalle en precisión sobre los segmentos de voz para los diferentes locutores, así como la precisión de los segmentos de música.

Figura 10. Ejemplo de marcación de tiempos segmentos de música y voz para catalogación manual y automática de archivo de audio.



A través de las anteriores figuras, observamos que el uso de la inteligencia artificial es una de las herramientas que nos permite realizar la catalogación automática de archivos sonoros.

CONCLUSIONES

En este texto se presentó un análisis de lo realizado en RTVC, durante varios años, para conformar el trabajo y el procedimiento de la catalogación que se realiza en la actualidad, con los retos y desafíos tecnológicos que representa la participación de la inteligencia artificial en esta área, a fin de dar respuesta a la pregunta inicial, a través de los trabajos que se realizan en el Laboratorio de Radio e Inteligencia Artificial (LABRIA) de RTVC, conforme a las experiencias del proyecto Mex-Culture.

Así como, describir una parte de la automatización del proceso de catalogación con los métodos de la IA y redes neuronales profundas, con base en protocolos internacionales para su indexación, interoperabilidad y difusión de la información del contenido sonoro.

REFERENCIAS

- A. Arratia, D. Brausin, M. De Giusti, A. López, S. Martínez, P. Rodríguez, P. Suarez, "Hacia una red iberoamericana para la integración, protección y difusión del patrimonio sonoro y audiovisual". REDAUVI, 1st ed. 2013.
- R. De la Cruz Modino, A. Rodríguez Darías and A. Doreste Alonso. "Paisajes sonoros y patrimonio cultural inmaterial convención sonora". *La Fonoteca del mar*, Revista de Turismo y Patrimonio Cultural, no. 10, pp. 211-215, 2012.
- M. Espejo Mojica. "El paisaje sonoro y la música en la red cultural". 2017.
- MaguaRED. 2017. "Audioteca De agua, viento y verdor, por el reconocimiento de la diversidad desde nuestros primeros años". J. Morejón Saravia. "Segmentación de audio y de locutores para recuperación de información multimedia y su aplicación a videos de información turística", Ingeniería de Telecomunicación, Universidad Autónoma de Madrid, 2011.
- Rodríguez Reséndiz, Perla Olivia, "La preservación digital sonora", *Investigación bibliotecológica*, no. 68, pp. 173-195, 2014.

Romo Melo, L., “Síntesis de voz aplicada a la rehabilitación de pacientes con labio y paladar hendido corregido”. Magister en Automatización Industrial, Universidad Nacional de Colombia, 2005.

UNESCO. (2018). PERSIST: la estrategia digital de la UNESCO para la sostenibilidad de la información.

Apuntes sobre la inteligencia artificial y su aplicación en los archivos sonoros

GEORGINA SANABRIA MEDINA

INTRODUCCIÓN

El presente trabajo tiene como punto clave reflexionar sobre la relación intrínseca que ha tenido el ser humano con la tecnología, ya que es este vínculo histórico el que ha permitido que con el paso del tiempo, se hayan desarrollado sistemas computacionales que tienen la capacidad de emular el comportamiento humano. A éstos, los procrea la Inteligencia Artificial (IA).

Aunque es cierto que el uso de la IA es muy cercano a los aspectos comerciales, no podemos negar que son las universidades y los centros de investigación quienes, de manera recurrente, apelan por su aplicación para la resolución de problemas que atañen a la sociedad, uno de ellos es la preservación de la memoria sonora de la humanidad. De este modo, en las siguientes líneas se plantean algunas ventajas que conlleva el uso de la IA en las tareas que se desarrollan dentro de los archivos sonoros.

EL SER HUMANO Y SU RELACIÓN CON LA TECNOLOGÍA

El ser humano ha estado ligado a la tecnología debido a que ésta es un producto de sus necesidades de supervivencia. Las herramientas que ha elaborado con el paso del tiempo han servido no sólo para superarse, mejor aún, con ellas ha logrado perfeccionarse y favorecer el progreso de la humanidad.

Desde la antigüedad, los homínidos se valieron de lo que tenían a su alrededor para diseñar herramientas que les ayudaron a resolver sus problemas de supervivencia, por ejemplo, podemos recordar los primeros instrumentos que crearon y con los cuales lograron cazar animales, elaborar sus prendas de vestir y construir sus chozas. Años más tarde, con el uso de los metales, perfeccionaron sus artilugios y fueron más eficientes en sus labores de agricultura, ganadería y transporte (Ordoñez, 2007). En otras palabras, con ayuda de esas herramientas, estaban sentando las bases para garantizar su seguridad y bienestar.

El ser humano, con la evolución de su pensamiento y por el interés de alcanzar un mejor nivel de vida, diseñó máquinas que lo sustituyeron en la realización de trabajos pesados y rutinarios, con ello, durante la segunda mitad del siglo XVIII, se dio paso al inicio de la industria. El referente principal es la máquina de vapor que reemplazó por completo las tareas que se hacían a mano, centradas en la fuerza de los seres humanos o con ayuda de los animales, esto desencadenó el descubrimiento de nuevas fuentes y tipos de combustibles, pero, sobre todo, favoreció la creación de grandes fábricas que seguirán este modelo de la sustitución de la energía muscular por la mecánica (Brynjolfsson y McAfee 2013).

Desde ese periodo histórico, las máquinas transformaron para siempre la vida del ser humano y, por lo tanto, la relación que tenía con su entorno. Los avances tecnológicos que derivaron de ello se extendieron en todos los ámbitos: económico, social, cultural, ambiental y político.

Como se ha señalado, hasta ese momento las herramientas que se crearon sirvieron primero para la supervivencia, posteriormente para sustituir los trabajos pesados y después para mecanizarlos. Pero no fue sino hasta mediados del siglo XX que se diseñó un aparato capaz de auxiliar al ser humano, ya no en cuestiones manuales, sino en las intelectuales, debido a la necesidad de realizar cálculos de manera eficiente, rápida, pero, sobre todo, automática, dado que las matemáticas se habían vuelto mucho más complejas. Es así como aparece la computadora, la cual cambió al mundo por completo (Orbe 2017).

De ser un aparato programado para hacer cálculos precisos, pasó a utilizarse también para almacenar información textual y gráfica. Las computadoras, sin duda, multiplicaron la producción y acumulación de información, y mejoraron también el procesamiento de los datos.

El impacto que trajo consigo la computadora fue tal que Brynjolfsson y McAfee (2013) han mencionado que nos encontramos en la segunda era de las máquinas, ellos refieren que la Revolución Industrial fue la primera y que ese ha sido el momento de mayor transformación que se ha visto en el mundo:

Las computadoras y otros avances digitales están haciendo por el poder mental lo que el motor de vapor y sus descendientes hicieron por el poder de los músculos. Nos están permitiendo traspasar las limitaciones previas y nos están llevando a un nuevo territorio (Brynjolsson y McAfee 2013, 14).

La aparición de las computadoras generó que se comenzara a hablar de máquinas inteligentes. Ese nuevo territorio que exponen Brynjolfsson y McAfee es el que tiene como centro a la IA, la cual se ha ido perfeccionando poco a poco (del mismo modo que ocurrió con el motor de vapor) para lograr contribuir en nuestro progreso y desarrollo.

LA INTELIGENCIA ARTIFICIAL

A John McCarthy se le atribuye haber acuñado, en los años 50, el término inteligencia artificial, el cual fue dado a conocer durante una conferencia que organizó en el Dartmouth College en New Hampshire (Estados Unidos, 1956). Este encuentro que duró seis semanas, se considera el germen de la disciplina, y es de destacar que “se reunieron los mejores investigadores en este campo, dando lugar a la era moderna de la IA y a alcanzar el estatus de ciencia.” (García Serrano 2012, 6).

De acuerdo con López de Mántaras y Meseguer González (2017), la idea de McCarthy para realizar esta conferencia era “establecer que cualquier aspecto del aprendizaje o cualquier otro rasgo de la inteligencia podía, en principio, ser descrito con un nivel de detalle

suficiente para ser simulado en una máquina.” (Mántaras y Meseguer González 2017,. 20-21).

En relación con la idea de McCarthy, si entendemos al aprendizaje como la capacidad que tenemos los seres humanos para la adquisición de habilidades cognitivas y esto es gracias a la inteligencia, es decir, a que razonamos, planificamos y resolvemos problemas; entonces en esencia lo que planteó fue poner casi al mismo nivel intelectual al ser humano y a las máquinas, pues, de acuerdo con lo que pensaba, estas últimas serían capaces de realizar cualquier acción que el ser humano genere por medio del aprendizaje.

De ahí que la primera definición postulada en esa conferencia sobre la IA es “hacer que una máquina se comporte como lo haría el ser humano, de tal manera que se la podrá llamar inteligente.” (Banda Gamboa 2014, 6).

Minsky, padre fundador de esta disciplina, fue otro científico que también mostró una pronta tesis sobre la expresión. Del mismo modo que McCarthy, estaba convencido de que se podían crear máquinas tan inteligentes como lo son las personas. Muestra de ello es la declaración que hizo sobre la IA, la cual considera como “la ciencia de crear máquinas que hacen cosas que requerirían inteligencia si las hicieran los hombres.” (1968, V).

En el mismo enfoque de identificar a la inteligencia artificial como sistemas que pueden pensar, tal como lo hacen los seres humanos, Haugeland (1985) concibió a la IA como “máquinas dotadas de mente, en el más amplio sentido literal.” (Haugeland 1985, 2).

