

Documentación de las Ciencias de la Información

ISSN: 0210-4210

 EDICIONES
COMPLUTENSE<https://dx.doi.org/10.5209/dcin.68262>

Nuevos formatos de audio, viejos problemas

Mario Tascón¹

Recibido: 9 de marzo de 2020 / Aceptado: 5 de mayo de 2020

Resumen. La era digital ha multiplicado las posibilidades de escuchar audios por diferentes canales y soportes y ha multiplicado la oferta casi hasta el infinito. Sin embargo, esta revolución tecnológica no facilita, sino que dificulta el reto de la conservación de esta gigantesca producción. El presente artículo hace referencia a innovaciones que son ya de consumo masivo y a los problemas con que se enfrentan empresas y particulares a la hora de preservar su patrimonio sonoro.

Palabras clave: Altavoces inteligentes, podcast, audiolibros, Internet, patrimonio sonoro.

[en] New audio formats, the old same problems

Abstract. The digital era has multiplied the possibilities of listening to audios through the latest multimedia resources and the offer has increased almost to infinity. However, this technological revolution does not facilitate, but hinders the challenge of the conservation of this huge production. This article refers to innovations that are already of mass consumption and the problems faced by companies and individuals when preserving their sound heritage.

Keywords: Smart speakers, podcast, audiobooks, Internet, sound heritage.

Sumario: 1. Introducción. 2. El renacer del podcast. 3. Los audiolibros. 4. La conservación de los programas y materiales. 5. Conclusiones. 6. Bibliografía y webgrafía.

Cómo citar: Tascón, M. (2020): Nuevos formatos de audio, viejos problemas, en *Documentación de Ciencias de la Información* 43, 3-8.

1. Introducción

Los contenidos de audio siempre han sido queridos y apreciados por el público pero, desde su inicio en la distribución por radio o a través de discos a principios del siglo XX, han mostrado algunas dificultades para ser conservados y recuperados. La llegada de la digitalización de contenidos en las radios, coincidiendo casi con la aparición de internet parecía ofrecer un futuro de esperanza para almacenar de una forma lógica y sencilla la producción, pero la realidad ha sido el aumento de las dificultades para hacerlo por el incremento de la producción en todo el mundo y por la aparición de nuevos formatos en el campo de aquello que puede ser escuchado. Nos podemos encontrar que en la era del comienzo de la producción masiva de audio hay un agujero negro a la hora de poder recuperar las voces de principios del siglo XXI, como ya pasó con las del XX.

Este artículo toma como principal referencia el informe “¿Cómo suena tu marca? Estudio de la voz y el audio en España”, realizado por la consultora Prodigioso Volcán y la SEIM (Soluciones Estratégicas de Investigación de Mercados) y publicado en noviembre de 2019. Este estudio, basado en 1.053 entrevistas en línea y destinado a empresas y marcas, tenía por objeto ofrecer un retrato-robot del usuario de altavoces inteligentes y del consumidor de nuevos formatos de audio, como los podcast o los audiolibros. Lo que aquí se ofrece es un resumen de los resultados obtenidos, así como una reflexión acerca del panorama que se abre en este nuevo ecosistema que tiene a la voz como protagonista.

Los asistentes de voz y los altavoces inteligentes

Desde que en el año 2011 Apple lanzara su asistente de voz Siri, el consumo y producción de audio ha sufrido

¹ amtascon@gmail.com
Director de la consultora Prodigioso Volcán

varios giros inesperados que se suman a los que ya iba experimentando desde la aparición de internet, los sistemas de descargas de música y la invención de los podcast.

Siri es una inteligencia artificial con funciones de asistente personal, con su propia personalidad (no le gusta que le insultes), y que funciona en los aparatos de la marca californiana Apple (teléfonos, ordenadores y altavoces inteligentes). Esta aplicación utiliza sistemas de procesamiento del lenguaje natural para entender y responder preguntas así como hacer recomendaciones a los usuarios que le consultan.

Otras marcas han lanzado sus propias inteligencias artificiales como Cortana (Microsoft) o Alexa (Amazon). Sobre la continuidad de Cortana había dudas a fecha de cierre de este artículo.

Una de las principales diferencias de estas aplicaciones es el uso intensivo de la voz y el audio para comunicarse con sus dueños.

