

Aplicaciones de inteligencia artificial en archivos media

Artificial intelligence applications in media archives

Virginia Bazán-Gil

Note: This article can be read in its English original version on:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87392>

Cómo citar este artículo.

Este artículo es una traducción. Por favor cite el original inglés:

Bazán-Gil, Virginia (2023). "Artificial intelligence applications in media archives". *Profesional de la información*, v. 32, n. 5, e320517.

<https://doi.org/10.3145/epi.2023.sep.17>

Artículo recibido el 07-06-2023
Aceptación definitiva: 31-08-2023



Virginia Bazán-Gil

<https://orcid.org/0000-0003-4920-2212>

Radiotelevisión Española. Fondo Documental RTVE

Universidad Carlos III de Madrid

Facultad de Humanidades, Comunicación y Documentación

Departamento de Biblioteconomía y Documentación

Madrid, 128. 28903 Getafe (Madrid), Spain

virginia.bazan@rtve.es

Resumen

Este trabajo presenta una panorámica internacional del uso de la Inteligencia artificial en el contexto de los archivos media de las empresas de televisión, organizaciones preservadoras del patrimonio audiovisual y agencias de comunicación. Para ello, se ha realizado una revisión exhaustiva de fuentes especializadas, esencialmente estudios de caso presentados en conferencias internacionales y seminarios, así como los resultados de la encuesta sobre el uso de la Inteligencia artificial en archivos media organizada por la *Federación Internacional de Archivos de Televisión (FIAT/IFTA)*. Una vez definidas las tecnologías más utilizadas e identificadas las fases de la cadena de producción en las que se emplean, se examinan las aplicaciones concretas de estas tecnologías en los archivos de televisión, organizaciones preservadoras del patrimonio audiovisual, agencias de prensa y proyectos de innovación que reúnen a empresas tecnológicas y medios. También se discuten desafíos relacionados con la implementación de la IA en los archivos, la necesidad de conjuntos de datos para el desarrollo de modelos de lenguaje y la importancia de un uso responsable de la tecnología.

Palabras clave

Inteligencia artificial; IA; Archivos media; Archivos de televisión; Archivos audiovisuales; Documentación audiovisual; Tecnologías del habla; Procesamiento del lenguaje natural; Metadatos; Media; Radio; Televisión; *FIAT/IFTA*.

Abstract

The aim of this paper is to present an international overview of the use of artificial intelligence in the context of media archives in broadcasters, preservation institutions and press agencies, through a comprehensive analysis of sources primarily focusing on case studies presented at international conferences and seminars, together with the results of the survey on the use of artificial intelligence conducted by *FIAT/IFTA*. Once the most commonly used technologies have been defined and we have identified the stages of the production workflow in which they are used, we will discuss the specific applications of these technologies in television archives, audiovisual heritage preservation organisations, press agencies and innovation projects where technology vendors and media companies collaborate. Finally, we will deal with the challenges related to the implementation of AI in media archives, the need for datasets in the development of language models, and the relevance of a sensible use of technology.

Keywords

Artificial intelligence; AI; Media archives; Television archives; Audiovisual archives; Audiovisual documentation; Speech technologies; Natural language processing; Metadata; Media; Radio; Television; *FIAT/IFTA*.



1. Introducción

En los últimos años la inteligencia artificial (IA) ha irrumpido en muchos sectores de la sociedad, llegando incluso a ser objeto de debate público dada la popularidad alcanzada por sistemas de IA generativa como *ChatGPT* (Open AI, 2022) o *Dall-E* (Open AI, 2023). Los medios de comunicación no han permanecido ajenos y han adoptado estas tecnologías en todas las fases del proceso periodístico, desde la recogida de información, a la producción automatizada de contenidos y la distribución y relación con la audiencia (Sánchez-García *et al.*, 2023).

En el ámbito de la producción de contenidos, la IA permite el análisis masivo de datos para comprender hechos informativos relevantes en los que se generan grandes volúmenes de información, como puedan ser los procesos electorales o acontecimientos deportivos como los Juegos Olímpicos. Aunque muy lentamente, la generación automática de noticias empieza a abrirse camino en las redacciones, dibujando un futuro en el que los profesionales darán el valor añadido a la información generada por algoritmos a partir de datos masivos estructurados. También en el ámbito de la producción, la IA se ha convertido en un elemento esencial para verificar la información, incluyendo la detección de imágenes, audios o videos falsos generados mediante algoritmos.

En cuanto a la distribución de los contenidos, la IA permite la creación de alertas y recomendaciones de acuerdo con los perfiles de los usuarios, la generación de subtítulos y de traducciones automáticas. Son elementos esenciales en los medios públicos para garantizar la accesibilidad a la información de distintos colectivos, como puedan ser las personas sordas. Los subtítulos y los propios textos informativos, procesados de forma automática, facilitan la detección de protagonistas y palabras clave que se usan como etiquetas y que hacen posible el descubrimiento de contenidos, mejoran su posicionamiento y atraen a un mayor número de usuarios.

2. Objetivos y metodología

El objetivo de este trabajo es presentar una panorámica internacional del uso de la IA en el contexto de los archivos media de las empresas de televisión, organizaciones preservadoras del patrimonio audiovisual y agencias de prensa. Además, se presentan proyectos de innovación que reúnen a empresas tecnológicas con medios de comunicación para profundizar en la aplicación de estas mismas tecnologías. Para ello, se ha realizado un análisis exhaustivo de fuentes, esencialmente estudios de caso presentados en conferencias internacionales como las de la *Federación Internacional de Archivos de Televisión (FIAT/IFTA)* o seminarios y sesiones de los grupos trabajo de la *Unión Europea de Radiodifusión (EBU/UER)*, desde 2013 hasta la actualidad. Se trata de fuentes de información que están únicamente al alcance de los profesionales del sector ya que, en casos como *EBU/UER*, los seminarios y la documentación generada por los distintos grupos son sólo accesibles a sus miembros, mientras que *FIAT/IFTA* suele mantener un embargo de un año sobre los contenidos de sus conferencias anuales y/o seminarios. Esta visión se completa con los resultados de una encuesta sobre el uso de la IA promovida por *FIAT/IFTA* en abril de 2023. El objetivo de esta encuesta, elaborada por la autora y la *Media Management Commission (MMC)*, es conocer el uso que los archivos de los medios de comunicación están haciendo de la Inteligencia artificial. Los resultados iniciales se ha presentaron en el seminario de la *MMC* de *FIAT/IFTA* en mayo de 2023.

3. Una panorámica de la situación actual

De acuerdo con el estudio realizado en 2022 por *Radiotelevisione Italiana (RAI)* en el contexto del proyecto *AI4Media*, el uso de la Inteligencia artificial en los medios de comunicación ha dejado de ser una tendencia novedosa y, sin embargo, dista mucho de ser una práctica industrial establecida (Bruccoleri *et al.*, 2022; *AI4Media*, 2023). En este sentido, todavía no es posible determinar cuándo se va a disponer de funcionalidades plenamente operativas y de alto nivel cualitativo. Es más, algunas tareas que se consideran esenciales no están bien cubiertas por las actuales aplicaciones, siendo necesario el desarrollo de unas nuevas. Según los datos aportados por este informe, las tecnologías tienen un enorme potencial para apoyar la cadena de valor de los medios de comunicación y podrían dar un gran impulso en términos de calidad y creatividad al trabajo de los medios, si bien no deberían reemplazar el trabajo humano. Finalmente, la fiabilidad de la IA es un factor esencial que condiciona su aplicación en los medios de comunicación, por lo que se considera fundamental la elaboración de aplicaciones fiables que respeten la privacidad de los usuarios y cumplan la normativa de protección de datos.

En el ámbito de los archivos media, los seminarios bianuales de la *Media Management Commission (MMC)* de *FIAT/IFTA* se han considerado como un marco de referencia esencial para intercambiar ideas y conocimiento sobre gestión de archivos, media, metadatos y avances tecnológicos (Green; Gupta 2019).