Estas definiciones nos permiten percibir que durante un tiempo se ha forjado la idea de que la inteligencia artificial es el significado de máquinas que, con el tiempo, logran pensar y actuar de la misma forma en que lo hacen las personas. Para algunos esto parece ser algo exagerado, de hecho, Ganascia (2018) se ha referido a que estas posturas fueron fruto “de la euforia de los orígenes y de los primeros éxitos [de tal modo] que los investigadores dieron rienda suelta a su imaginación con algunas declaraciones precipitadas, por las cuales han sido criticados severamente desde entonces.” (Ganascia 2018, 7), incluso va más allá y asegura que el concepto de inteligencia artificial “al principio fue inventado, probablemente para llamar la atención.” (Ganascia 2018, 7).

Circundada por parte del seno científico y cobijada de otras corrientes intelectuales como la computación, la informática y la cibernética, surgieron más definiciones, entre las que destacan las siguientes cuatro:

- “El ‘aprendizaje de las máquinas’ en realidad no es más que la confluencia de diversas disciplinas relacionadas como la informática, la estadística y la matemática.” (Rodríguez 2018, 107).
- “La IA se basa en algoritmos que resuelven problemas. Estos algoritmos son de muy distinto tipo: algunos realizan búsquedas, otros siguen algún tipo de razonamiento, otros aprenden, algunos siguen reglas lógicas, mientras que otros se basan en probabilidades.” (Oribe 2017,91)
- Se puede definir a la “IA como un conjunto de técnicas, algoritmos y herramientas que nos permiten resolver problemas para los que, *a priori*, es necesario cierto grado de inteligencia, en el sentido de que son problemas que suponen un desafío incluso para el cerebro humano.” (García Serrano 2012, 5).
- La “IA trata de hacer programas que permitan a las computadoras hacer ese tipo de cosas que la mente humana puede hacer.” (Boden 1990, 1).

La tecnología no deja de avanzar, quizá sea por eso que resulte complejo llegar a un consenso sobre la expresión de la inteligencia artificial. Lo cierto es que lo que hemos estado presenciando con el tiempo ha sido un progreso de esta disciplina y, por ende, de su definición.

Con las contribuciones hechas hasta el momento podemos rescatar y resumir que la IA es la ciencia que se encarga de la creación de sistemas computacionales que permiten que máquinas procesen ingentes cantidades de datos, con los cuales hagan predicciones lo más acertadas posibles que contribuyan a resolver dificultades físicas e intelectuales del ser humano, incluso de una forma en la que ellos no lo podrían hacer.

No hay que olvidar que estas máquinas aprenden de los humanos (de la programación que realiza la persona) y en definitiva no se preocupan de lo que es ser “inteligente” sino de hacer cosas inteligentes (Oribe 2017, 88).

Este repaso sobre algunas definiciones, antiguas y actuales nos permite comprender qué es la IA y cómo se le ha visto desde diferentes épocas, pero también una contribución que se ha desprendido de todo esto es que “la misma IA nos ha enseñado que la mente es mucho más rica de lo que los psicólogos se habían imaginado en un principio.” (Boden 2017,12).

LA INTELIGENCIA ARTIFICIAL Y SU APLICACIÓN EN LOS ARCHIVOS SONOROS

Pese a que desde la década de los años 50 del siglo pasado se comenzaron a establecer las primeras bases de la IA, hoy a 20 años de que inició el siglo XXI, hablar sobre este tema sigue resultando ser algo novedoso. No hemos dejado de sorprendernos sobre las diversas aplicaciones que ha tenido en nuestro ámbito cotidiano y continuamos buscando otras formas de sacarle provecho en cuestiones más especializadas.

La IA trabaja con datos y entre más datos tenga es mejor, ya que los necesita para que las máquinas logren realizar la tarea establecida con éxito. En este sentido, los archivos sonoros digitales son un nicho muy favorable para la inteligencia artificial debido a la gran cantidad de datos que se encuentran en los documentos sonoros resguardados.

El uso de la tecnología ha ayudado a que los documentos sonoros llegaran a tener un mayor alcance entre los usuarios, ya que gracias a la digitalización hubo un incremento en el acceso al contenido, lo cual ha permitido, entre otras cosas, ensanchar y enriquecer el estudio del sonido como objeto digital (Rodríguez Reséndiz 2017).

También hay que hacer mención en que, de acuerdo con Rodríguez Reséndiz (2017), desde la primera década del siglo XXI ha habido un crecimiento exponencial de los documentos sonoros de origen digital, esto ha provocado que las bibliotecas y los archivos tengan un nuevo tipo de documento sonoro que ofrece contenido creado ya no sólo por músicos o radiodifusoras, sino por la sociedad civil, los movimientos sociales, los artistas independientes, entre muchos otros que se han vuelto generadores de contenido sonoro en formato digital.

Todo ello, ha traído consigo la multiplicación de los documentos sonoros (se deben considerar los que ya existían de origen analógico

y que se han ido digitalizando, y también aquellos, cuyo origen es digital). De ahí que el uso de la IA en los archivos sonoros cobra una especial relevancia, pues es posible aplicarla en las distintas tareas que se realizan y en los servicios que se ofrecen, algunos ejemplos son los siguientes:

- **Digitalización:** Es posible llevar a cabo la creación de “algoritmos de control de calidad para identificar defectos específicos en los archivos resultantes de la digitalización” (Rezzonico, 2019, p. 1). Para un humano eso implicaría muchos años de trabajo, los algoritmos detectarían estas fallas en la digitalización de una manera más rápida y con mayor precisión.
- **Catalogación:** Es una de las tareas más arduas que se realizan en un archivo, debido a la importancia que tiene, dado que la descripción y la extracción de datos clave de un documento es fundamental para su posterior recuperación. La cantidad abrumadora de datos que requiere ser procesada se ha vuelto una tarea que demanda mucho tiempo humano, aunado al alto porcentaje de errores que puede haber durante su realización, por ejemplo, los ortotipográficos. Con el uso de la IA se podría automatizar gran parte de este proceso, en este sentido sería posible transcribir de voz a texto, extraer palabras clave y categorizar el contenido, lo que reduciría enormemente los errores y el tiempo que invierte un catalogador en cada uno de los documentos con los que trabaja. “Esa persona ya no tendrá que ver y escuchar cada contenido que se quiera almacenar en las bases de datos, sino que a través de inteligencia artificial estamos creando un software que catalogue de manera automática y genere fichas bibliográficas.” (Radio Nacional de Colombia 2019, 1).
- **Búsqueda y recuperación de información:** ¿cuánto tiempo llevaría a una persona buscar, de manera tradicional, audios sobre los movimientos feministas en México que se encuentran en un acervo de más de 200 mil documentos digitalizados?, ¿cuántos audios estaría dejando de escuchar por falta de tiempo para hacerlo? La IA permite hacer una búsqueda rápida en todo el acervo y logra recuperar la información lo más cercana

posible a lo que el usuario necesita. De acuerdo con Altón (2018), la IA sirve para “encontrar gemas ocultas, programas y contenido que de otra manera no se verían de nuevo o se pasarían por alto. [Además] podemos utilizar la velocidad de las máquinas para ayudarnos, ya que la mirada de una computadora es mucho más rápida que la nuestra” (Altón 2018, 1).

- **Servicios de referencia:** a través del aprendizaje automático, los algoritmos de la IA pueden tener en cuenta los comportamientos de búsqueda de los usuarios, así como sus gustos e intereses, de este modo es posible brindarles una mejor experiencia con el archivo. Esto permitiría obtener un filtrado de datos en fracciones de segundos, en relación con la información que se tiene de cada consulta que realiza el usuario, y con esto, evitar, hacer la depuración manual. La ventaja que tiene la IA en este aspecto es que se logra satisfacer de manera más rápida la demanda de información de los usuarios, a través de una mejor orientación y asistencia en la búsqueda del archivo, la posibilidad de hacer accesibles materiales que están enfocados a sus intereses o necesidades particulares, y aconsejar sobre la selección más adecuada de documentos.

CONCLUSIONES

La inteligencia artificial nos está ayudando a realizar tareas que, hasta hace unos años, eran exclusivas del ser humano. Su uso cobra especial relevancia debido a la cantidad de información que se genera actualmente, ya que con ella es mucho más fácil manejar grandes volúmenes de datos.

En cuanto a su aplicación en los archivos sonoros, es necesario reiterar la importancia de su uso, porque de este modo el manejo y procesamiento de los datos que se encuentran en los documentos sonoros se desarrollará de una forma más rápida y permitirá reducir errores en las tareas propias que se realizan en estos lugares, como es el caso de la digitalización, la catalogación, la búsqueda y recuperación de la información, y los servicios de referencia.

Debido al número de parámetros aplicables en la IA es posible que los usuarios tengan un mejor aprovechamiento del acervo, con ello, la recuperación de la información no se verá mermada por el crecimiento acelerado de los documentos sonoros digitales, al contrario, suscitará ensanchar su acceso.

El hecho de automatizar el procesamiento de datos de los documentos sonoros con algoritmos de aprendizaje automático, usados en la inteligencia artificial, permitirá incrementar la productividad de los archivos sonoros, enriquecerá la eficacia de los departamentos involucrados en las distintas áreas de preservación y mejorará la ejecución de las tareas a desarrollar.

El futuro de la IA en los archivos sonoros depende de aprender a integrarla en los trabajos sustanciales de preservación, ya que su uso favorecerá a quienes laboran en los archivos sonoros y su impacto será en beneficio de la sociedad.

REFERENCIAS

Álvarez Álvarez J. (2013). *Aproximación crítica a la inteligencia artificial: claves filosóficas y prospectivas de futuro*. Madrid: Universidad Francisco de Vitoria.

Boden, M. (1990). *The Philosophy of Artificial Intelligence*, Oxford University Press: Oxford.

Boden, M. (2017). *Inteligencia artificial*. Madrid: Turner Publicaciones.