Al lanzamiento de este software se ha añadido el de los altavoces inteligentes, un nuevo tipo de dispositivo que suele sumar el altavoz y el micrófono y que permite conectar con instrucciones de voz los dispositivos del hogar así como hacer de altavoz de teléfonos, teles y aparatos musicales. En junio de 2018 llegó a España Google Home, el altavoz inteligente del gigante tecnológico. Cuatro meses después, en octubre, Amazon puso a la venta sus altavoces Echo también en nuestro país. Hoy, más de dos millones de internautas españoles usan uno de estos altavoces inteligentes. Un número en aumento.

La generalización de las interfaces de voz no solo ha revolucionado la forma de relacionarnos con la tecnología, sino que también replantea cómo se realiza una búsqueda e incluso se compra un producto. Igual que los ratones externos han empezado a convivir con los *trackpads* y *touchpads*, el teclado se ve ahora repartiendo tiempo con las órdenes dadas con la voz.

Pero aunque los altavoces inteligentes resultan el reclamo comercial más evidente de las tecnologías por voz, la gran mayoría de las interacciones orales tienen normalmente una vía más directa: el teléfono móvil.

En España, el 50,5% de los internautas utiliza asistentes de voz a través de uno o varios dispositivos, sobre todo tabletas y teléfonos inteligentes.

Los asistentes de voz más utilizados en nuestro país a través de móvil o tableta son Google Assistant, seguido de Siri, de Apple, y a buena distancia ya de Amazon, Alexa y Cortana. La ventaja de Google en este terreno es aplastante. La compañía es propietaria del sistema operativo mayoritario para teléfonos Android, que incluye Google Assistant preinstalado y que tiene en España una cuota de mercado del 90% frente al 8,9% de iPhones de Apple, que también tienen Siri instalado por defecto. Esta ventaja de partida se refleja por tanto en el dominio de Google entre los asistentes más utilizados.

Los asistentes de voz se utilizan, fundamentalmente, a través del teléfono inteligente. Le siguen la tableta o iPad y muy de cerca el coche, donde la integración es cada vez mayor. Tanto Amazon como Google han cerrado acuerdos con marcas como Volvo o Seat para incluir los asistentes de forma nativa en sus vehículos. No obs-

tante, otras compañías han optado por la creación de sus propios asistentes. Mercedes, BMW, Hyundai o Kia son algunos ejemplos.

El auge actual de la voz, sin embargo, no es independiente del desarrollo de otras tecnologías. El empuje extraordinario en los últimos años de campos como la inteligencia artificial, la computación en la nube, y el reconocimiento y síntesis de voz, entre otros, son en buena parte responsables de la facilidad con la que somos capaces de interactuar con teléfonos móviles y otros dispositivos electrónicos.

El otro gran ingrediente de su expansión son unas barreras de acceso relativamente bajas con unos precios de venta, por lo general, reducidos en comparación con otras herramientas tecnológicas como los móviles de última generación. Las empresas interesadas quieren entrar en este mercado y dominarlo porque consideran que es el centro desde el que se pueden desarrollar otros negocios relacionados con la domótica y, por supuesto, con el ocio y la información.

Esta combinación, liderada e incentivada sobre todo por compañías como Amazon y Google, explica la rápida expansión en los últimos años de este nuevo ecosistema expandido de voz y audio.

Amazon superaba en septiembre de 2019 las 100.000 aplicaciones de voz para su asistente Alexa. A principios de 2019, Google contaba con 4.253, un 150% más que un año antes. Estos números abundan en la importancia de un campo que para otros operadores tradicionales (fabricantes de equipos, radios y productoras) estaba pasando desapercibido.

El auge actual de la voz, sin embargo, no es independiente del desarrollo de otras tecnologías. El empuje extraordinario en los últimos años de campos como la inteligencia artificial, la computación en la nube, y el reconocimiento y síntesis de voz, entre otros, son en buena parte responsables de la facilidad con la que somos capaces de interactuar con teléfonos móviles y otros dispositivos electrónicos.

¿Por qué a los usuarios les gusta el asistente a pesar de que a veces, todavía, no les entiende? Porque la interacción por voz es más rápida, permite la multitarea y, por ello, es más cómoda. Quizás las mismas ventajas que siempre ha presentado la radio frente al resto de medios con el añadido de que, ahora, las instrucciones pueden hacerse con la propia voz frente a encender y apagar el aparato o sintonizar una emisora, además de que se amplía a otras instrucciones posibles en otros dispositivos conectados (desde la televisión al sistema de control del coche).