- En 2007, en Viena (Austria) se presentan los primeros avances tecnológicos en catalogación automatizada, detección de escenas, reconocimiento de caras y objetos, reconocimiento automático de caracteres y generación automática de subtítulos.
- En 2013, en Hilversum (Holanda), IA y anotación automática son objeto de debate y reflexión y los archivos media empiezan a preguntarse por el futuro papel de los catalogadores, un debate que continúa vigente en la actualidad (*FIAT/IFTA Media Management Commission*, 2013).
- En 2017, en Lugano (Suiza), y de la mano de la segunda generación de *MAM (Media Asset Manager)*, se presentan las primeras pruebas piloto realizadas en archivos de televisión y, de nuevo, se discute el futuro rol de los documentalistas, esta vez como parte implicada en la generación de modelos o la supervisión de los resultados (*FIAT/IFTA Media Management Commission*, 2017).

- En 2019, en Estocolmo (Suecia), se conmemoran los 20 años de los seminarios de la MMC y, una vez más, se pone de manifiesto la capacidad de adaptación de una profesión que continúa haciéndose preguntas similares en un contexto de cambio tecnológico constante (FIAT/IFTA Media Management Commission, 2019).
- En 2023, en Dublín (Irlanda), con numerosos archivos usando soluciones de IA en producción, se debate sobre la ya incuestionable necesidad de convivencia entre los algoritmos y las personas (FIAT/IFTA Media Management Commission, 2023).

En las redacciones, la IA facilita la recuperación de forma inmediata de contenidos para la producción. En el archivo hace accesibles colecciones con niveles de catalogación muy bajos

En la actualidad, en los archivos audiovisuales de los medios de comunicación, la IA se está aplicando tanto en la producción como en el archivo. En el archivo de producción se usa para el análisis de señales y material original con el fin de facilitar la recuperación y uso de las imágenes de forma inmediata y, muy especialmente, en las redacciones de informativos. En el archivo las soluciones se basan en la recuperación de colecciones cuyo nivel de descripción no es suficiente para garantizar su reutilización.

Las distintas soluciones varían según la organización y pueden responder a proyectos de innovación, pruebas de concepto o a proyectos bien delimitados en cuanto a su alcance temporal. Es interesante señalar cómo los proyectos de IA en los archivos se llevan a cabo por equipos interdisciplinarios en los que participan distintas áreas de las organizaciones. Así mismo, el material de archivo tiene un valor esencial en la elaboración de soluciones que se aplican en los sitios web o en las plataformas de vídeo bajo demanda de las empresas de televisión sin que a veces exista un retorno en el propio archivo, como veremos en algunos de los casos de uso que presentamos a continuación.

4. ¿De qué tecnologías hablamos?

Pero, antes de profundizar en las aplicaciones concretas, es importante definir de qué tecnologías hablamos cuando se trata de IA aplicada a archivos media. Vamos a detenernos algo más en ellas:

- Tecnologías del habla y del audio: se trata de un conjunto de tecnologías que permiten el reconocimiento automático del habla y su transcripción a texto, incluyendo el reconocimiento del idioma y la identificación de locutores, así como la detección de algunas características asociadas a éstos, como son el género, la edad o el estado anímico. Estas tecnologías permiten también conocer el entorno acústico en el que se habla, la detección de locución, música y silencio.
- Procesamiento del lenguaje natural: es el conjunto de técnicas que permiten comprender la estructura y significado de los textos. Mediante su aplicación es posible la detección de entidades nombradas en los textos, la identificación de palabras clave o la clasificación automática de los contenidos. Estas tecnologías se usan también para la generación de textos, resúmenes, etc. mediante IA generativa.
- Visión artificial: se trata de la rama de la IA que permite a los sistemas extraer información significativa a partir de imágenes digitales o videos. Su uso en los archivos de los medios de comunicación se aplica esencialmente al reconocimiento facial y de identidad, reconocimiento de logos y objetos, reconocimiento de subtítulos (*optical character recognition*, OCR), segmentación por escenas y planos, resúmenes en imágenes y generación automática de contenidos.

Estas tecnologías se están implementando tanto en los medios de comunicación públicos y privados, como en instituciones preservadoras del patrimonio audiovisual y en agencias de noticias. En los últimos años, e impulsados por los fondos de resiliencia europeos, son numerosos los proyectos de innovación que cuentan entre sus participantes con cadenas de televisión que aportan datos, conocimiento y experiencia a estos proyectos. Se produce así una sinergia con la industria que busca comprender mejor el mercado y crear productos para un sector que está experimentando una gran transformación y que está definiendo cómo será su futuro.

5. Tres actores fundamentales: Instituciones preservadoras, cadenas de televisión y agencias de noticias

Existen tres agentes fundamentales que desarrollan proyectos de IA para la gestión, preservación y explotación de las colecciones audiovisuales. Nos detenemos a continuación en cada uno de ellos.

5.1. Instituciones preservadoras

En Europa existen dos organizaciones que son un claro referente en lo que a preservación del patrimonio audiovisual se refiere: el *Institut national de l'audiovisuel* (INA) en Francia y el *Netherlands Institute for Sound & Vision* (NISV) en Holanda.

5.1.1. Institut national de l'audiovisuel (INA)

El *Institut national de l'audiovisuel*, más conocido como INA, es una institución pública creada en 1975 con el objetivo de preservar el patrimonio audiovisual francés, crear contenidos, investigar y transferir conocimientos en los ámbitos audiovisual y digital (INA, 2023a). El INA es responsable del depósito legal audiovisual y de la web en Francia y, como parte de su misión, debe promover y facilitar el acceso a las colecciones que preserva, tanto a profesionales del sector como a investigadores o al público en general. Las colecciones de INA incluyen la producción de 179 canales de televisión y emisoras de radio además de sitios web y cuentas en redes sociales (INA, 2023b). Desde un punto de vista práctico esto

supone la gestión de grandes cantidades de metadatos heterogéneos y de muy diversa procedencia, a veces inexactos o en ocasiones simplemente inexistentes. En este contexto, la IA se presenta como un instrumento adecuado para mejorar la descripción y asegurar la accesibilidad de estos contenidos. En este sentido podemos destacar varios proyectos de innovación llevados a cabo en los últimos años como son *NOA* y *Trombinos*.

El proyecto *NOA* supone la aplicación de técnicas de visión artificial y procesamiento del lenguaje natural para la segmentación de la emisión de una cadena de televisión en programas, la segmentación de estos programas informativos en noticias y la identificación de los temas tratados mediante el análisis de los rótulos (**Martin; Segura, 2021; Couteux; Segura, 2023**). La visión artificial se aplica para el reconocimiento de cabeceras, logos, presentadores y títulos de crédito, de forma que es posible identificar el comienzo y el final de cada programa y asociarlo a una cabecera. La presencia y ausencia del presentador en pantalla facilita la segmentación de los programas en sus distintas noticias y el análisis de los subtítulos permiten identificar los principales temas tratados en las noticias. Toda la información generada de manera automática se valida posteriormente por un profesional de forma manual.

Trombinos es un proyecto de reconocimiento facial basado en un modelo de *IBM* creado por el *INA*. El algoritmo ha sido entrenado con 62 millones de caras correspondientes a 70.000 personas. Estas imágenes se han obtenido tanto de programas de televisión como de imágenes recuperadas mediante motores de búsqueda en internet. Los contenidos son procesados usando el modelo y los resultados se devuelven con un nivel de confianza determinado y son validados manual o automáticamente. Cada una de las personas reconocidas se asocia a un registro de autoridad que incluye enlaces a fuentes de datos externas como *DBpedia*. En la actualidad el *INA* trabaja en un modelo con menos sesgos en el que exista una mayor representatividad de caras en cuanto a género y raza (**Petit, 2022**).

Por otra parte, las soluciones basadas en IA han tenido un papel relevante en el desarrollo de *data.ina.fr*, un portal que pretende fomentar el conocimiento de las colecciones del *INA* mediante la analítica de datos (**Roche-Dioré, 2023**). En este portal, *Trombinos* e *INA Speech Segmenter* aportan el reconocimiento facial y la segmentación del audio respectivamente. *Vocapia* realiza la transcripción de audio a texto y *Textrazor* el reconocimiento de entidades. El proceso de análisis de los contenidos de forma masiva a través de la plataforma de IA ha permitido generar grandes cantidades de datos que pueden facilitar estudios posteriores relacionados con los medios de comunicación.