Brynjolfsson, E. y McAfee, A. (2016). *La segunda era de las máquinas. Trabajo, progreso y prosperidad*. Buenos Aires: Temas.

García Serrano, A. (2017). *Inteligencia artificial: fundamentos, práctica y aplicaciones*. México: Alfaomega.

Haton, J. (1991). *La inteligencia artificial: una aproximación*. Buenos Aires, Barcelona y México: Paidós.

López de Mántaras, R (2017). *Inteligencia artificial*. Madrid: CSIC – Catarata.

Minsky, M. (1968). *Semantic information processing*, MIT Press: Cambridge, Massachusetts.

Ordóñez, Leonardo. (2007). El desarrollo tecnológico en la historia. Areté, 19(2), 187-210. Recuperado el 08 de septiembre de 2019, de http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1016-&lng=es&tlng=es.

Oribe, A. (2017). *Una mirada al futuro. Inteligencia, abundancia, empleo y sociedad*. México: Alfaomega.

Radio Nacional de Colombia (2019). Inteligencia Artificial: la apuesta por visibilizar la memoria sonora del país. Recuperado el 02 de febrero de 2020, de <https://www.radionacional.co/noticias/actualidad/inteligencia-artificial-senal-memoria>.

Rezzonico, P. (2019). Inteligencia Artificial al servicio de los archivos audiovisuales RTS en: *FIAT-IFTA*. Recuperado el 06 de febrero 2020, de URL: <http://fiatifta.org/index.php/artificial-intelligence-at-the-service-of-the-rt-s-audiovisual-archives/>.

Rodríguez, P. (2018). *Inteligencia artificial. Cómo cambiará el mundo (y tu vida)*. Barcelona: Ediciones Deusto.

Rodríguez Reséndiz, P. (2017). Propuesta de principios que se deben tener en cuenta para la preservación de documentos sonoros de origen digital. *Anales De Documentación*, 20(2). Disponible en: <https://doi.org/10.6018/analesdoc.20.2.272181>.

Identifying Metadata Elements in Photographic Repositories by Semantic Segmentation of Images with Deep Learning

SINAÍ LÓPEZ-CASTILLO

Center of Innovation in Competitive Technologies (CIATEC, A.C.)

ISNARDO REDUCINDO

Faculty of Information Science, Autonomous University of San Luis Potosí

FRANCISCO BENITA

Singapore University of Technology and Design

INTRODUCTION

Over time, humans have expressed information on different media in order to leave evidence of his activities, interests and way of thinking. The ways in which humans express the information have been evolving and adapting according to the different needs they have, among them we found walls, clay tablets, papyrus, parchment, paper, etc. However, one of the most interesting media of the nineteenth century is photography, because it allowed to leave evidence of an event with the exact context in which it happened. Various areas of science put their attention on photography as it gave the possibility of providing realities embodied in images due its probative character. Before the World War I, the value that users gave to photography was mainly artistic, but after the war the large number of news and people interested in visualizing in detail what was happening photography began to obtain a documentary character and becomes an indispensable communication form. Therefore, the information containing in the photographs attained historical and evidential value. At this point, the photography began to reach many parts of the world, revolutionizing the way people communicate, evidencing and informing others about a variety of events. Archival

documents and photographic archives started to spread worldwide. Moreover, the generalized access to the photographic cameras in the last decades increased the volume of photographs that existed in different private and public organizations and other parties that used this support as archival document related to its infrastructure, monuments, cultural objects, and so forth. In recent years, the access to these repositories became more complex due to the vast amount of photography stored by the institutions. Therefore, different standards and regulations guidelines started to emerge with the aim to organize and standardize the photography archival.

On the other hand, in the wake of the rapid expansion of Information Systems, mobile and web technologies, photography collections were evolving with the constant change of technologies. Digital photographic repertoires and digital photographic archives of public organizations have grown rapidly as a response to such innovative new technologies. To date, digital repositories are the common framework for users to consult photography archival. However, in order to preserve and have a well organization of the photographic archival through digital repositories, the use of established standard metadata records for image description is required (Crawford 1984). The metadata fields provide different access points such that users can retrieve the photographs by applying different filtering criteria.

In this work, we propose an automated image description method for the photographic archival. In order to obtain a semantic description of the photographs, we provide a normalized access points through metadata elements using a deep learning algorithm for image segmentation. In the next sections, we give more details about appropriate image description, theoretical terms and technologies employed in this work, and then we explain our proposed framework and show the obtained results to validate our approach.

BACKGROUND

The digital repertoires facilitate documents organization and its access, but it is necessary to provide them with an adequate description so that image retrieval is simple and fast. The description, in

both physical and digital formats, is a complicated process that must be carried out with caution. However, the large documentary volumes together with the little time that organizations have for organizing their repositories usually leads to outputs that complicate the retrieval process and make the photography description inefficient. For example, photography description is often performed by individuals that are not familiar with the description and standardization of terms through Knowledge Organization Structures (KOS) such as undergraduate interns and other non-information professionals. In addition, when the image description task is carried out by multiple non-professional individuals within a short period of time, it is likely that the quality of the recorded information will be inadequate.

Despite working in a digital environment much of work today related to the archival (image) description is manual. On top of this, the old practice (Floyd and Oram 1992) of using undergraduate student employees to perform the routinized task of images description is still common nowadays. Hence, one of the challenges faced by public organizations such as government agencies or universities is the issue relating to the standardization of image description.

In the context of Information Systems, ontologies, thesaurus, vocabularies, terminologies and other types of KOS are widely used to facilitate and standardize the archival image description. In particular, ontologies have been viewed as a new type of vocabulary that ensure semantic interoperability with other vocabularies (e.g., the capacity of two or more systems to exchange and use the information that has been exchanged) through its ISO standard (ISO 25964-2). Semantic heterogeneity causes serious problems as it might happen that the same image might have different description depending on the person who is executing the task. It also might happen that the same semantic term is used to describe completely different properties. For example, one can think of “plant” which can refer either to factory or herbal. Semantic interoperability is an ontology-based approach that deals with this problem (Doerr 2003).

The KOS is a generic term used for referring to a wide range of items that models mutual relationships between elements of the repository (e.g. subject headings, thesauri, classification schemes and

ontologies). They are characterized by different specific structures and functions, and they are used in a plurality of contexts by diverse communities. Although different in nature, what they all have in common is that they are designed to support the organization of knowledge and information in order to make their management and retrieval of information easier (Mazzocchi 2018). Despite the tools provided by regulations, standards, KOS and technologies, the description of photographic collections and repositories has not evolved, that is, it still requires the support of many collaborators for its description and the result remains the same.

Another of the advantages that technology evolution allows is the image labelling using tools that allow us to know the most relevant characteristics of a photograph (image) by the scene segmentation. Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels). Then, after the implementation of computer vision tools the image is characterized by means of a conceptual representation using labelling (Peres et al. 2010). In recent years several powerful tools performing image segmentation (*SEGNET* by (Badrinarayanan, Kendall, and Cipolla 2017) or *DeepLab* by (Chen et al. 2017)) and object detection (Mask R-CNN by (He et al. 2017)) have been developed. Image segmentation has achieved very good results in the field of medicine, urbanism, construction and robotics. Image segmentation is a well-studied computer vision task which has been significantly improved with the boom of convolutional neural networks (Krizhevsky, Sutskever, and Hinton 2012). This technique has been used to look at the problem of automatically describing archival images, at least in its conceptual-representational mode (Frigui and Caudill 2006).

With the analysis of the needs in the photographic collections and the advantages provided by the image labelling, the opportunity offered by using the latter tool is identified in order to start automating the description process. This approach not only helps to reduce the time needed to describe and image but also improves the normalization description and leads to a better retrieval performance. In this work, we chose the Dublin Core standard (Weibel 1997). which is a simple and effective 15-element set for describing images. The 15

core elements are: title, description, subject (keyword), type, relations, creator, contributor, publisher, rights, date, format, identifier and language.

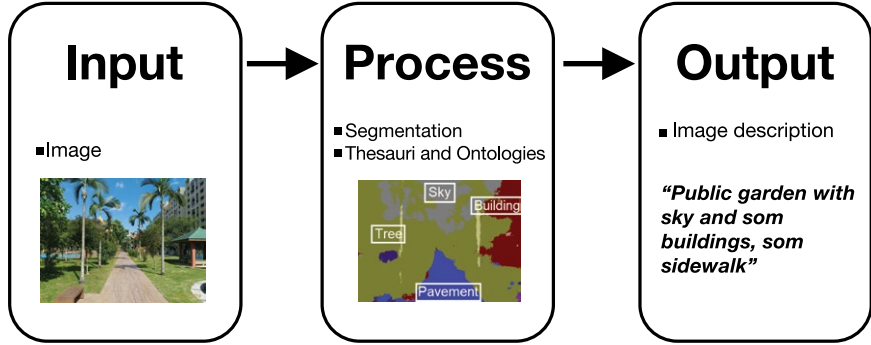
Nevertheless, the image segmentation does not provide the necessary elements to define all the metadata that make up the standard. Conversely, the tool can support the definition of the most time-consuming elements such as title and description. That is, from the elements that make up the image, image segmentation may allow the creation of access points so that users can identify the images with respect to their content.