Se está pues modificando lo que antes solo era un canal unidireccional del audio hacia el oyente con la vuelta que este puede ejercer no solo desde lo mecánico sino desde la propia voz que es ahora quien puede dar órdenes.

Es muy interesante ver qué uso hacen las personas de estos aparatos ya que, en el fondo, la mayoría del tiempo lo utilizan de forma similar a la forma en la que antes consumían la radio solo que, ahora, también mezclan un uso similar al que antes hacían en el ordenador o teléfono cuando consultaban cosas concretas en Google. En-

tre escuchar (música, noticias, radio y podcast) y buscar respuestas se concentra la inmensa mayoría del tiempo de uso de estos dispositivos. Parece por ello que mezclan el papel del receptor de radio, del altavoz de música y del buscador de internet utilizado ahora sin manos.



Figura 1. Para qué se utiliza el altavoz inteligente².

Fuente: Informe “¿Cómo suena tu marca?”. Prodigioso Volcán/ SEIM. 2019

A pesar de todas las ventajas sobre el aumento de uso de estos dispositivos han aparecido algunos nubarrones relacionados con la privacidad. Amazon, en abril de 2019, reconoció en un comunicado que sus empleados tenían acceso a algunas conversaciones grabadas directamente desde los dispositivos: “anotamos un número extremadamente pequeño de interacciones de un conjunto aleatorio de clientes para mejorar la experiencia del cliente, siempre después de haber utilizado la palabra de activación de un dispositivo”. “Esta información nos ayuda a entrenar nuestros sistemas de reconocimiento de voz y comprensión del lenguaje natural, para que Alexa pueda entender mejor sus peticiones y garantizar que el servicio funcione bien para todos”.

Pero también Apple ha tenido que disculparse por un uso cuestionable del sistema de grabación y cambiar su sistema de trabajo tras realizar “una profunda revisión” de sus prácticas y políticas. La compañía aseguró al diario *El País* que ya no conservará por defecto grabaciones de audio de las interacciones con Siri y que seguirá usando transcripciones generadas por ordenador para ayudar a mejorar al asistente³.

² “¿Cómo suena tu marca? Estado de la voz y el audio en España”. Prodigioso Volcán / Soluciones Estratégicas de Investigación de Mercados (SEIM). Diciembre, 2019. <https://www.prodigiosovolcan.com/sismogramas/informe-voz/>

³ Rubio, Isabel. “Apple se disculpa por escuchar conversaciones privadas por Siri”. *El País*, 29 de agosto de 2019. Edición digital. https://elpais.com/tecnologia/2019/08/29/actualidad/1567086732_801829.html

El miedo de los usuarios a tener un espía conectado permanentemente en sus salones o dormitorios extiende una cierta sombra sobre el desarrollo de estos aparatos. Miedos acrecentados por otras noticias rocambolescas como cuando a un robot de cocina comercializado en los supermercados Lidl unos hackers le encontraron en su interior un micrófono.

2. El renacer del podcast

El panorama del consumo de audio cambió completamente cuando aparecieron los *podcast* en el comienzo de este siglo. Lo que inicialmente fue solo un avance técnico para poder recibir y escuchar programas de radio troceados, una especie de radio a la carta, ha tenido después un extraordinario desarrollo hasta llegar con fuerza a nuestros días. Lo que inicialmente surgió como una nueva forma de distribuir lo que se producía en la radio convencional para que los usuarios lo pudieran escuchar cuando les pareciera mejor se ha convertido en una industria en sí misma que ha venido a agitar el panorama del audio.

Pero el empujón definitivo lo daría “Serial”, un programa de periodismo de investigación presentado por Sarah Koenig. “Serial” narra historias reales en múltiples episodios, tiene diferentes temporadas y ocupó el número uno en iTunes incluso antes de su debut, permaneciendo allí durante varias semanas. A partir de septiembre de 2018, los episodios de las temporadas uno y dos se han descargado más de 340 millones de veces, estableciendo el récord mundial actual. “Serial” mostró con mucha claridad el potencial que tenía este nuevo formato y, a partir de ahí, en todo el mundo se produjo una eclosión de productoras, programas, plataformas e interés.