5.1.2. The Netherlands Institute for Sound and Vision (NISV)

El *Netherlands Institute for Sound and Vision*, es la institución responsable de mantener vivo el patrimonio de los medios públicos de comunicación en Holanda, garantizar su preservación y ponerlo a disposición de la sociedad en su conjunto (*Netherlands Institute for Sound and Vision, 2023*). El *NISV* apostó de forma temprana por la generación automática de metadatos y la integración con fuentes de datos externas para facilitar la explotación de sus colecciones. En 2012 comenzó a utilizar técnicas de transcripción automática de audio a texto (*S2T*), aunque con una alta incidencia de errores. En paralelo aplicaron soluciones para el reconocimiento de entidades nombradas (*NERD*), práctica que abandonaron en 2019 al considerar que no respondía a sus necesidades en un entorno real de producción. Este mismo año empezaron a desarrollar el reconocimiento de voces y el reconocimiento facial para facilitar el etiquetado de sus contenidos (**Manders, 2019**).

El actual modelo de reconocimiento facial en *NISV* se basa en el tesoro de personalidades en el que únicamente se recogen personajes públicos de acuerdo a una política de privacidad bien definida respetuosa con la *Reglamento General de Protección de Datos (RGPD)* (**Manders, 2022**). El equipo del proyecto definió desde el principio los niveles de confianza esperados, de forma que sólo las caras identificadas con un nivel de confianza superior al 90% son ingestadas en el sistema. En la actualidad el sistema ha alcanzado un 95% de confianza, reconoce a 3 de cada 4 personas y es capaz de etiquetar hasta un 50% de las caras que aparecen en un programa de televisión. A pesar de los buenos resultados preocupa su escalabilidad, sus limitaciones relacionadas con el uso del tesoro como elemento esencial con los sesgos que esto implica, la dificultad para identificar personalidades emergentes y la falta de modelos faciales en el conjunto de datos, lo que puede suponer falsas identificaciones.

A pesar de haber adoptado todas estas tecnologías y de contar con un área de innovación y desarrollo potente, en 2022 la mayor parte de los metadatos que *NISV* gestionaba no se habían producido de manera automática (**Manders, 2022; Manders; Wigham, 2021**).

5.2. Las cadenas de televisión

Las cadenas de televisión se enfrentan a diario a problemas relacionados con la inmediatez de la producción y con la necesidad de hacer accesible un gran volumen de contenidos de muy diversa procedencia que se ingesta a diario en los sistemas de producción. En el caso de las televisiones públicas europeas, con archivos que se remontan a la década de los 60 del siglo XX, con colecciones de origen diverso y con políticas cambiantes de catalogación, la IA se presenta como una oportunidad para mejorar la eficiencia, aumentar la reutilización de los fondos de archivo y evitar la repetición de tareas. Veamos a continuación algunos ejemplos.

5.2.1. YLE

Yleisradio Oy (YLE), la cadena pública finlandesa, ha sido pionera en el testeo de soluciones basadas en IA (**Selkälä, 2017**). La aplicación de técnicas automáticas para la generación de metadatos se planteó como una forma de mejorar la accesibilidad

y reutilización de unas colecciones completamente digitalizadas y de las que, en ocasiones, no había suficientes datos. Para explorar las posibilidades de la IA YLE formó un equipo de trabajo multidisciplinar en el que el Archivo se integró con miembros de la redacción, operaciones y multimedia. Este equipo exploró las posibilidades del reconocimiento de imágenes, segmentación por escenas, reconocimiento de objetos y caras y reconocimiento óptico de caracteres (OCR) sobre programas de actualidad, por ser los más utilizados. Este proyecto piloto, realizado en 2016, no demostró de forma efectiva una reducción del tiempo dedicado al análisis y demostró algunas limitaciones como la imposibilidad de reconocer la identidad de las personas que aparecían en imagen por falta de tiempo para el entrenamiento del algoritmo.

Ese mismo año realizaron pruebas de transcripción automática de audio a texto y su clasificación para el archivo de radio. Los resultados fueron buenos tanto en transcripción como en reconocimiento de entidades nombradas, si bien no profundizaron en la segmentación por hablantes.

Este piloto, lejos de ofrecer soluciones tecnológicas reales, permitió al equipo del archivo de YLE reflexionar sobre sus necesidades, especialmente sobre qué parte del trabajo puede automatizarse y qué parte debe abordarse manualmente. En la actualidad los esfuerzos de YLE se centran en la generación automática de metadatos para la producción y la publicación de contenidos en la web, la creación automática de textos a partir de datos y el subtítulo automático. El objetivo final es que todos los contenidos producidos por la cadena sean analizados automáticamente y es en YLE Areena, la plataforma de video bajo demanda, donde se está realizando un mayor esfuerzo (Viljanen, 2022).

5.2.2. ARD

El consorcio de radio televisiones públicas de Alemania, ARD, creó en 2017 un grupo de trabajo con el objetivo de identificar oportunidades y crear casos de uso para la introducción de herramientas y métodos de IA en la producción diaria. De acuerdo con los planteamientos de este grupo, el carácter regional y la variedad de dominios son elementos determinantes en la implementación de estas soluciones y al mismo tiempo, una herramienta esencial para generar metadatos adecuados para nuevas plataformas, distintos usuarios, personalizar o recomendar contenidos (Wenger-Glemser, 2019).

Entre las cadenas que forman parte de ARD, podemos destacar el uso de la IA en la televisión de Baviera, Bayerischer Rundfunk (BR). BR ha desarrollado un modelo de reconocimiento facial entrenado con datos y casos de uso propios (Schreiber, 2022). En una primera fase del proyecto se preparó el conjunto de datos para el entrenamiento usando para ello escenas con caras en las que aparecían rótulos identificativos extraídos de dos programas informativos de 30 y 15 minutos de duración. El resultado de este proceso fue un modelo de datos exigente que requería de tasas de acierto muy altas para incorporar los metadatos extraídos al archivo. En febrero de 2022 BR había analizado 55.000 imágenes del archivo llegando a detectar 3.000 clases diferentes. El control de calidad humano se consideró un elemento esencial para detectar falsos positivos, así como las imágenes que los generaban.

En esta misma línea, BR ha entrenado un modelo capaz de reconocer edificios históricos y centros políticos y financieros relevantes del ámbito regional (Förster, 2023). El modelo de datos, generado de forma automática a partir de herramientas *open source*, ha utilizado 255 programas subtítulos sobre los que se aplicaron procesos de reconocimiento de entidades y desambiguación. El objetivo es optimizar el modelo e integrarlo con otras soluciones actualmente en uso.

5.2.3. Radio Télévision Suisse (RTS)

Desde 2018 la Radio Télévision Suisse (RTS), ha abordado de manera sistemática la catalogación de su archivo aplicando técnicas de IA. Para ello, ha desarrollado una interfaz basada en tecnologías de código abierto capaz de transcribir audio a texto, realizar reconocimiento facial y clasificar imágenes de forma automática durante el proceso de ingesta de las imágenes en el archivo (Rezzonico, 2020). La clasificación automática se ha aplicado con éxito sobre colecciones de deportes sin catalogar, lo que les ha permitido conocer, al menos, la temática de cada una de las grabaciones. El reconocimiento facial se ha realizado sobre una base de datos de 5.000 personajes públicos suizos teniendo en cuenta factores como la duración del plano y en niveles de confianza altos para evitar los falsos positivos, es decir, confundir a un personaje con otro en función de sus rasgos faciales (Bouchet; Ducret, 2019). De esta manera, sólo las caras reconocidas con un nivel de confianza superior al 85% son integradas en el archivo. Una de las principales características de esta herramienta es que permite no sólo la obtención de metadatos, sino también la recuperación por imágenes o búsqueda visual por caras, escenas, monumentos y edificios. Esta funcionalidad es especialmente relevante en el sector de los archivos de televisión cuando se trata de localizar escenas que han sido embargadas por motivos de derechos o por orden judicial. La interfaz RTS.ai puede obtener la transcripción automática del audio a texto hasta un total de 10.000 horas en 2 años. En la actualidad, RSI trabaja en el reconocimiento de hablantes, la integración de la transcripción de audio a texto con el reconocimiento facial, la valoración estética de fotografías y videos y la descripción de acciones (Sonderegger, 2023).