PROPOSED FRAMEWORK

To test our approach, we used 24 photographs of a residential area in Singapore. The “description” element of the Dublin Core standard was generated in three different ways. First, a final-year undergraduate student from library management background, i.e., “Person”, described each image according to her own criteria. Note that this process is commonly adopted by public institutions as it is inexpensive and fast, but it leads to semantic heterogeneity issues. Second, a graduate student of archival management, i.e., “Archivist”, described the images using a thesaurus. The Archivist description can be considered as a standardized description that generates normalized access points. This is also a common practice within institutions that aim improving the quality of their digital repositories. Third, an expert in library management and archival management systems, i.e., “Target”, described the images following the standardized thesaurus and ontologies (UNESCO, n.d.). This Target constitutes the gold standard of the normalized descriptions. Fourth, captions were generated after implementing a Deep Learning tool of image segmentation. In this work open source SegNet algorithm was employed (Badrinarayanan, Kendall, and Cipolla 2017). SegNet allows the identification of 12 different classes related to outdoor environments, namely, sky, building, pole, road, road marking, pavement, tree, sign symbol, fence, vehicle, pedestrian and bike. But the convolutional neural network can be trained to identify new segments such as person, different ani-

mals, everyday items, food or indoor furniture (chairs, lamps, tables, floor, wall, etc.). *Figure 1* illustrates the description approach based on image segmentation and Deep Learning techniques.

Figure 1. Process description.



Next, the semantic similarity among all three descriptions versus the Target description (expert normalized description) is performed adapting the semantic interoperability metrics detailed (Yahia, Aubry, and Panetto 2012). Given a set A composed by n semantic concepts c_i ($i = 1, \dots, n$), we can define the lexical AL and nonlexical ANL subsets as follows:

$$A_L = \{c_i | c \text{ is a lexical concept } \forall c_i \in A\}$$

$$A_{NL} = \{c_i | c \text{ is a nonlexical concept } \forall c_i \in A\}$$

such that

$$A_L \cup A_{NL} = A \quad \text{and} \quad A_L \cap A_{NL} = \emptyset$$

We can define a lexical concept as a term that can be written down, consist of letters, numbers, among other characters. A non-lexical concept is a concept that cannot be written down and is named by lexical concepts (Lezoche, Aubry, and Panetto 2012). Then, given two semantic sets A and B, we obtain the semantic relations set R as follows

$$R = \{ \langle c_i^b, (c_1^a, \dots, c_j^a) \rangle \mid \text{sem}(c_i^b) \subseteq \text{sem}(c_1^a, \dots, c_j^a) \forall c_j^a \in A, \forall c_i^b \in B \}$$

where $\text{sem}(c)$ represents the semantic interpretation of the concept c . Now, removing the non-lexical subset of A we can define the lexical relations set R_L as follows

$$R_L = \{ \langle c_i^b, (c_1^a, \dots, c_j^a) \rangle \mid \text{sem}(c_i^b) \subseteq \text{sem}(c_1^a, \dots, c_j^a) \forall c_j^a \in A, \forall c_i^b \in B_L \}$$

Finally, we defined the semantic similarity potential V between two semantic sets A and B as

$$V_{A \rightarrow B} = \frac{|R_L|}{|R|}$$

where $|\cdot|$ represents the set cardinality. It is important to note that the semantic similarity potential is not bijective, for this reason we compute both sides potential in all cases.

RESULTS

In order to appreciate the differences between the tree description cases (under graduated student (Person), graduated student (Archivist) and automated (SegNet)) versus the target (expert normalized description) Figure 2 displays the boxplot of the semantic similarity potential between all six types of descriptors (both sides of potential).

By visual inspection it is not clear if mean values are different among types of descriptors, hence Table 1 summarizes the p-values from the Kruskal-Wallis test. This is a non-parametric test than help us to decide whether types of descriptors came from the same distribution. The table compares all three combinations of descriptions and, with the exception of pairs (Target-Person, Person-Target), (Target-Person, Automated-Target), (Target-Archivist, Person-Target) and (Target-Automated, Archivist-Target), all other pair of combination between descriptors came from the different distributions. In other words, we found evidence that the description generated by the different individuals and automated tools is semantically different.

Figure 2. Boxplot of semantic similarity potential.

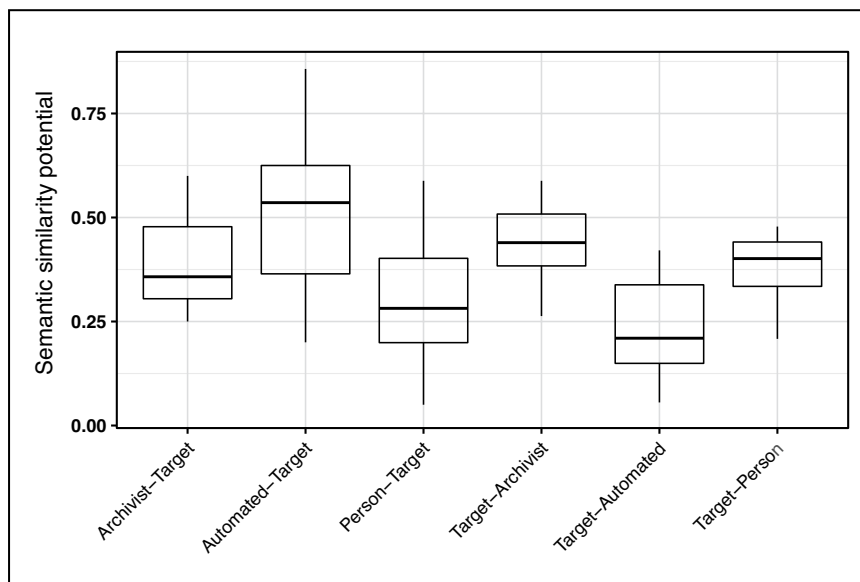


Table 1. P-values of the Kruskal-Wallis test of semantic similarity potentials

Note: Estimates in bold are statistically significant (p-value < 0.05).

| | Person-Target | Archivist-Target | Automated-Target |
|------------------|---------------|------------------|------------------|
| Person-Target | 0.0664 | 0.0005 | 0.0928 |
| Archivist-Target | 0.9589 | 0.0487 | 0.0004 |
| Automated-Target | 0.0108 | 0.2307 | 0.0004 |

In addition to the presented semantic similarity potential results, it is worth noticing that the automated photographs description took a computation time of 2.79 seconds for all 24 images. Alternatively, the description generated by the by under graduated student, e.g., “Person”, required 4 hours (one day of work), the graduated normalized student description, i.e., “Archivist”, required 16 hours (four days of work), and the expert description, (i.e., “Target”), took 8 hours (two days of work).

CONCLUDING REMARKS

In this work, we propose the use of image segmentation combined with Deep Learning techniques as a feasible framework to derive automated metadata descriptors for photographs. The experimental results reveal that our automated descriptors have high semantic similarity compared with expert normalized description which evidence the effectivity of our approach. Furthermore, the automated description is not only considerably faster (couple of seconds vs couple of hours) but also could help improving knowledge organization of the digital photographs repositories as it provides well normalized access points. Our approach also removes the ambiguities in the descriptions that can arise when the task is performed individuals that are non-experts in the field of photograph documentation such as undergraduate students.

Finally, a natural extension of this work is the improvement of the semantic concepts in the automated description by adding other characteristics of the objects located in the images such as colours or position in the scene. In doing so, the convolutional neural networks need to be trained to detect other type of objects. The implementation of natural language processing techniques is also required in order to derive more elaborated image descriptors.

BIBLIOGRAPHY

- Badrinarayanan, V., A. Kendall, and R. Cipolla. (2017). "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39 (12): 2481-95.
- Chen, L. C., G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. (2017). "Deeplab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected Crfs." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40 (4): 834-48.

- Crawford, W. (1984). *MARC for Library Use: Understanding the USMARC Formats*. USA: Knowledge Industry Publications.
- Doerr, M. (2003). "The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata." *AI Magazine* 24 (3): 75.
- Floyd, B., and R. Oram. (1992). "Learning by Doing: Undergraduates as Employees in Archives." *The American Archivist* 55 (3): 440-52.
- Frigui, H., and J. Caudill. (2006). "Unsupervised Image Segmentation and Annotation for Content-Based Image Retrieval." *Proceedings of the IEEE International Conference on Fuzzy Systems (ICFS 2006)*, 72-77.
- He, K., G. Gkioxari, P. Dollár, and R. Girshick. (2017). "Mask R-CNN." *Proceedings of the IEEE International Conference on Computer Vision (ICCV 2017)*, 2961-69.
- Krizhevsky, A., I. Sutskever, and G.E. Hinton. (2012). "Imagenet Classification with Deep Convolutional Neural Networks." *Advances in Neural Information Processing Systems (NIPS 2012)*, 1097-1105.
- Lezoche, Mario, Alexis Aubry, and Hervé Panetto. (2012). "Formal Fact-Oriented Model Transformations for Cooperative Information Systems Semantic Conceptualisation." *Lecture Notes in Business Information Processing*, 117-31. https://doi.org/10.1007/978-3-642-29958-2_8.
- Mazzocchi, Fulvio. (2018). "Knowledge Organization System (KOS)." *ISKO Encyclopedia of Knowledge Organization*. Birger Hjørland and Claudio Gnol.

- Peres, F A, F R Oliveira, L A Neves, and M F Godoy. (2010).
“Automatic Segmentation of Digital Images Applied in
Cardiac Medical Images.” *Pan American Health Care Ex-
changes*, 38–42. [https://doi.org/10.1109/PAHCE.2010.
5474606](https://doi.org/10.1109/PAHCE.2010.5474606).
- UNESCO. n.d. “UNESCO Thesaurus.” [http://vocabularies.
unesco.org/browser/thesaurus/es/?clang=en](http://vocabularies.unesco.org/browser/thesaurus/es/?clang=en).
- Weibel, S. (1997). “The Dublin Core: A Simple Content
Description Model for Electronic Resources.” *Bulletin
of the American Society for Information Science and
Technology* 24 (1): 9–11.
- Yahia, Esma, Alexis Aubry, and Hervé Panetto. (2012).
“Formal Measures for Semantic Interoperability Assess-
ment in Cooperative Enterprise Information Systems.”
Computers in Industry 63 (5): 443–57. [https://doi.
org/10.1016/j.compind.2012.01.010](https://doi.org/10.1016/j.compind.2012.01.010).