Pero incluso en español hay lugares de podcast que se han convertido en mitos de los fanáticos (que los hay) de este formato, como Radio Ambulante: un sitio web que publica crónicas latinoamericanas en audio en español. Lo presenta el escritor peruano Daniel Alarcón. Las primeras emisiones se lanzaron en 2012 y, desde 2016, Radio Ambulante es distribuido semanalmente por la National Public Radio.

En nuestro país, además de iniciativas como La Cafetera de RadioCable, la cadena SER ha sido la que ha realizado la apuesta más grande por la producción de productos propios de podcast en Podium Podcast y ahora en Podium Studios. El proyecto, que se presentó en 2016, tiene actualmente un catálogo muy amplio. Parte del éxito de este formato se hizo evidente en los premios Ondas de PRISA Radio, cuando se otorgó el Premio al Mejor Podcast a “Las tres muertes de mi padre” del periodista Pablo Romero.

Pero ¿escuchan los oyentes programas distintos a los de la radio? la respuesta es sí, de manera creciente. Por temáticas, y según el estudio ya citado, en podcast parecen destacar los contenidos de entretenimiento (44,4%), humor (34,6%) y cultura (34%).

Si atendemos a las plataformas utilizadas, es llamativo el impacto de Spotify, que se posiciona como líder. La plataforma de *streaming* fue de las últimas en activar

una estrategia específica para este formato, pero ha hecho una fuerte apuesta. En el gráfico se puede ver el tipo de contenidos más consumidos en podcast en España a principios de 2020.

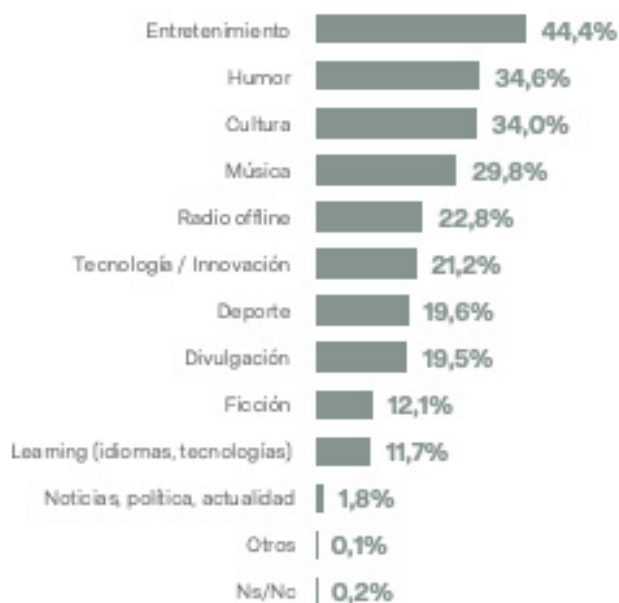


Figura 2. Temáticas favoritas del consumidor de podcast.
Fuente: Informe “¿Cómo suena tu marca?”. Prodigioso Volcán/ SEIM. 2019

3. Los audiolibros

Los audiolibros, que históricamente siempre han representado en nuestro país un pequeño porcentaje de ventas en el mercado de libros, han emergido muy recientemente como un nicho robusto. El formato se ha convertido en un complemento cada vez más habitual, hasta el punto de competir con los dos productos que reinan en España dentro del sector, el libro y su versión digital, el *e-book*.

El audiolibro es una grabación de un libro leído en voz alta. Puede ser una lectura completa o una versión resumida.

En el mundo anglosajón los audios hablados procedentes de libros han estado disponibles en escuelas y bibliotecas desde los años 30. Tuvieron un gran auge en la era de las cintas de cassette y, luego, de los CDs. En la década de 1980 se podían encontrar en las estanterías de las librerías de Inglaterra y Estados Unidos.

Pero hoy el 27% de los internautas en España escucha audiolibros. No obstante, el grado de conocimiento es bastante superior al de uso. El 71,6% de sus usuarios asegura que su consumo es muy esporádico.