5.2.4. BBC

La British Broadcasting Corporation (BBC), ha sido una de las cadenas públicas europeas que antes ha reflexionado sobre las consecuencias que la IA tiene en la producción de contenidos. BBC 4.1 surge como un proyecto del área de Investigación y Desarrollo de la cadena pública británica con un doble objetivo:

- comprender cómo la IA puede influir en el futuro de la producción audiovisual;
- trabajar con los creadores de contenidos para preparar a la cadena para el futuro.

Entre el 4 y el 5 de septiembre de 2018 la programación de *BBC Four*, canal especializado en contenidos culturales, fue generada por un algoritmo. Este algoritmo se había entrenado para detectar los contenidos más relevantes para el canal a partir de los programas emitidos por la cadena en el pasado (BBC, 2018a). Para ello analizó las descripciones de los programas y los temas tratados. Sobre un total de 270.000 programas, el algoritmo identificó los 150 que consideró más relevantes para su emisión y fueron los programadores los que, a partir de esta selección, configuraron la parrilla final. De esta experiencia surgió además el proyecto “*Made by machine: when AI met the archive*” (BBC, 2018b), una serie de 4 capítulos creados mediante IA con la intención de mostrar a los espectadores cómo piensan las máquinas. A partir de los 150 programas más relevantes se aplicaron técnicas de reconocimiento de objetos y escenas, procesamiento del lenguaje natural (PLN) para el análisis de subtítulos y de dinamismo. El resultado, en forma de 4 microespacios, fue emitido en el canal en 2018 junto a la presentación de *Hanna Fry*, matemática y divulgadora británica, y declaraciones de archiveros de la BBC.

Otro de los proyectos de uso de IA en la BBC es “*Starfruit tagger*” (Berger; Armstrong, 2022; *Starfruit tagger*, 2023). La BBC conserva muy pocas de las grabaciones originales de los boletines de noticias emitidos durante sus primeros 50 años de existencia de su emisora de radio, sin embargo, conserva los guiones de estos boletines entre 1937 y 1955. Estos guiones se digitalizaron y procesaron aplicando OCR. Mediante el uso de un sistema de etiquetado automático, entrenado previamente con contenidos de deportes y noticias indizadas manualmente por el equipo editorial de la BBC, se extrajeron en una fase posterior nombres de personas, lugares, organizaciones y eventos.

5.2.5. TV2

El canal comercial noruego TV2 comenzó en 2021 a realizar pruebas de transcripción automática de audio a texto y subtitulación sobre programas de actualidad y material original de informativos y programas (Tverberg, 2021). Este proyecto contó con la colaboración de periodistas y personal de operaciones que, de manera subjetiva, evaluaba la calidad de la transcripción generada a partir de *Speechmatics* (2023), *Azure* (Microsoft, 2023) y otros servicios. El rendimiento se medía además de forma objetiva mediante el cálculo de la tasa de error por palabras (conocida por su acrónimo en inglés WER). Ante la pregunta ¿cómo valoras la calidad de los subtítulos generados automáticamente frente a los subtítulos manuales? Un 9,1% de los usuarios los consideraba muy buenos, un 10,3% buenos, un 21% pensaba que la calidad era suficiente, un 39,4% creía que eran de poca calidad y un 21,2% de muy poca calidad. Desde el punto de vista de los programadores, los principales retos para este tipo de servicios son los dialectos, el reconocimiento de entidades, la segmentación por hablantes y la tasa de error por palabras que, aun siendo cada vez menor, puede tener importancia según el contexto en el que se produzca el error.

TV2 ha realizado también pruebas de concepto con *CLIP* (2023), la red neuronal desarrollada por *OpenAI*, sobre 5.000 contenidos del archivo con apenas metadatos descriptivos (Steskal, 2023). Los resultados obtenidos para el reconocimiento de personas y objetos fueron buenos sin necesidad de un entrenamiento específico, si bien detectaron que los *keyframes* devueltos no eran siempre representativos del video completo y que, a pesar del buen rendimiento, el sistema no será capaz de responder a búsquedas complejas de información.

5.2.6. VRT

La radio televisión pública flamenca de Bélgica, VRT, ha apostado por la elaboración de sus propios modelos de IA para la segmentación automática por escenas de los contenidos publicados en su web y el enriquecimiento de los metadatos de archivo. La aplicación de estas soluciones tiene como objetivo mejorar la eficiencia, aumentar la reutilización, favorecer la recomendación y evitar la repetición de tareas por parte de los equipos de edición, multimedia y archivo. El proyecto se ha llevado a cabo en dos fases, una de preprocesamiento y entrenamiento y una segunda fase de aplicación. Durante la primera fase, se identificaron las fuentes de información más relevantes para entrenar al algoritmo de segmentación, esto es, información procedente de la emisión, subtítulo, escaletas, reconocimiento facial y valores RGB (*red, green and blue*). Todos los datos obtenidos sirvieron para realizar una representación vectorial de los mismos y definir las transiciones entre escenas midiendo los valores RGB. Finalmente se definieron pares de escenas similares y no similares. Una vez entrenado el algoritmo, y durante la fase de aplicación, se extrajo la información de los programas y se aplicó el modelo cuyos resultados fueron validados finalmente por humanos (Daniels; Degryse, 2021).

Además, VRT usa IA en producción para el enriquecimiento de metadatos a partir de la aplicación de OCR en el tercio inferior de las imágenes en pantalla donde son capaces de detectar los nombres de las personas en pantalla. La información obtenida mediante OCR es procesada empleando PLN para obtener entidades y palabras clave y, sobre ella se aplican filtros que permiten depurar información irrelevante (Daniels, 2023).

5.2.7. SVT

Sveriges Television, la televisión pública de Suecia, conocida como SVT, ha realizado distintas pruebas de concepto en su archivo de producción *Mark* considerado, además, como una plataforma para desarrollar la innovación dentro de la propia compañía (Åstrand; Ståhl, 2023). Estas pruebas parten de la ausencia de metadatos en un porcentaje importante de contenidos y utilizan técnicas de IA que permiten su recuperación. Concretamente han realizado pruebas de búsqueda de imágenes mediante *CLIP*, reconocimiento facial para políticos locales, OCR para la identificación de los creadores y equipos técnicos de los programas en los títulos de crédito y, finalmente, la transcripción de audio a texto y el reconocimiento de entidades para la identificación de los protagonistas y temas tratados. El objetivo a medio plazo es identificar casos de uso reales y llevarlos a producción, incorporando únicamente metadatos con un nivel de confianza alto.

5.2.8. *Asharq News*

Asharq News (Battrick, 2022; Battrick; Petitpont, 2022) la cadena de noticias árabe multiplataforma fundada en noviembre de 2020, es probablemente uno de los pocos ejemplos en el mundo en el que el uso de IA fue incluido como parte del plan fundacional de la cadena. El proyecto, realizado en 18 meses, supuso la integración de su sistema de producción *AVID Media Central* con la tecnología de *Newsbridge* con el objetivo de generar metadatos para 1.600 horas de emisión mensuales y material original en inglés y árabe. La prueba de concepto previa a la implantación del servicio contó con la implicación de distintos grupos de usuarios. Estos no sólo participaron en entrenamiento de los modelos de reconocimiento facial y transcripción y traducción automática, sino que, además, identificaron los distintos casos de uso y definieron las tasas de acierto necesarias para la integración. La complejidad de gestionar de forma consistente metadatos en inglés y árabe, la dificultad de entrenar un modelo de transcripción automática del habla para un idioma que cuenta con numerosos dialectos como es el árabe, la especialización de la cadena en temas políticos y militares, la necesidad de generar modelos de reconocimiento facial para personalidades del mundo árabe, le valió a *Asharq* el reconocimiento internacional con el *FIAT/IFTA Media Management Award* en 2022 (FIAT/IFTA, 2022).