Propuesta metodológica y de análisis computacional para identificar el proceso fotográfico en fotografías históricas del siglo XIX y XX

GUSTAVO LOZANO SAN JUAN

Instituto de Investigaciones Estéticas. Universidad Nacional Autónoma de México.

RODRIGO COLÍN RIVERA

Laboratorio Audiovisual de Investigación Social, Instituto Mora.

OBJETIVO

El objetivo de este proyecto es diseñar una metodología de trabajo que permita a los archivistas, historiadores, conservadores y otros profesionales de los archivos, identificar correctamente la técnica de las fotografías producidas durante los siglos XIX y XX.

ANTECEDENTES

Como sabemos, el catálogo es el instrumento que describe ordenadamente y de forma individualizada las unidades documentales de una serie o conjunto documental (Heredia, 1991), las operaciones necesarias para la catalogación de una fotografía son: la elaboración de un lenguaje documental, el análisis de las unidades documentales y la sistematización de la información (Boadas, 2001). El análisis documental se refiere tanto al contenido como al continente, y el aspecto más relevante a documentar en este último caso, es la identificación de la técnica o proceso fotográfico.

Figura 1. Diversos procesos fotográficos



En la bibliografía sobre el tema, la identificación del proceso fotográfico no se presenta como un procedimiento sistemático compuesto por pasos claramente delimitados, sino como la aplicación de habilidades de observación y reconocimiento que son adquiridas de manera empírica a lo largo de años de práctica y estudio. Bajo el planteamiento

tradicional, es obligatorio que dichas habilidades se acompañen de un conocimiento sólido de la cronología de los procesos fotográficos y de sus principales hitos tecnológicos, requerimientos que plantean una curva de aprendizaje pronunciada.

En los cursos de capacitación y actualización existentes, la identificación de los procesos fotográficos se enseña de forma casi personalizada en grupos de tamaño pequeño y en clases de duración limitada. El instructor es generalmente un profesional de la conservación especializado en fotografías y de manera invariable se requiere de una colección de estudio variada que permita a los participantes desarrollar las habilidades de observación y análisis descritas previamente.

Tabla 1. Procesos fotográficos utilizados en los siglos XIX y XX

| Imágenes de cámara | Impresiones | Negativos | Transparencias |
|---------------------------|-------------------------------------|------------------------------------------|---------------------------------------------------|
| Daguerrotipo | Cianotipo | Negativo de colodión húmedo | Transparencia plata gelatina sobre vidrio |
| Ambrotipo | Albúmina | Negativo de placa seca de gelatina | Transparencia procesos de pantalla aditiva |
| Ferrotipo | Colodión de impresión directa | Plata gelatina sobre nitrato de celulosa | Transparencia cromógena sobre acetato de celulosa |
| | Plata gelatina de impresión directa | Plata gelatina sobre acetato de celulosa | Transparencia cromógena sobre poliéster |
| | Platinotipo | Plata gelatina sobre poliéster | |
| | Impresión plata gelatina | Cromógeno sobre acetato de celulosa | |

| | | | |
|--|--------------------------------------|---------------------------|--|
| | Impresión por difusión de plata | Cromógeno sobre poliéster | |
| | Impresión cromógena | | |
| | Impresión por difusión de colorantes | | |
| | Impresión por blanqueo de colorantes | | |

JUSTIFICACIÓN

El proceso fotográfico se refiere tanto a los materiales como a la manera en que estos se combinan -artesanal o industrialmente- para crear la fotografía. Su documentación es importante porque da cuenta de la evolución tecnológica de prácticas extintas hoy en día y porque algunos procesos requieren de medidas de conservación especiales, como por ejemplo, el almacenamiento a bajas temperaturas para las impresiones, transparencias y negativos cromógenos o las medidas de prevención y combate de incendios para los negativos con soporte de nitrato de celulosa.

Desafortunadamente los vehículos tradicionales para la enseñanza-aprendizaje de los conocimientos y las habilidades requeridos para la identificación del proceso fotográfico, presentan varias limitantes que impiden su disseminación entre los profesionales de los archivos.

DESARROLLO

La propuesta que aquí se plantea pretende aliviar algunas de las limitaciones antes mencionadas, ya que se trata de una metodología de fácil acceso para los profesionales de los archivos, por medio de la cual es posible identificar el proceso de una fotografía de interés,

entre una gama de 30 alternativas utilizadas a lo largo de los siglos XIX y XX. Algunas de sus características más relevantes son que se pueden aplicar de forma individual sin necesidad de un instructor, no es necesaria, una colección de estudio, no depende del estudio previo de la historia técnica de la fotografía. Adicionalmente, permite a usuarios que aplican la metodología por primera vez, obtener resultados positivos y facilitar la diseminación de un conocimiento de acceso limitado.

A continuación se describe la manera en que esta propuesta fue desarrollada.

INVESTIGACIÓN BIBLIOGRÁFICA

En la primera etapa de este trabajo se hizo una búsqueda bibliográfica que permitiera identificar las principales características físicas que definen a cada uno de los procesos fotográficos (Reilly, 1986) (Sepiades, 2003) (Aasbø, 2003) (Barra, 2005) (Image Permanence Institute, 2019).

Tabla 2. Características físicas.

| | |
|------------------|-----------------------------|
| Soporte primario | Tonalidad |
| Iluminación | Brillo |
| Polaridad | Superficie |
| Tono | Texto |
| Fecha | Deterioro |
| Estratigrafía | Tonalidad |
| Magnificación | Particularidades del Objeto |

SÍNTESIS DE LA INFORMACIÓN

Posteriormente se sintetizó la información y se capturó en formato tabular, en las filas se colocaron los diferentes procesos fotográficos, en las columnas las diferentes características físicas a identificar y en las celdas en donde ambas se intersectan se colocaron las caracterís-

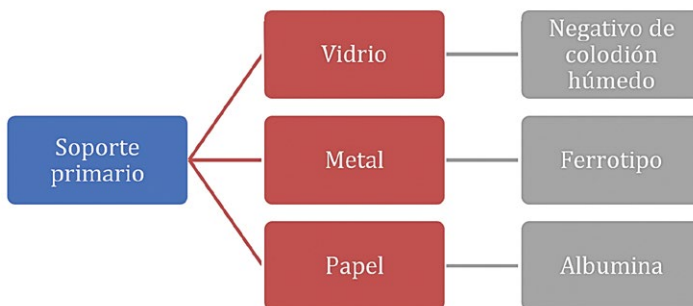
ticas concretas que definen a cada proceso. Por ejemplo, para el caso del proceso fotográfico negativo de colodión húmedo, la característica soporte primario corresponde a la opción vidrio, para el proceso ferrotipo corresponde a la opción metal y para el proceso albumina corresponde a la opción papel.

Tabla 3. Ejemplo de información en formato tabular.

| | Soporte primario |
|-----------------------------|------------------|
| Negativo de colodión húmedo | Vidrio |
| Ferrotipo | Metal |
| Albumina | Papel |

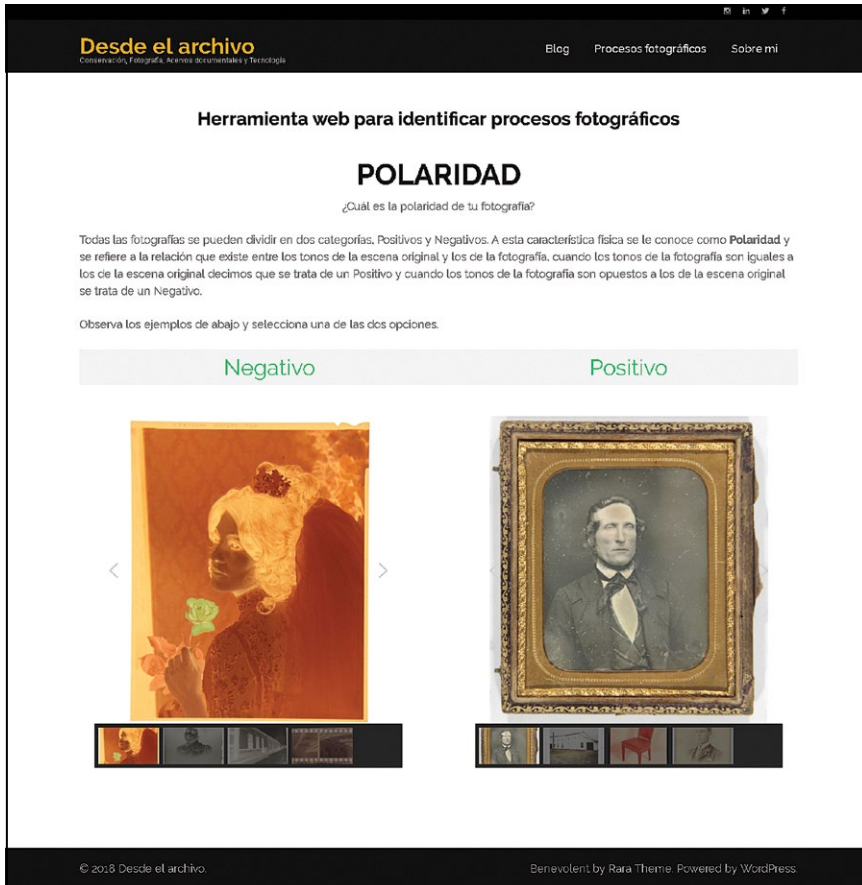
Esta misma información fue trasladada posteriormente a un árbol de decisiones, un modelo predictivo utilizado en las ciencias de datos para clasificar entidades con base en sus atributos; así el soporte primario que en la tabla se encontraba como encabezado de la segunda columna se convirtió en un nodo de decisión del cual se derivan tres ramas: vidrio, metal y papel, que corresponden a las opciones posibles para esa característica, finalmente los tres procesos fotográficos que en la tabla se encontraban en la primera columna se convierten en las hojas del árbol y corresponden a la clasificación final.