El formato que originalmente estaba pensado más para personas discapacitadas (la ONCE tiene un gran catálogo) fue luego ampliándose, pero ha sido la llegada de nuevas plataformas la que está consiguiendo un gran salto en el consumo. Aunque en el arranque de este mercado han sido Apple e iVoox los que lo han liderado, ha sido de nuevo Spotify uno de los vectores de creci-

miento, pero también ha sido determinante la irrupción de Storytel, una empresa de origen sueco que ofrece un servicio de streaming de audiolibros, e-books y podcasts a través de su propia plataforma de suscripción. Storytel cuenta con más de un millón de suscriptores en el mundo y un catálogo de casi 500.000 títulos. El quinto en discordia es Audible, una empresa norteamericana, actualmente propiedad de Amazon, que ofrece audiolibros, programas de radio y televisión, podcasts y versiones en audio de algunas revistas y periódicos, como *The New York Times*. Audible es también el mayor productor mundial de audiolibros. A sus contenidos solo se puede acceder siendo suscriptor o cliente de Amazon.

En cualquier caso, frente al podcast o los contenidos de voz y audio servidos por los altavoces inteligentes, los audiolibros representan un porcentaje más pequeño aunque creciente.

Muy relacionada con la producción de libros hablados está la posibilidad de lectura con computadora y voz sintética que cada vez está más avanzada y resulta más difícil de distinguir de la natural. Además, en muchos casos, los textos pueden venir marcados con sistemas de etiquetas especiales para ayudar a una pronunciación más “humana”. Muchas páginas web incorporan como ayuda botones que lanzan una lectura sintética de los contenidos de la página.

4. La conservación de los programas y materiales

Los coordinadores de esta revista, en su *call for papers*, alertaban del problema que he señalado al comenzar el artículo:

“La imposibilidad de conservar la producción de audio por falta de soportes físicos, ausencia de profesionales cualificados y desidia de las propias empresas radiofónicas en la preservación de sus fondos, ha impedido recuperar las emisiones desde su implantación en España, en 1924, hasta bien entrado el siglo XX. Cientos de miles de horas de radio se han perdido irremediablemente y jamás podremos escucharlas, aunque se hacen esfuerzos por reconstruir sus programaciones para tener una cierta idea de cómo sonaba entonces este medio: el más cercano, el más inmediato, pero también el más fugaz”.

Este problema –que parece increíble en tiempos en los que toda la información del mundo se encuentra disponible tras el teclado de un teléfono móvil–, ofrece una sorprendente paradoja. Nunca hemos tenido tanta información accesible, pero esta es efímera y no encuentra siempre la mejor forma para perdurar. Gran número de plataformas; ausencia de estándares en la grabación y digitalización; materiales en bruto en formatos solo de audio digital sin meta, informaciones ni transcripciones textuales de contenidos; sistemas de etiquetas diferentes y una cierta dejadez en las diferentes organizaciones productoras de contenidos por archivarlos una vez emitidos o consumidos: todo lo anterior parece que ha conseguido el efecto advertido en la carta.

Por una parte, a la hora de etiquetar un audio —en el caso de hacerse—, se suelen destacar aquellas etiquetas que ayuden a promocionarlo en las plataformas. Se eligen las que puedan darle una mayor visibilidad a la hora de ubicarse en Apple o iVox, también las relacionadas con el contexto actual de ese documento, pero no se colocan las que, tiempo después, podrían ayudar a encontrarlo. Por otro lado, el propio proceso de etiquetado suele estar condicionado por las herramientas con las que estos audios se suben a los portales, sistemas editoriales o distribuidores. Estas herramientas se diseñan desde el inicio para esa función, pero no para una recuperación posterior.

El último movimiento de Google anunciando que integraría los podcast reproducibles entre los resultados de búsqueda para ordenadores y también en su versión móvil, simplifica mucho más para los usuarios el consumo de estos archivos y favorece la etiquetación basada en que los audios se encuentren de forma fácil para consumir a través de un buscador, pero no para ser accesibles a lo largo del tiempo.

Hoy, cuando una persona busca un podcast, Google le muestra en pantalla los tres episodios más recientes y le permite escucharlos a través del reproductor web, similar a la aplicación Google Podcast del sistema operativo Android. Sin necesidad de abrir otro software y de forma intuitiva, el usuario puede escuchar lo elegido fácilmente, incluso en los reproductores de los coches. La sencillez de uso segmenta de algún modo la forma de archivar en beneficio del consumo inmediato. Predominarán las etiquetas que lo faciliten.