5.2.9. *IRIB*

En este mismo contexto, cabe destacar el modelo centralizado de datos y el uso intensivo de la IA que realiza la televisión pública iraní, *IRIB*. La dificultad de encontrar datos para entrenar modelos propios ha llevado a esta cadena a generar una factoría de datos en la que se llevan a cabo los proyectos de etiquetado de imágenes y textos y verificación, así como el pre-procesado automático y validación de datos que se aplicarán después tanto en los sistemas de producción como en el archivo (Ghanbari, 2022).

Esta factoría permite a *IRIB* elaborar herramientas propias que alimentan los distintos servicios basados en IA para el archivo, los canales de televisión, la web y las áreas de administración y finanzas que se nutren también de los datos que generan otras áreas. Estas herramientas se aplican para la generación de resúmenes y detección de entidades sobre texto, la anotación de video y la creación de resúmenes en imágenes, la transcripción de audio a texto, la conversión de texto en audio, la detección de género y la identificación de hablantes.

En Europa, *France Télévision* y la *Radio Televisión Pública Italiana (RAI)* han apostado por enfoques similares.

5.2.10. *France Television*

DAIA es el área de *France Télévision* dedicada a la gobernanza de datos y tiene como objetivo asegurar la interoperabilidad y la disponibilidad de esos datos para los distintos departamentos que forman la compañía. Para ello cuenta con una interfaz de conocimiento que analiza los datos, los comparte con los distintos *datasets* y los traduce cuando no existe coherencia entre ellos, es decir, los convierte en datos interoperables. La base de este sistema es una ontología común que hace posible el entendimiento entre estos *datasets*. Adicionalmente, cuando los datos que existen son insuficientes, se aplica IA de código abierto para generarlos a partir de los títulos de crédito o de los propios contenidos (Parmentier, 2021).

5.2.11. *Radiotelevisione Italiana (RAI)*

En el caso de la *RAI* la generación de conjuntos de datos destinadas al aprendizaje automático se entiende como un elemento clave para la integración de la IA en los procesos de la radiotelevisión pública italiana (Messina, 2021). En este sentido, los metadatos generados por el archivo o las áreas de emisiones y producción, por poner sólo algunos ejemplos, tienen que someterse a un proceso de extracción, filtrado y adaptación para ser útiles en el entrenamiento de modelos. *RAI Media Cognitive Service Platform* es la herramienta que permite la ingesta de los datos, la anotación de contenidos aplicando soluciones en la nube o modelos propios, la validación y enriquecimiento de los datos y la creación de colecciones ya depuradas y adecuadas para el desarrollo de los modelos que después se pondrán en producción. Esto supone que la *RAI* es independiente en sus desarrollos basados en IA de terceros y es capaz de generar sus propios modelos de datos y aplicarlos de forma específica a sus necesidades (Messina; Montagnuolo, 2023).

5.2.12. *Radio Televisión Española (RTVE)*

En España, *RTVE* y *Atresmedia* han sido pioneras en la implantación de la IA en sus archivos. En el caso de *RTVE*, las primeras aproximaciones se llevan a cabo con la creación de la *Cátedra RTVE Universidad de Zaragoza* (Cátedra RTVE Universidad de Zaragoza, 2017). En 2018, y como parte de las actividades de esta cátedra, se publica por primera vez la *RTVE Database*, un conjunto de datos anotados procedentes de programas emitidos en televisión que constituyen la base para los *RTVE Al-bayzin Challenges* (Lleida-Solano et al., 2022). Estos retos, que reúnen a grupos de investigación nacionales e internacionales, han permitido poner a prueba los sistemas más avanzados en transcripción automática y reconocimiento multimodal con casos de uso preparados por el archivo de *RTVE*. En el ámbito de la producción en 2021 se adjudica el expediente para el metadatado automático de contenidos del Fondo Documental de *RTVE* (RTVE, 2021), un servicio en la nube que, a través de un integrador tecnológico, permite el metadatado automático de audio y video de 11.000 horas de contenidos en *RTVE*. Este servicio será sustituido en octubre de 2023 por otro en el que se incluyen nuevas funcionalidades como el análisis y traducción automática desde el catalán y en el que tiene más relevancia el análisis de imagen (RTVE, 2023).

Además, *RTVE* trabaja en un proyecto similar para el archivo de *Radio Nacional de España (RNE)*, que permitirá la transcripción, clasificación automática y extracción de entidades para 190 horas de emisión de *Radio 1* y *Radio 5*.

5.2.13. Atresmedia

Atresmedia puso en marcha en 2019 un proyecto de catalogación automática supervisada. A través de este servicio *on-premise* el archivo de *Atresmedia* obtiene la transcripción de 40 horas diarias de contenido entre material original, piezas de noticias y programas completos subtítulos. Los metadatos generados se integran en el MAM donde se corrigen y complementan con catalogación humana. Este proyecto, que aspira a la transformación del perfil profesional de los documentalistas, de procesador a generador de contenidos, fue reconocido, entre otros, con el premio *Excellence in Media Management* de FIAT/IFTA en 2021 (López-de-Quintana, 2021; López-de-Quintana; León-Carpio, 2021)

Otras televisiones autonómicas españolas, como *Aragón TV* (*Aragón Noticias*, 2021) o la *Televisió de Catalunya CCMA*, han llevado a cabo pilotos de catalogación automatizada, siendo esta última la que ha incorporado a su MAM la transcripción automática en catalán.

5.3. Agencias de noticias

5.3.1. Associated Press

En el ámbito de los medios, las agencias de noticias no han permanecido al margen de los avances tecnológicos. *Associated Press* (AP), tras una prueba piloto y un trabajo de 8 meses ha integrado, de la mano de la empresa belga *Limecraft*, IA multimodal para el análisis tanto de señales en directo como de material grabado, lo que en cifras supone unos 700 clips de video mensuales (Coppéjans, 2021). La tecnología multimodal integra, por una parte, visión artificial para la identificación y segmentación de escenas, reconocimiento facial y de identidad, reconocimiento de roles y actitudes y, por otra, tecnologías del habla que facilitan la detección de idioma y transcripción automática. Los resultados se integran en una interfaz y se muestran con el nivel de detalle adecuado para AP (Verwaest, 2022).

5.3.2. Reuters

La agencia *Reuters*, por su parte, anunciaba en 2020 la aplicación de técnicas de IA sobre un millón de clips de su archivo desde el año 1986 a la actualidad (*Reuters Staff*, 2020). Este proyecto, financiado por *Google DNI Fund*, ha permitido la transcripción automática de voz a texto, la traducción a 11 idiomas y el reconocimiento de personajes internacionales, pero además ha permitido a *Reuters* comprender mejor su archivo, las políticas de análisis aplicadas en el pasado y cómo éstas han afectado a la accesibilidad de sus contenidos (*Reuters*, 2023). De esta forma han podido mejorar la capacidad tecnológica interna y determinar qué tipo de contenidos pueden analizarse automáticamente con un nivel de confianza alto y cuáles no.

6. Uniendo fuerzas: proyectos de innovación

Como hemos visto, son muchas las empresas de radio y televisión que han llevado a cabo pilotos con recursos internos o puesto en producción distintas soluciones basadas en IA. Pero, además, en los últimos cinco años, hemos asistido a un auge importante de los proyectos de innovación en los que empresas, institutos de investigación y medios de comunicación se dan la mano para avanzar en la aplicación de la IA a los procesos de trabajo.

6.1. VIVA

El proyecto *VIVA* ha supuesto la creación de una herramienta capaz de implementar métodos de recuperación de vídeo basados en modelos de aprendizaje profundo (Mühling *et al.*, 2022). En él han participado investigadores del *TIB – Leibniz Information Centre for Science and Technology*– y la *University of Marburg* (Alemania) junto a profesionales de *ARD*. El objetivo es facilitar la recuperación por conceptos o personalidades en archivos media y retroalimentar el modelo de aprendizaje profundo a medida que aparecen personalidades emergentes o nuevas necesidades. Se ha probado sobre cuatro casos de uso en el contexto de una colección de vídeos históricos del *Archivo Alemán de Radiodifusión*, un conjunto de unas 34.000 horas de grabaciones de televisión de la antigua República Democrática Alemana.