Figura 2. Ejemplo de información en formato de árbol de decisiones.



El modelo del árbol permitió elaborar un cuestionario (<http://desde-elarchivo.com/id-procesos-fotograficos/>) que a través de una serie de preguntas dirigidas guía al usuario paso a paso a través de la metodología hasta llegar al conjunto de fotografías que le permiten identificar el proceso fotográfico de su fotografía.

Figura 3. Interfaz de usuario de la página web.



En su estado actual el árbol cuenta con 40 nodos, 96 ramas y 30 hojas por lo que optimizarlo es sumamente importante. Para ello se propo-

ne también un análisis mediante conjuntos de datos representativos con diversas propiedades de las fotografías que serán procesados mediante herramientas y algoritmos de inteligencia artificial. El propósito es generar un árbol de decisiones alternativo a la metodología archivística para encontrar similitudes y determinar las propiedades de mayor relevancia al clasificar las fotografías.

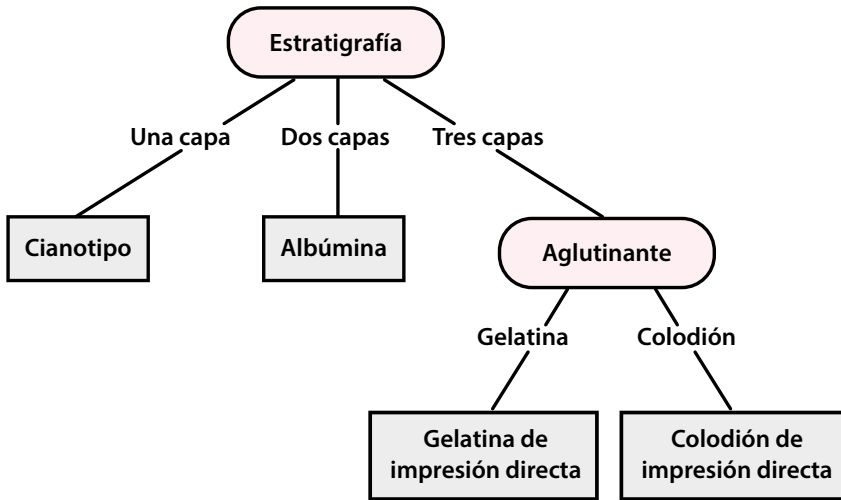
El tipo de análisis computacional propuesto deriva de una rama de la inteligencia artificial denominada Machine Learning (ML) o aprendizaje de máquina, donde el énfasis es aprender de los datos, sin que esta capacidad de aprendizaje esté explícitamente programada. Y de esta manera, utilizar los árboles de decisión como un modelo de predicción de un valor objetivo basado en un conjunto de datos variables.

La generación de árboles de decisión de manera automática o asistida por algoritmos computacionales permite desarrollar tareas de clasificación, cuya aplicación es realizar un análisis computacional para la identificación del proceso fotográfico en fotografías históricas del siglo XIX y XX. Al referirnos al término *identificación*, y bajo este contexto, es válido indicar que el término *predicción* es empleado como sinónimo, ya que el propósito es aproximar (con el menor margen de error posible) un conjunto de datos a un modelo de árbol de decisión que nos permita realizar dicha identificación fotográfica y además, brinde un punto de comparación en cuanto a las características físicas más relevantes de este conjunto particular de fotografías.

HACER PREDICCIONES

Al revisar otro ejemplo simplificado de árbol de decisión para la identificación de procesos fotográficos, se aprecia la raíz en la parte superior y el árbol crece hasta los nodos terminales u hojas en la parte inferior. Esta representación es equivalente a tener el nodo raíz del lado izquierdo y hacerlo crecer hacia la derecha.

Figura 4. Ejemplo sencillo de árbol de decisión con nodo raíz en la parte superior.



Es interesante notar que cada uno de los nodos intermedios (en color rosa) del árbol de decisión representa una característica física de la fotografía mientras que las hojas (en color gris) representan procesos fotográficos que determinan de manera definitiva la categoría a la que pertenece. Como se puede apreciar, utilizar esta estructura de datos para tareas de clasificación es bastante claro e intuitivo, en contraste con otros modelos como las redes neuronales, considerados modelos de *caja negra* ya que usualmente es difícil de explicar en términos simples por qué y cómo son hechas las predicciones o clasificaciones (Gerón, 2017).

VENTAJAS Y LIMITACIONES

Como se ha mostrado, una ventaja importante de los árboles de decisión es que el proceso de clasificación es lo suficiente fácil de llevar a cabo teniendo la estructura del árbol, y se puede realizar manualmente sin la necesidad de una computadora. Los datos de entrada que se procesan usualmente derivan de conjuntos o bases de datos que requieren poca intervención y se pueden exportar en formatos de archi-

vo convenientes como CSV (valores separados por coma). Otra ventaja importante es que el desempeño es bueno con cantidades razonables de poder de cómputo; si se tiene un conjunto de datos grande, el proceso de aprendizaje y construcción del modelo es manejable (Bell, 2015).

A pesar de las ventajas también es importante identificar inconvenientes, y uno de los más importantes es el hecho de que se pueden crear modelos muy complejos generados a partir de conjuntos de datos complejos. La manera de contrarrestar este efecto es revisar y acotar los valores a usar en categorías, lo cual producirá un modelo más refinado y concreto. Hay que estar consciente de que este proceso no siempre es viable, de ser el caso, el modelo resultante puede ser mucho más grande de lo esperado. Puede ser que utilizar otro método de inteligencia artificial se adapte mejor a las necesidades del conjunto de datos, por ejemplo, redes neuronales o máquinas de soporte vectorial.

ALGORITMOS

ID3

El algoritmo ID3 (Iterative Dichotomiser 3) fue inventado por Ross Quinlan, investigador en ciencias de la computación en el área de minería de datos y teoría de la decisión, para crear árboles a partir de bases de datos. Al calcular la entropía --unidad de medida del desorden en un conjunto de datos, según la teoría de la información-- para cada atributo del conjunto de datos, esto permite dividir en subconjuntos basados en el valor de mínima entropía, y construir de manera recursiva cada nodo del árbol de decisión.

En conjunción, ID3 utiliza también el método de ganancia de información, la medida de las diferencias de entropía antes y después de que un atributo divide en subconjuntos de datos, para determinar el nodo raíz en cada llamada recursiva del algoritmo.

C4.5

Ross Quinlan efectuó mejoras en su algoritmo original y creó el algoritmo C4.5 que de igual manera emplea el método de ganancia de

información, pero hace más evidente que el árbol resultante puede ser usado como modelo de clasificación.

Entre las mejoras más notables respecto a ID3 se encuentra la habilidad de trabajar con atributos continuos, como valores numéricos o fechas. También permite trabajar con valores de atributos vacíos o “huecos” en la base de datos sin que esto afecte la ganancia de información de un atributo durante la construcción del árbol. Y por último, el árbol creado con C4.5 es recortado después de su creación, es decir, algunas ramificaciones del árbol son modificadas por nodos terminales cuando contribuyen a simplificar el árbol (Bell, 2015).

PROCESO DE GENERACIÓN AUTOMÁTICA DE ÁRBOLES DE DECISIONES

Es usual en el área de ML, que el ciclo de acciones a tomar para generar modelos a partir de conjuntos de datos involucra los siguientes pasos: obtención de datos, preparación de éstos, ejecución de la heurística o algoritmo y presentar los resultados.

Se emplea el software de código abierto WEKA (<https://www.cs.waikato.ac.nz/ml/weka>) para la generación de árboles de decisión. Dicho software cuenta con una colección de algoritmos y herramientas de pre procesamiento de datos que permiten trabajar de manera flexible con conjuntos de datos, facilita la experimentación, la obtención de resultados estadísticos y visualización tanto de los datos de entrada como de los resultados obtenidos.

OBTENCIÓN DE DATOS

Los datos que se emplearon en este trabajo son una acotada selección de diversas características fotográficas de la colección de estudio perteneciente al área de conservación del Archivo Fotográfico Manuel Toussaint del Instituto de Investigaciones Estéticas de la UNAM. Todas las características físicas de los procesos fotográficos a identificar se registraron en una hoja de cálculo que sirve de insumo inicial en esta etapa del proceso.