Este problema tiene un agravante: quienes producen programas de audio fuera de los circuitos profesionales de las emisoras de radio se encuentran con que, a la hora de subir sus archivos, la mayoría de las plataformas exige muy poco para que éstos se puedan escuchar: un título, categoría, unas etiquetas genéricas, una foto y marcando el formato e idioma ya está. Toda indicación y consejo sobre esto se reduce a la información disponible en internet de otros usuarios que, como ya avisamos, solo se centran en que los archivos suban a los primeros lugares de los resultados de búsqueda de las páginas.

Por otro lado sigue sin haber un consenso amplio sobre las taxonomías que hay que utilizar y en una buena parte de las productoras de audio se emplean conjuntos propios de palabras y conceptos que no se relacionan con ningún otro de fuera de su ámbito empresarial o local. Muchos departamentos documentales han extendido los sistemas de clasificación que ya tenían en otra época y apenas se vislumbran en el horizonte otros nuevos que contemplen todos los cambios necesarios.

Tampoco colabora, de momento, el análisis y generación automática de sistemas de ayuda para la búsqueda con programas de inteligencia artificial ya que, en muchos casos, por costes e incluso porque la tecnología no está suficientemente madura para conseguir resultados precisos, quedan fuera del alcance de las organizaciones. Desarrollos científicos como la taxonomía denominada Audio Signal Classification (ASC) están muy

alejados de ofrecer soluciones útiles en el entorno de la radio y los podcast.

Y, por supuesto, el aumento en el número de archivos que se suben cada segundo a las redes ha convertido el mundo digital en un universo en el que es francamente difícil atender con detalle a cada uno de los ficheros que circulan por el mismo. Por ello, estos ficheros adquieren la etiquetación que los propios buscadores extraen del contexto de los archivos. Google, por ejemplo, cuando localiza un archivo en alguna página web toma también la información de esa página deduciendo que los datos que rodean a un audio incrustado de alguna forma lo describen.

Los sistemas de audiodescripción, pensados en un inicio especialmente para que los usuarios con discapacidad visual o limitada pudieran ver la narración en pantalla o, incluso, cantar leyendo las letras en modelos de karaoke, han sido de ayuda pero no son suficientes ni todos los audios los incorporan en origen, ya que hacerlo es un proceso —sobre todo cuando es manual—, tedioso y laborioso.

La organización española CIVIO, cuya misión es acabar con la opacidad en las instituciones vigilando a los poderes públicos, ha presentado recientemente un proyecto, financiado por Google, que usa activamente esa audiodescripción que existe en los telediarios españoles y permite realizar varias tareas como declaran en su página web:

- Acceder a las transcripciones del Telediario 1 (el que se emite a las 15.00 horas) y del Telediario 2 (21.00h) de Televisión Española desde 2014 hasta la actualidad.
- Buscar menciones a términos, personas, entidades y lugares en los contenidos de estos informativos.
- Comprobar visualmente la evolución del número de referencias a un término concreto a lo largo del tiempo.
- Encontrar el fragmento exacto del informativo en el que se hizo mención al objeto de tu búsqueda.
- Descargar los resultados de tu búsqueda en formato .csv (un formato para manejar en bases de datos y poder hacer estudios propios).

Esto supone un gran avance para desenterrar de cara al público mucho material de los telediarios que, de otra forma, permanecería oculto o sería de muy difícil recuperación. Y ha sido posible porque desde 2014 los telediarios tienen la obligación de incluir esa audiodescripción.

Donde también parece haber adelantos es en el software que convierte voz a texto de manera automática. Este es un campo en el que los grandes emporios tecnológicos invierten muchos millones de dólares y que lideran IBM, Microsoft y Google. El software que consigue escuchar a una persona hablando y transcribe lo que dice funciona cada vez con una precisión mayor y recoge con más fidelidad la palabra pronunciada. Watson en el caso de IBM o los avances de Azure de Microsoft son capaces desde 2017 de reconocer las palabras y oraciones emitidas por una única persona al mismo nivel que un

ser humano. Microsoft asegura que hoy ya puede hacerlo distinguiendo incluso entre voces diferentes en el caso de que se esté produciendo un diálogo. Siguen, en cualquier caso, existiendo algunos problemas con los acrónimos y las expresiones propias de determinados dominios de especialidad humana, pero el software avanza también para resolver esos casos.