6.2. Europeana Subtitled

Otro proyecto relevante en este ámbito es *Europeana Subtitled*, iniciativa que ha reunido a un consorcio de 7 cadenas públicas de televisión europeas junto a la *Fondazione Bruno Kessler (FBK)* y *Translated* con el objetivo de elaborar modelos automáticos de transcripción de habla a texto, traducción y subtítulos automática que permitan mejorar la accesibilidad de los contenidos audiovisuales en colecciones como *Europeana* (Lewis; Jarret, 2023). Este proyecto ha permitido la publicación en la biblioteca digital europea de 8.000 vídeos subtítulos al inglés, en torno al tema *Broadcasting Europe*, como una forma de dar a conocer los cambios sociales que se han producido en Europa desde los años 30.

6.3. AI4Media

El proyecto *AI4Media*, al que ya hemos hecho referencia, tiene como objetivo elaborar herramientas innovadoras para que el sector media pueda afrontar los retos actuales (*AI4Media*, 2023). Para ello, ha definido siete casos de uso industriales que incluyen desde las redes sociales y la desinformación hasta el apoyo a la redacción mediante la creación automática de noticias. Entre estos casos de uso cabe destacar la aplicación de técnicas de visión artificial y metadatos automático de material de archivo para dar apoyo a las coberturas informativas de acontecimientos inesperados, en las que la inmediatez y la calidad son elementos diferenciadores. En este proyecto colaboran 9 universidades, 9 centros de investigación y 12 empresas del sector entre organismos preservadores del patrimonio, empresas de radio y televisión y proveedores tecnológicos.

6.4. Tailored Media

Tailored Media es un proyecto de innovación liderado por *Joanneum Research* con la colaboración de *ORF* y *Austrian Mediatheque* (Bailer; Bauer; Rottermann, 2021). Tiene como objetivo generar mediante técnicas de visión artificial y procesamiento del lenguaje natural, metadatos relevantes de forma automática, integrarlos en los actuales procesos de trabajo y elaborar procesos orientados a los usuarios a partir de estos metadatos.

7. Resultados de la encuesta FIAT/IFTA sobre el uso de IA en archivos media

En abril de 2023 *FIAT/IFTA* lanzaba una encuesta sobre el uso de la IA en los archivos media (*FIAT/IFTA*, 2023). La encuesta tiene como objetivo obtener una panorámica sobre el grado de implantación de la IA. De forma más concreta, se pretende conocer:

- qué tecnologías y aplicaciones se están usando, cuál es su grado de implantación;
- cómo se están integrando los metadatos en los sistemas de gestión de los archivos;
- cuál se espera que sea la evolución de la catalogación de contenidos.

En la encuesta han participado 54 organizaciones, esencialmente cadenas de televisión y archivos audiovisuales nacionales o regionales europeos:

- un 61% (33) de las organizaciones que ha respondido está aplicando IA en su archivo;
- un 9% (5) la está usando únicamente en el archivo de producción;
- un 28% (15) tiene previsto usarla en el futuro;
- sólo un 2% (1) no prevé usarla en el futuro.

En base a las respuestas obtenidas, se puede afirmar que las organizaciones están usando o planeando usar una combinación de distintas tecnologías, existiendo una ligera predilección por:

- tecnologías del audio y del habla (37%, 33) en primer lugar;
- visión artificial (34%, 31) en segundo lugar;
- finalmente, por el procesamiento de lenguaje natural (29%, 26).

Si desglosamos los datos por tecnologías en producción o en fase de planificación, la tendencia se mantiene en las organizaciones que están planificando la integración de IA, mientras que hay un mayor uso de la visión artificial (27) que de tecnologías del audio y del habla (25) y que del procesamiento del lenguaje natural (20) en las organizaciones que ya las emplean.

El gráfico 1 muestra el porcentaje de uso de aplicaciones concretas relacionadas con las tecnologías del habla y del audio. El gráfico 2 muestra las aplicaciones concretas relacionadas con el procesamiento del lenguaje natural. El gráfico 3 muestra el porcentaje de uso de aplicaciones relacionadas con la visión artificial.

En un porcentaje muy significativo de los casos, las tecnologías están plenamente integradas en el proceso de producción, mientras que, en menor porcentaje, son proyectos de innovación que se van a implantar próximamente, o son pruebas de concepto para comprender el alcance y limitaciones de la tecnología.

Otro de los factores relevantes que recoge esta encuesta es el origen de las tecnologías que actualmente están aplicando los archivos media:

- 26% (13), utiliza soluciones de código abierto implementadas por la propia organización;
- 26% (13), tecnología de terceros (tipo *Azure*, *Amazon*, *IBM*, etc.) implementada por la propia organización;
- 26% (13), recurren a los servicios de terceros que integran tecnologías propietarias;
- 23% (11), también recurren a terceros que a su vez integran tecnologías de otros.



Gráfico 1. Porcentaje de uso de aplicaciones de tecnologías del habla y del audio. Fuente: *FIAT/IFTA* (2023)

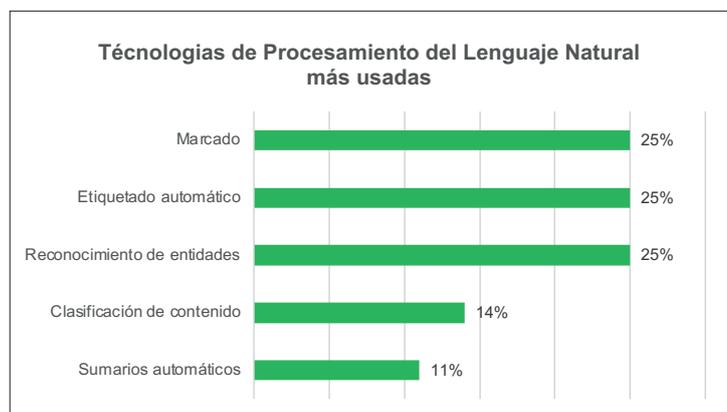


Gráfico 2. Porcentaje de uso de aplicaciones de procesamiento del lenguaje natural. Fuente: *FIAT/IFTA* (2023)

En cuanto a cómo se muestran los datos a los usuarios finales en los sistemas de gestión de archivo:

- 57% (24) no indica la procedencia de los metadatos;
- 12% (5) lo indica con ventanas emergentes;
- 31% (13) lo hace constar en la misma pantalla de visualización de datos.

Finalmente, la mayor parte de las organizaciones que han respondido a la encuesta considera que la catalogación manual va a descender entre un 25% y un 50% en los próximos 5 años.

8. Conclusiones

Desde 2007 los profesionales de los archivos media están reflexionando sobre la aplicación de tecnologías relacionadas con la IA y sobre la evolución futura de su trabajo, especialmente en actividades como la anotación de contenidos. Los primeros proyectos de IA han sido impulsados por las áreas de innovación de las compañías y desarrollados por equipos multidisciplinares. En muchos casos se ha recurrido a *start-ups*, grupos de investigación o compañías especializadas, tal es el caso de *Yle*, *RTVE*, *RSI*. Algunas organizaciones han apostado por programas a medida basados en soluciones *open source*, especialmente a partir de 2019.

La aplicación de la IA tiene como principal objetivo mejorar la eficiencia de los procesos, aumentar la reutilización de los contenidos, favorecer la recomendación en los sitios web y evitar la repetición de tareas por parte de distintos grupos usuarios que, hasta ese momento, generaban la misma información sobre los mismos contenidos en distintos momentos de la cadena de producción (*VRT*, *IRIB*, *RAI*, *France TV*). Las tecnologías del habla y del audio pueden considerarse ampliamente implantadas, aunque existen algunos retos que superar relacionados con la falta de conjuntos de datos para el desarrollo de modelos de lenguaje en idiomas con bajos recursos. La elaboración de modelos basados en datos propios bien estructurados convive con el uso de tecnologías suministradas por terceros, ya sean tecnologías propietarias o comerciales adaptadas a las necesidades específicas de las cadenas.