Figura 6. Mismo conjunto de datos en diferentes formatos, de izquierda a derecha se muestra una hoja de cálculo, un archivo CSV y un archivo ARFF.

| Características físicas - Imp - X | | | |
|-----------------------------------|----------|-----------|-----------|
| 1 | Proceso | Tipología | Aguinante |
| 2 | Albúmina | Impresión | Brillo |
| 3 | Albúmina | Impresión | Albúmina |
| 4 | Albúmina | Impresión | Semi mate |
| 5 | Albúmina | Impresión | Semi mate |
| 6 | Albúmina | Impresión | Semi mate |
| 7 | Albúmina | Impresión | Semi mate |
| 8 | Albúmina | Impresión | Semi mate |
| 9 | Albúmina | Impresión | Semi mate |
| 10 | Albúmina | Impresión | Semi mate |
| 11 | Albúmina | Impresión | Semi mate |
| 12 | Albúmina | Impresión | Semi mate |
| 13 | Albúmina | Impresión | Semi mate |
| 14 | Albúmina | Impresión | Semi mate |
| 15 | Albúmina | Impresión | Semi mate |
| 16 | Albúmina | Impresión | Semi mate |
| 17 | Albúmina | Impresión | Semi mate |
| 18 | Albúmina | Impresión | Semi mate |
| 19 | Albúmina | Impresión | Semi mate |
| 20 | Albúmina | Impresión | Semi mate |
| 21 | Albúmina | Impresión | Semi mate |
| 22 | Albúmina | Impresión | Semi mate |
| 23 | Albúmina | Impresión | Semi mate |
| 24 | Albúmina | Impresión | Semi mate |

| Características físicas - Imp - X | | | |
|-----------------------------------|----------|-----------|-----------|
| 1 | Proceso | Tipología | Aguinante |
| 2 | Albúmina | Impresión | Brillo |
| 3 | Albúmina | Impresión | Albúmina |
| 4 | Albúmina | Impresión | Semi mate |
| 5 | Albúmina | Impresión | Semi mate |
| 6 | Albúmina | Impresión | Semi mate |
| 7 | Albúmina | Impresión | Semi mate |
| 8 | Albúmina | Impresión | Semi mate |
| 9 | Albúmina | Impresión | Semi mate |
| 10 | Albúmina | Impresión | Semi mate |
| 11 | Albúmina | Impresión | Semi mate |
| 12 | Albúmina | Impresión | Semi mate |
| 13 | Albúmina | Impresión | Semi mate |
| 14 | Albúmina | Impresión | Semi mate |
| 15 | Albúmina | Impresión | Semi mate |
| 16 | Albúmina | Impresión | Semi mate |
| 17 | Albúmina | Impresión | Semi mate |
| 18 | Albúmina | Impresión | Semi mate |
| 19 | Albúmina | Impresión | Semi mate |
| 20 | Albúmina | Impresión | Semi mate |
| 21 | Albúmina | Impresión | Semi mate |
| 22 | Albúmina | Impresión | Semi mate |
| 23 | Albúmina | Impresión | Semi mate |
| 24 | Albúmina | Impresión | Semi mate |

| Características físicas - Imp - X | | | |
|-----------------------------------|----------|-----------|-----------|
| 1 | Proceso | Tipología | Aguinante |
| 2 | Albúmina | Impresión | Brillo |
| 3 | Albúmina | Impresión | Albúmina |
| 4 | Albúmina | Impresión | Semi mate |
| 5 | Albúmina | Impresión | Semi mate |
| 6 | Albúmina | Impresión | Semi mate |
| 7 | Albúmina | Impresión | Semi mate |
| 8 | Albúmina | Impresión | Semi mate |
| 9 | Albúmina | Impresión | Semi mate |
| 10 | Albúmina | Impresión | Semi mate |
| 11 | Albúmina | Impresión | Semi mate |
| 12 | Albúmina | Impresión | Semi mate |
| 13 | Albúmina | Impresión | Semi mate |
| 14 | Albúmina | Impresión | Semi mate |
| 15 | Albúmina | Impresión | Semi mate |
| 16 | Albúmina | Impresión | Semi mate |
| 17 | Albúmina | Impresión | Semi mate |
| 18 | Albúmina | Impresión | Semi mate |
| 19 | Albúmina | Impresión | Semi mate |
| 20 | Albúmina | Impresión | Semi mate |
| 21 | Albúmina | Impresión | Semi mate |
| 22 | Albúmina | Impresión | Semi mate |
| 23 | Albúmina | Impresión | Semi mate |
| 24 | Albúmina | Impresión | Semi mate |

PREPARACIÓN DE LOS DATOS

De las 27 columnas originales de la hoja de cálculo, se mantienen 15 columnas o atributos que representan conjuntos de información lo más acotados posibles en un rango de valores bien identificado. Las razones principales para descartar algunas de las columnas son la ausencia de información en campos opcionales (por ejemplo, subtipo de soporte), y la dispersión de datos en el caso de las columnas donde se escriben comentarios o anotaciones que suelen ser únicos (por ejemplo, particularidades del objeto fotográfico).

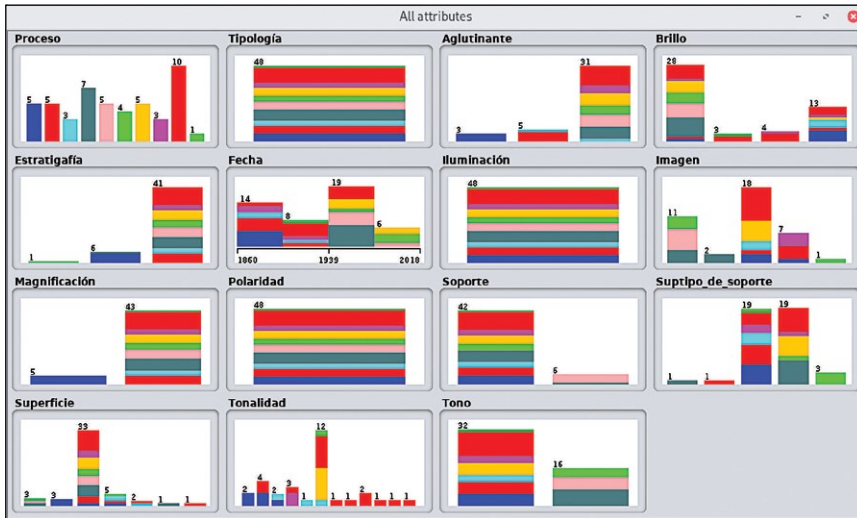
El método de almacenamiento nativo de WEKA es a través de archivos en formato ARFF. Este formato es muy similar a un archivo con formato de valores separado por coma (CSV), con la diferencia de que se le agregan encabezados que dan cuenta de las columnas o atributos, así como de los valores que pueden tener.

Este proceso de conversión de formato es simple de llevar a cabo gracias a que diversos programas permiten exportar bases de datos u hojas de cálculo en el formato CSV. Sin embargo, esta tarea se dificulta si los datos no están estandarizados, contienen errores ortográficos u otras inconsistencias. Por tal razón, la importancia de definir un vocabulario común es vital para un correcto funcionamiento del algoritmo y a su vez, para mantener consistencia e integridad de los datos.

EJECUCIÓN DEL ALGORITMO

Para utilizar el algoritmo C4.5 en WEKA, se hace uso de la versión de código abierto llamada algoritmo J4.8 (versión ligeramente mejorada del algoritmo C4.5 revisión 8 que está implementando en el lenguaje de programación Java). Con esto en consideración, el procedimiento es bastante directo, el primer paso consiste en seleccionar el archivo ARFF desde el explorador de WEKA para apreciar el conjunto de datos en su totalidad.

Figura 7. Gráficas de todas las propiedades a utilizar. Cada color representa un proceso fotográfico diferente.



En seguida se selecciona la pestaña de clasificación y se selecciona de la lista de opciones el algoritmo J48, se asignan los parámetros correspondientes, se selecciona el atributo o columna que representa las clases (en este caso el proceso fotográfico), se ejecuta el algoritmo y se observan los resultados que se obtienen.

Es crucial determinar los parámetros adecuados para obtener el mejor árbol de decisión. Si bien la cantidad de parámetros es acotada, son suficientes para generar diversidad de resultados que solamente pueden ser considerados mejores o peores, en comparación con el porcentaje de registros clasificados correctamente, según un conjunto de prueba (el cual puede ser el mismo conjunto de datos).

Los parámetros más significativos empleados para lograr un árbol de decisión suficientemente compacto, pero que logre clasificar correctamente la mayor cantidad de elementos no siempre es evidente y requiere de experimentación y múltiples ejecuciones del algoritmo.

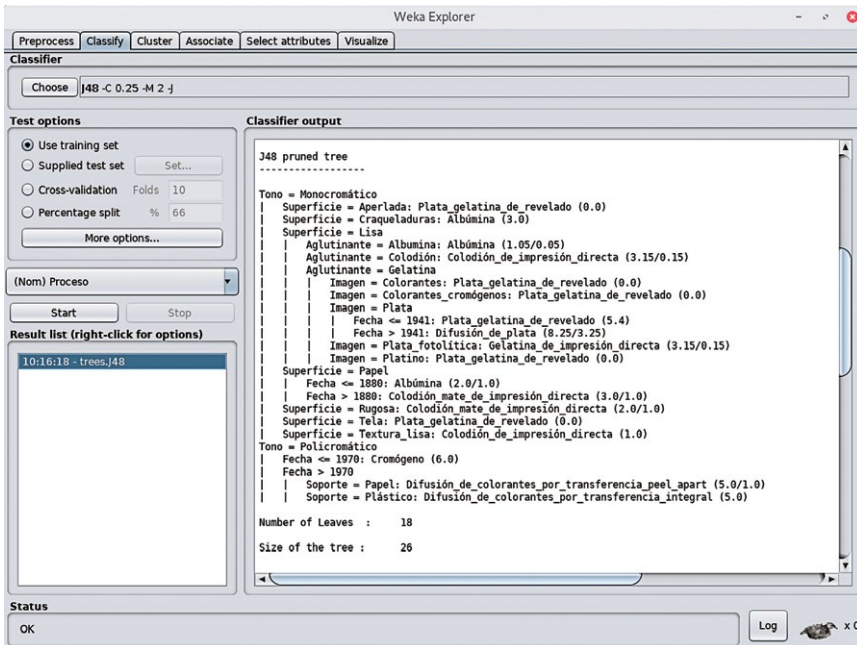
Tabla 4. Descripción de los parámetros relevantes en el algoritmo J4.8

| Parámetro | Descripción | Valor usado |
|---------------------------------------------|---------------------------------------------------------------------------------------------------|-------------|
| Unpruned tree (unpruned) | Usar árbol de decisión sin poda o recortes (genera árboles más grandes). | False |
| Reduced-error pruning (reducedErrorPruning) | Determina si se utiliza un método alternativo de poda (no necesariamente con mejores resultados). | False |
| Pruning confidence (confidenceFactor) | Factor que se utiliza para podar o recortar el árbol; a menor valor mayor poda. | 0.25 |
| Minimum number of instances (minNumObj) | Establece el mínimo número de instancias por hoja o el número mínimo de ramificaciones. | 2 |
| MDL-correction (useMDLcorrection) | Ajuste basado en MDL (mínima longitud de descripción) para calcular ganancia de información. | False |

RESULTADOS

El resultado es un árbol de 26 nodos, 25 ramas y 18 son nodos terminales u hojas, que permiten clasificar fotografías según los 15 atributos que se definieron en la base de datos. Se tiene un porcentaje de instancias clasificadas correctamente de 85.41%, que es el porcentaje más alto que se alcanzó con los parámetros que ya se mencionaron. Esto no representa un árbol ideal que clasifique correctamente el 100% de las muestras pero da cuenta de una estructura diferente a la concebida de manera no-automatizada y hace evidente, al recorrer el árbol desde la raíz hasta una hoja, que hay características de la fotografía que tienen mayor peso para lograr una rápida clasificación de la misma.