Todos estos logros realizados con herramientas de Aprendizaje Profundo de Inteligencia Artificial están permitiendo avances significativos al añadir esas descripciones a los archivos de audio que permiten, incluso, localizar el punto exacto donde se dijo una palabra, hacerlo de manera masiva e incluso separada de su origen, con lo que las propias plataformas pueden aplicarlo en directo. Programas como el popular Power Point empiezan también a tener la capacidad de añadir a sus archivos la transcripción automática del audio del conferenciante sobre las diapositivas que ha utilizado generando un nuevo archivo enriquecido y que es más fácil de encontrar.

En cambio, donde existe menos progreso es en cuanto a los sistemas de vocabularios controlados y taxonomías, que parecen avanzar a menor velocidad y de una forma más dispersa. La promesa de la web semántica y su aplicación documental no acaba de añadir muchas mejoras significativas a lo que ya se venía realizando con buenos sistemas de etiquetas. La esperanza de las ontologías que nos permitan tener un cierto “pensamiento lógico” que nos ayude parece lejana de momento. Como señala el profesor Codina: “Lo peor (de las ontologías) es que únicamente son una aspiración, un ideal, que de momento solo funciona en laboratorios y sitios experimentales”. Y el profesor remata: “Una ontología podría servir para proporcionar un sistema de búsqueda inteligente, capaz de realizar inferencias y actuar como si fuera una verdadera inteligencia artificial, pero aún no están implantadas en forma funcional”⁴.

6. Bibliografía y webgrafía

- Prodigioso Volcán / SEIM (Soluciones Estratégicas de Investigación de Mercados). (2019). “¿Cómo suena tu marca? Estado de la voz y el audio en España”. Recuperado de: <https://www.prodigiosovolcan.com/sismogramas/informe-voz/>
- Rubio, Isabel. “Apple se disculpa por escuchar conversaciones privadas por Siri”. *El País*, 29 de agosto de 2019. Edición digital. Recuperado de: https://elpais.com/tecnologia/2019/08/29/actualidad/1567086732_801829.html
- Amazon. Comunicado de prensa de 10 abril de 2019. Recuperado de: <https://www.bloomberg.com/news/articles/2019-04-10/is-anyone-listening-to-you-on-alex-a-global-team-reviews-audio>
- Codina, Lluís. “Taxonomías y ontologías: qué son y cómo se aplican en medios de comunicación”. Recuperado de: <https://www.lluiscodina.com/taxonomias-ontologias/>

La verdad es que a pesar de que llevamos diez años hablando de la web semántica, y de reconocer las posibles ventajas, más allá de algunos proyectos científicos y de la facilidad de conectar datos enlazados no parece todavía un campo que nos ayude demasiado en nuestro propósito de almacenar y recuperar audios. Sobre todo porque, cada vez más, las búsquedas se realizan en lenguaje natural.

5. Conclusiones

La voz como interfaz se está revelando como la forma más simple de interactuar. Es el principal medio de comunicación entre humanos desde hace miles de años y nuestro cerebro ha evolucionado para que seamos capaces de identificarla y de expresarnos con ella. Por encima de otras vías de relación, la voz representa un grado de accesibilidad muy superior al de cualquier tecnología anterior. El usuario no tiene que aprender un nuevo método, simplemente tiene que hacer algo que ya adquirió con su propio aprendizaje vital: hablar.

Sin embargo, detrás de ese acto sencillo es necesario un potente desarrollo tecnológico. Las empresas, marcas o negocios que busquen un lugar en este ecosistema deberán ofrecer eficacia y eficiencia en sus aplicaciones para satisfacer al usuario con el menor número de pasos posible. La inteligencia artificial, la computación en la nube y el reconocimiento y síntesis de voz aumentarán las posibilidades de la interacción oral. Mientras en la búsqueda de una interfaz gráfica se ofrecen al usuario decenas de opciones, en la búsqueda por voz se darán uno o como mucho dos resultados a elegir. Por eso las marcas pelearán por estar en tan restringida oferta: una batalla que está todavía en su infancia y cuyas reglas se van a ir escribiendo sobre la marcha.

⁴ Codina, Lluís. “Taxonomías y ontologías: qué son y cómo se aplican en medios de comunicación”. <https://www.lluiscodina.com/taxonomias-ontologias/>