La aplicación de técnicas de visión artificial se concreta, por el momento, en el reconocimiento facial sobre colecciones controladas con el objetivo de reconocer personajes públicos de interés especialmente en el ámbito regional, donde las soluciones de las grandes tecnológicas no son suficientes. En estos casos se imponen modelos con tasas de acierto exigentes para evitar los falsos positivos, con políticas de protección de datos también estrictas evitando trabajar sobre imágenes de niños o personas no públicas (*NISV*, *RSI*, *BR*).

El uso de IA para generar contenidos de forma automática a partir de material de archivo es, por el momento, anecdótico.

Finalmente, la aplicación de la IA en el archivo constituye un verdadero reto en cuanto a: escalabilidad, integración con los sistemas de producción y archivo y reorganización del trabajo de los documentalistas, tanto en las unidades de análisis como en las de atención a usuarios. Además, la integración de los metadatos en los sistemas de gestión de los archivos es una tarea compleja con un impacto considerable en la búsqueda y selección de los resultados, tanto para los usuarios expertos como para los usuarios finales que utilizan cada vez más estas herramientas sin intermediarios profesionales.

En conclusión, el uso de la IA en los archivos media presenta enormes oportunidades para mejorar la accesibilidad y la producción de nuevos contenidos basados en material de archivo. Sin embargo, plantea importantes consideraciones relacionadas con su integración en los flujos de trabajo, de calidad de los datos generados y de fiabilidad. A medida que se extienda su aplicación será crucial llevar a cabo investigaciones adicionales y fomentar la colaboración entre profesionales de diversas disciplinas para aprovechar todo el potencial de la IA, asegurando su integración responsable.

9. Referencias

AI4Media (2023). *The AI4Media project*.
<https://www.ai4media.eu>

Aragón Noticias (2021). "Aragón TV participa en la prueba piloto de una herramienta de inteligencia artificial". *Aragón noticias*, 13 noviembre.
<https://www.cartv.es/aragonnoticias/sociedad/aragon-tv-participa-en-la-prueba-piloto-de-una-herramienta-de-inteligencia-artificial-6130>

Åstrand, Mikaela; Ståhl, Sally (2023). "Finding without tagging: AI experiments for improved findability in the new media archive at SVT". In: *EBU Data technology seminar 2023*.
<https://tech.ebu.ch/publications/ai-experiments-for-improved-findability-in-the-svt-media-archive>

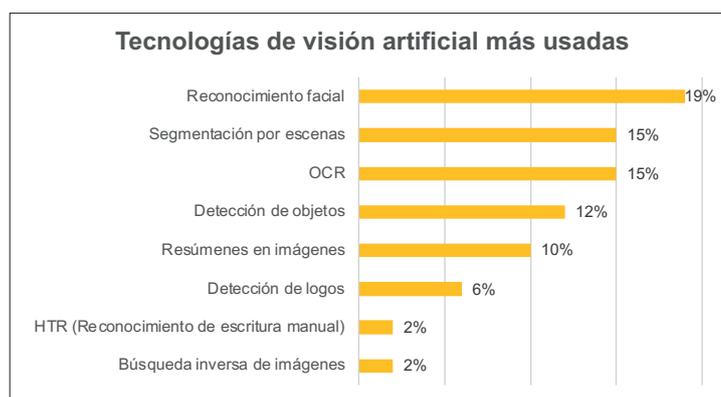


Gráfico 3. Porcentaje de uso de aplicaciones relacionadas con la visión artificial. Fuente: FIAT/IFTA (2023)

- Bailer, Wermer; Bauer, Christoph; Rottermann, Gernot** (2021). "Analysing used needs for automatic metadata creation and advanced search in the Tailored Media Project". In: *EBU MDN workshop 2021*.
<https://tech.ebu.ch/publications/analysing-user-needs-for-automatic-metadata-creation-and-advanced-search-in-the-tailored-media-project>
- Batrick, Kathey** (2022). "Case study: AI indexing. Assessing and selecting AI services". In: *FIAT/IFTA world conference: Archive out of the box!*
<https://fiatiftaworldconference2022.sched.com/event/149Cd>
- Batrick, Kathey; Petitpont, Frederic** (2022). "How Asharq News is leveraging multimodal AI indexing with Newsbridge: Inside the back-end, operational & technical workflows of a multilingual news channel". In: *FIAT/IFTA world conference: Archive out of the box!*
<https://fiatifta.org/world-conference-2022-recordings-available>
- BBC (2018a). *AI TV on BBC 4.1: BBC Four is giving an AI control of your TV for two nights*.
<https://www.bbc.co.uk/programmes/p06jrfcc>
- BBC (2018b). *Made by machine: When AI met the archive*. BBC Four.
<https://www.bbc.co.uk/programmes/b0bhwk3p>
- Berger, Jake; Armstrong, Andy** (2022). "BBC radio news scripts 1937-1995: Using an automated tagger to enable journeys across time and space". In: *EBU MDN workshop 2022*.
<https://tech.ebu.ch/publications/bbc-radio-news-scripts-1937--1995-using-an-automated-tagger-to-enable-journeys-across-time-and-space>
- Bouchet, Leonard; Ducret, Sebastien** (2019). "A visual feature extraction pipeline and its applications for radio television Suisse". In: *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles, part IX*, pp. 51-59.
<https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>
- Bruccoleri, Angelo; Iacoviello, Roberto; Messina, Alberto; Metta, Sabino; Montagnuolo, Maurizio; Negro, Fulvio** (2022). *AI in vision: High quality video production & content automation*. RAI. Radiotelevisione Italiana. Centre for Research, Technological Innovation and Experimentation. AI4media.
<https://www.ai4media.eu/whitepapers/ai-in-vision-high-quality-video-production-content-automation>
- Cátedra RTVE Universidad de Zaragoza (2017).
<https://catedrartve.unizar.es>
- Coppejans, Charlotte** (2021). *How artificial intelligence slashes editing time for Associated Press (AP)*.
<https://www.limecraft.com/how-the-associated-press-is-using-ai-to-create-automated-shot-lists>
- Couteux, Anne; Segura, Olivio** (2023). "News channel automatic segmentation". In: *FIAT/IFTA media management seminar 2023*.
<https://fiatifta.org/seminar/media-management-seminar-2023>
- Daniels, Marijn** (2023). "Metadata enrichment using lower third character recognition". In: *FIAT/IFTA media management seminar 2023*.
<https://fiatifta.org/seminar/media-management-seminar-2023>
- Daniels, Marijn; Degryse, Jasper** (2021). "How VRT automated the segmentation of programmes with AI". In: *EBU MDN workshop 2021*.
<https://tech.ebu.ch/publications/how-vrt-automated-the-segmentation-of-programs-with-ai>
- FIAT/IFTA (2022). *Archive Achievement Awards 2022*.
<https://fiatifta.org/awards-2022>
- FIAT/IFTA (2023). *AI Survey. The results are now available*.
<https://fiatifta.org/ai-survey-results>
- FIAT/IFTA Media Management Commission (2013). *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles. Part VI. Metadata as the cornerstone of digital archiving*.
<https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>
- FIAT/IFTA Media Management Commission (2017). *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles. Part VIII. Embracing automation – enhancing discoverability*.
<https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>
- FIAT/IFTA Media Management Commission (2019). *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles. Part IX. Game changers? From automation to curation - Futureproofing AV content*.
<https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>

FIAT/IFTA Media Management Commission (2023). *MMC Seminar. Download the slides.*

<https://fiatifta.org/mmc-seminar-2023-download-the-slides>

Förster, Constantin (2023). "Generation of training data for landmark recognition in videos using named entity recognition". In: *EBU data technology seminar 2023.*

<https://tech.ebu.ch/publications/landmarkner--generating-training-data-for-landmark-recognition>

Ghanbari, Shirin (2022). "Strategy and challenges for developing AI applications". In: *EBU AIM community meeting 2022.*

Green, Eva-Lis; Gupta, Jacqui (2019). "20 year of MMC seminars: changing sceneries, changing roles 1998-2019". In: *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles, part IX*, pp. 7-17.