Figura 8. Resultados mostrados en WEKA. Se aprecia una versión textual del árbol de decisión generado.

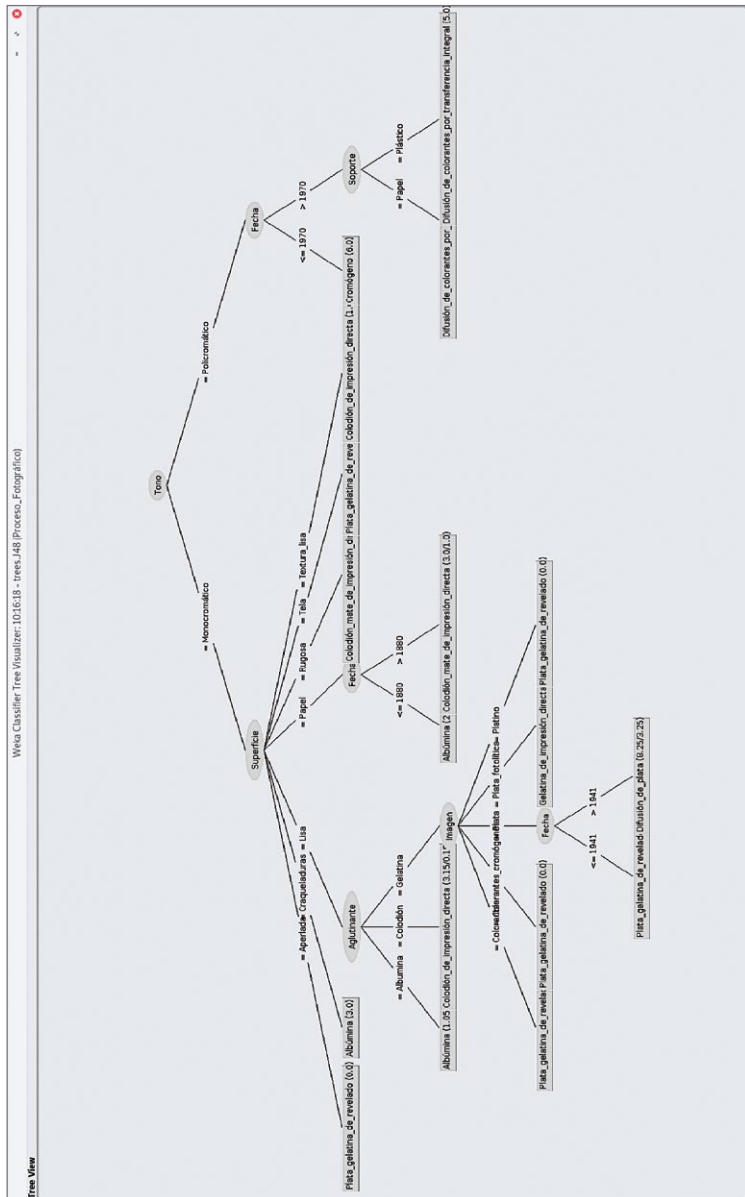


CONCLUSIONES

Los conocimientos, experiencia y habilidades de especialistas, combinados con técnicas computacionales, nos permitieron obtener una metodología más robusta que cualquiera de estas dos opciones por separado. Se evidenciaron las diferencias entre las posibles maneras de construir un árbol de decisión y se revaloró la importancia de algunas características físicas de la fotografía. Una de las finalidades, buscar un balance entre un árbol compacto autogenerado y uno creado, a partir de la experiencia profesional, a fin de enriquecer también la práctica docente y de investigación.

A mayor cantidad de datos disponibles, el árbol de decisiones producido de manera automática, genera un modelo más refinado para contemplar diversos casos y mejorar la clasificación. Pero se debe

Figura 9. Representación gráfica del árbol de decisión final generado en WEKA.



tener cuidado con el problema del sobreajuste de datos, para lo cual se recomienda que los datos sean suficientemente representativos de las características físicas en las fotografías. Por lo tanto, no se debe descartar la opción de eliminar o agregar atributos y columnas en futuros procedimientos, ya que la experimentación y la comparación de diferentes versiones de árboles de decisión es de suma importancia en la labor comparativa.

Futuros trabajos encaminados en la misma temática pueden llegar a considerar una interfaz web a la que se le suministre el resultado de la generación automática de árboles de decisión, para mejorar la experiencia de usuario, e identificar de manera más rápida el proceso fotográfico, sin dejar de lado la parte didáctica con las explicaciones que ya se tienen. Así también, incluir otros tipos de análisis computacionales, como el uso de otras técnicas de Machine Learning o de análisis digital de la imagen asistida con herramientas de inteligencia artificial.

BIBLIOGRAFÍA

Aasbø, Kristin, y Edwin Klijn. (2003). *Sepiades: recommendations for cataloguing photographic collections: advisory report by the SEPIA Working Group on Descriptive Models for Photographic Collections*. Amsterdam: European Commission on Preservation and Access.

Barra Moulain, Paula Alicia e Ignacio Gutiérrez Rubalcava. (2005). *Normas catalográficas del Sistema Nacional de Fototecas del INAH*. México: Instituto Nacional de Antropología e Historia.

Bell, Jason. (2015). *Machine Learning: Hands-On for Developers and Technical Professionals*. [S.l.]: John Wiley.

Boadas i Raset, Joan, Lluís-Esteve Casellas y M. Àngels Suquet i Fontana. (2001). *Manual para la gestión de fondos y colecciones fotográficas*. Girona: CCG Ediciones.

Comité Técnico de Normalización Nacional de Documentación. (2016). Norma Mexicana NMX-R-069-SCFI-2016. Documentos fotográficos. Lineamientos para su Catalogación. México: Secretaría de Economía.

Witten, Ian H., y Eibe Frank. (2016). *Data Mining: Practical Machine Learning Tools and Techniques, Fourth Edition*. San Diego: Elsevier Science & Technology Books.

Géron, Aurélien. (2017). *Hands-On Machine Learning with Scikit-Learn and TensorFlow. Concepts, Tools, and Techniques to Build Intelligent Systems*. 1st ed. Sebastopol, CA: O'Reilly.

_____. (2019). *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*.

Heredia Herrera, Antonia. (1995). *Archivística general: teoría y práctica*. Sevilla: Diputación Provincial.

Image Permanence Institute. Graphics Atlas: Search Process. <http://www.graphicsatlas.org/identification/>.

Quinlan, Ross. (1986). *Induction of Decision Trees*. Machine Learning 1. 81-106. doi: <https://doi.org/10.1007/BF00116251>.

Reilly, James. (1986). *Care and identification of 19th-century photographic prints*. Rochester: Eastman Kodak Company.

Inteligencia artificial y datos masivos en archivos digitales sonoros y audiovisuales.

Instituto de Investigaciones Bibliotecológicas y de la Información/UNAM. La edición consta de 100 ejemplares. Coordinación editorial, Israel Chávez Reséndiz; revisión especializada, Angélica Valenzuela y Valeria Guzmán González; revisión de pruebas, Valeria Guzmán González; formación editorial, Oscar Fernando Arcos Casañas. Fue impreso en papel cultural de 90 gr. en los talleres de Grupo Fogra. Año de Juárez 223. Col. Granjas San Antonio. Alcaldía Iztapalapa. Ciudad de México. Se terminó de imprimir en 2020.

Desde la invención de la tecnología de la información, nuestra capacidad de generación de datos nunca había sido tan intensa y constante. Cada año se produce más información digital que el anterior. El uso y manejo de grandes volúmenes de información digital es un signo contemporáneo. La Inteligencia Artificial tiene una amplia gama de aplicaciones para enormes cantidades de datos, por lo que ambos conceptos van de la mano en pro de su almacenamiento y catalogación.

En los primeros años del siglo XXI, se han puesto en marcha las primeras aproximaciones teóricas y desarrollos tecnológicos de la IA en grandes volúmenes de datos digitales que se resguardan en archivos sonoros y audiovisuales. Los resultados de las experiencias pioneras son muy valiosos.

Este libro reúne estas experiencias, así como opiniones y aproximaciones al uso de Inteligencia Artificial para resguardar datos masivos en distintos ámbitos profesionales y académicos.



TECNOLOGÍAS
DE LA INFORMACIÓN

