

Inteligencia artificial y datos masivos en archivos digitales sonoros y audiovisuales

Perla Olivia Rodríguez Reséndiz
Coordinadora



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:

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: En trámite

Publicación dictaminada

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 aaa 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 Último

1000

Figura 2. Primer sistema de catalogación audiovisual Winisis.

Adición global de campo

Rango MFN (de/a) 1/32000

Etiqueta de campo

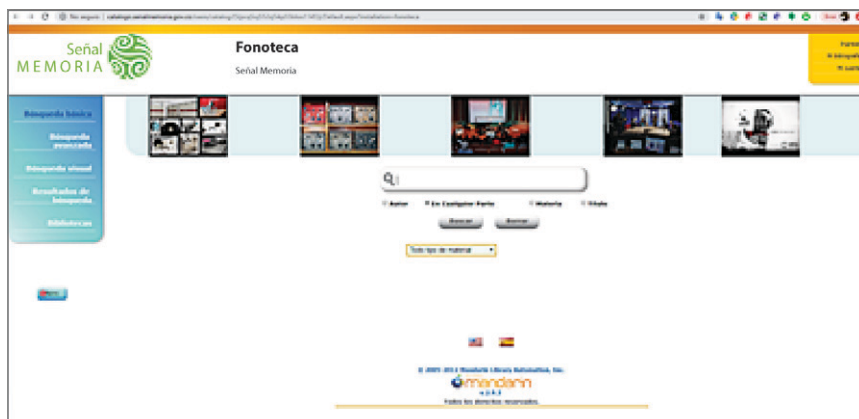
☐ Add if not present

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.

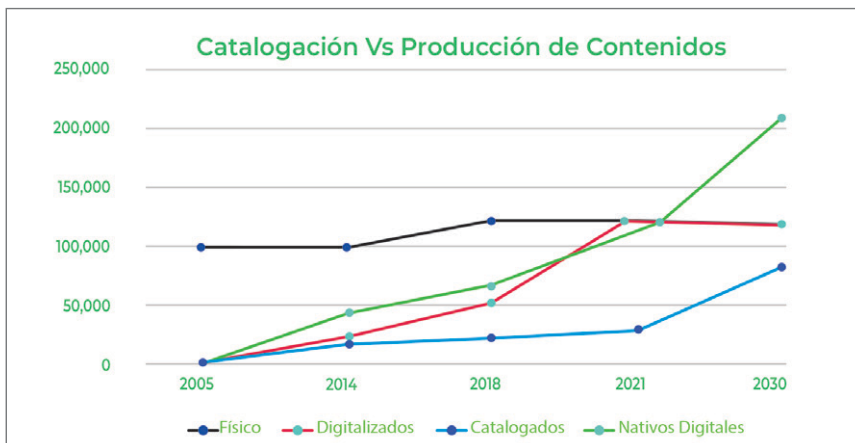


Figura 5. Nuevo sistema de catalogación KOHA, versión usuario (web).



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
	/650/653		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		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 marcador 70%
Enseñanza LABRIA	500	dc:seventyrange	Rango de tiempo 70%
		dc:initialthirty	Tiempo inicial marcador 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-SER001CPT06A	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-SER001CPT06A	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-SER001CPT06A	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-SER001CPT06A	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-SER001CPT06A	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-SER001CPT06A	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-SER001CPT06W	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-SER001CPT06W	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-SER001CPT06W	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-SER001CPT06W	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-SER001CPT06W	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.

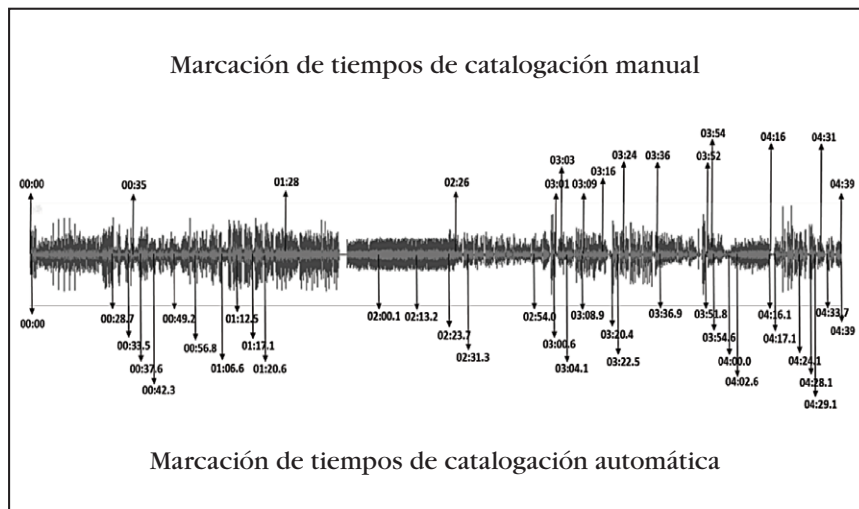
```
<format>.wav</format>
<duration>279.943</duration>
<frequency>48000</frequency>
<totalSamples>13437326</totalSamples>
<sampleRate>48000</sampleRate>
<nchannel>2</nchannel>
+<codec>
</media>
</content>
<body>

<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.

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.