<https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>

INA (2023a). *International affairs.* Institut National de l'Audiovisuel.

<https://www.ina.fr/institut-national-audiovisuel/international-affairs>

INA (2023b). *Collection, preservation, and documentation of audiovisual heritage.* Institut National de l'Audiovisuel.

<https://www.ina.fr/institut-national-audiovisuel/collection-preservation-and-documentation-of-audiovisual-heritage>

Lewis, Michelle; Jarret, Nicholas (2023). *Explore how Europeana Subtitled increased access to audiovisual heritage.*

<https://pro.europeana.eu/post/explore-how-europeana-subtitled-increased-access-to-audiovisual-heritage>

Lleida-Solano, Eduardo; Ortega-Giménez, Alfonso; Miguel, Antonio; Bazán-Gil, Virginia; Pérez-Cernuda, Carmen; De-Prada, Alberto (2022). *RTVE 2018, 2020 and 2022 database description.*

<http://catedrartve.unizar.es/reto2022/RTVE2022DB.pdf>

López-de-Quintana, Eugenio (2021). *AI algorithms for media cataloguing in Atresmedia Group.*

<https://fiatifta.org/awards-2021>

López-de-Quintana, Eugenio; León-Carpio, Antonio (2021). "Artificial intelligence for a role change in television archives: The Atresmedia–Etiqmedia experience". *Journal of digital media management*, v. 10 n. 2, pp. 177-187.

http://etiqmedia.com/ficheros/JDMM_10_2_JDMM0006_e_lopez_de_quintana_and_carpio-1.pdf

Manders, Tim (2019). "Harder, better, faster, stronger: Adding face recognition in the mix at NISV". In: *FIAT/IFTA world conference: AV archives in the all-media world.*

<https://fiatifta2019.sched.com/event/S0Sa/harder-better-faster-stronger-adding-face-recognitionin-the-mix-at-nisv>

Manders, Tim (2022). "Face recognition at the Netherlands Institute for Sound and Vision. Is it really harder, better, faster, stronger". In: *FIAT/IFTA media management seminar.*

<https://fiatifta.org/seminar/media-management-webinars-2022-2>

Manders, Tim; Wigham, Marie (2021). "More metadata, lots of links, but what do you do with them? Practical examples of the added value of automatic metadata and linked data for archive users". In: *2021 FIAT/IFTA world conference. Advancing the digital dividend.*

Martin, Camille; Segura, Olivio (2021). "Using AI tools to segment and describe broadcast live stream. In: *EBU MDN workshop 2021.*

<https://tech.ebu.ch/publications/using-ai-tools-to-segment-and-describe-broadcast-livestream>

Messina, Alberto (2021). "Enabling AI with dataset engineering". In: *EBU MDN workshop 2021.*

<https://tech.ebu.ch/publications/enabling-ai-with-dataset-engineering>

Messina, Alberto; Montagnuolo, Maurizio (2023). "Integrating open knowledge-bases and AI-tools". In: *EBU data technology seminar 2023.*

<https://tech.ebu.ch/publications/integrating-open-knowledge-bases-and-ai-tools>

Microsoft (2023). *Azure cognitive services documentation.*

<https://learn.microsoft.com/en-us/azure/cognitive-services>

Mühling, Markus; Korfhage, Nikolaus; Pustu-Iren, Kader; Bars, Joanna; Knapp, Mario; Bellafkir, Hicham; Vogelbacher, Markus; Schneider, Daniel; Hörth, Angelika; Ewerth, Ralph; Freisleben, Bernd (2022). "VIVA: visual information retrieval in video archives". *International journal on digital libraries*, v. 23, n. 4, pp. 319-333.

<https://doi.org/10.1007/s00799-022-00337-y>

Netherlands Institute for Sound and Vision (2023). *Mission and vision.*

<https://www.beeldengeluid.nl/en/about/mission-and-vision>

Open AI (2022). *Introducing ChatGPT.*

<https://openai.com/blog/chatgpt>

Open AI (2023). *Dalle-E 2*.
<https://openai.com/dall-e-2>

Parmentier, Matthieu (2021). "The challenge of data management and governance". In: *EBU MDN workshop 2021*.
<https://tech.ebu.ch/publications/governance-of-data-at-ftv-for-better-usage-and-valorisation-of-the-programs>

Petit, Thomas (2022). "Facial recognition: Trombinos". En: *FIAT/IFTA media management seminar*.
<https://www.youtube.com/watch?v=qM10EXb6988>

Reuters (2023). *Automated transcription, translation and identification*.
<https://liaison.reuters.com/help/reutersconnect#automated-transcription-translation-and-identification>

Reuters Staff (2020). *Reuters applies AI technology to 100 years of archive video to enable faster discovery, supported by Google DNI*.
<https://www.reuters.com/article/rpb-lavita-video-archive/reuters-applies-ai-technology-to-100years-of-archive-video-to-enable-faster-discovery-supported-by-google-dni-idUSKCN2591VO>

Rezzonico, Pietro (2020). "AI-based tools to help archive workflows". In: *EBU tech roundtable on archives 2020*.
https://tech.ebu.ch/publications/presentations/roundtable_2020_archive_workflows/RTS-archive_tools

Roche-Dioré, Axel (2023). "data.ina.fr a portal to promote media analytics". In: *EBU data technology seminar 2023*.
<https://tech.ebu.ch/publications/datainafr--a-portal-to-promote-media-analytics>

RTVE (2021). *Metadatado automático de contenidos del Fondo Documental de RTVE*.
<https://licitaciones.rtve.es/licitacion/licitaciones/detalle?id=1208797>

RTVE (2023). *Metadatado automático de contenidos del Fondo Documental RTVE*.
<https://licitaciones.rtve.es/licitacion/licitaciones/detalle?id=2200005>

Sánchez-García, Pilar; Merayo-Álvarez, Noemí; Calvo-Barbero, Carla; Diez-Gracia, Alba (2023). "Spanish technological development of artificial intelligence applied to journalism: companies and tools for documentation, production and distribution of information". *Profesional de la información*, v. 32, n. 2.
<https://doi.org/10.3145/epi.2023.mar.08>

Schreiber, Jonas (2022). "AI going local - conception and training of specialized Bavarian' AI models at Bayerischer Rundfunk". In: *EBU AI-AME 2022*.

Selkälä, Elina (2017). "Heading for AI". In: *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles, part VIII. Embracing automation - Enhancing discoverability*, pp. 169-189.
<https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>

Sonderregger, Janique (2023). "Speaker recognition service: Roles and expectations". In: *FIAT/IFTA media management seminar*.
<https://fiatifta.org/seminar/media-management-seminar-2023>

Speechmatics (2023). *The most accurate speech to text API*.
<https://www.speechmatics.com>

Startfruit Tagger (2023). *BBC Startfruit tagger*.
<http://starfruit.virt.ch.bbc.co.uk/#>

Steskal, Lubos (2023). "Using free-text queries to perform visual search in video archives". In: *EBU data technology seminar 2023*.
<https://tech.ebu.ch/publications/using-free-text-queries-to-perform-visual-search-in-video-archive>

Tverberg, Are (2021). "TV 2 Norway's AI hub and QA Monitor: A gateway to ASR services for a low-resource language". In: *EBU MDN workshop 2021*.
<https://tech.ebu.ch/publications/tv-2-ai-hub-and-qa-monitor--a-gateway-to-asr-services-for-a-low-resource-language>

Verwaest, Maarten (2022). "Using AI for automatic shot listing in news at Associated Press". In: *EBU MDN workshop 2022*.
<https://tech.ebu.ch/publications/using-ai-for-automatic-shot-listing-in-news---the-associated-press>

Viljanen, Kim (2022). "Content tagging in Yle Areena: Improving quality and process". In: *EBU MDN workshop, 2022*.
<https://tech.ebu.ch/publications/content-tagging-in-yle-areena---improving-quality-and-process>

Wenger-Glemser, Gabriele (2019). "Shaping the future Artificial Intelligence in AV archives". In: *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles, part IX*. pp. 61-78.
<https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>