

Proyecto Comma: metadatos sonoros

La BBC optimiza la recuperación de sus archivos radiofónicos

DIEGO RODRÍGUEZ GUTIÉRREZ

La cadena internacional, de referencia en todo el mundo, ultima el lanzamiento de una idea pionera en el ámbito de los documentos sonoros. Se trata de un sistema de gestión de archivos de audio basado en la obtención de metadatos a través de detectores y transcritores de voz. Teniendo en cuenta que la BBC comenzó a emitir en 1922, el éxito del proyecto puede facilitar la recuperación de un inmenso volumen documental que podría marcar un antes y un después en el trabajo tanto de periodistas y documentalistas internos como de todo tipo de profesionales interesados, ya que se pretende abrir el servicio a otras instituciones culturales como la Biblioteca Británica.

Rod Cooper, productor de la BBC, afirma que cuando llegó la digitalización, esta se tomó como la solución al problema de la recuperación de información de los grandes fondos radiofónicos. Sin embargo, sostiene que lo único que se había conseguido hasta ahora era sustituir un archivo

(continúa en pág. 37)



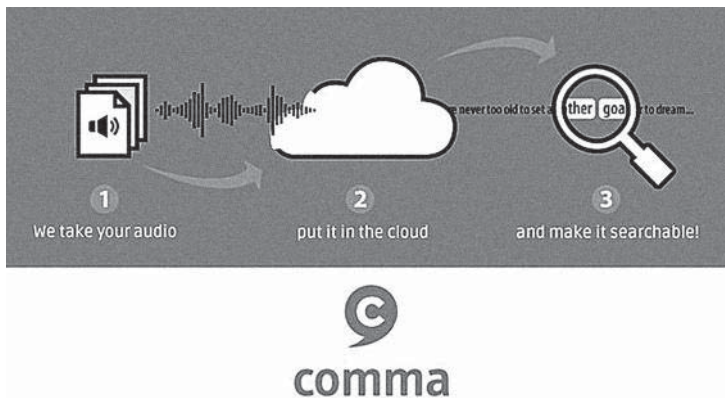
(viene de pág. 28)

físico polvoriento por un equivalente digital que presentaba las mismas dificultades en cuanto a la indexación de su contenido por muy fácil de enviar, compartir y escuchar que fuera. COMMA toma un archivo de audio digital y ana-

liza el sonido para extraer metadatos, trasladando el habla a texto para crear una transcripción, identificar los temas e incluso distinguir quién está hablando, usando la detección de voz. Según Cooper, el sistema sería "capaz de identificar los bits donde tú hablas en una entrevista". Esto sería

especialmente útil para muchos documentalistas del mundo de las comunicaciones que necesitan buscar audios de forma muy rápida. Para este productor solo hay un defecto: el sistema es demasiado exigente con los mínimos de calidad del audio de entrada para que pueda ser reconocido, identificado e indexado correctamente.

Se prevé que vea definitivamente la luz en la primavera de 2015. Hasta entonces los trabajadores implicados de los dos socios comerciales contratistas de las tareas, las empresas Kite y Somethin'Else, deberán concluir el desarrollo informático necesario para garantizar su viabilidad. Tras cinco años en el proyecto afirman que "tecnologías como el reconocimiento de voz



automatizado, la identificación del hablante y etiquetado automatizado han alcanzado un nivel de precisión en el que podemos empezar a obtener resultados impresionantes en muchos tipos de audio. Al analizar automáticamente los archivos de sonido y tomar de-

tos. Para tratar de construir un primer patrón sobre el que trabajar, se desarrollaron una serie de algoritmos de procesamiento de voz basados en la segmentación entre altavoz, identificación y detección de género, en la emisión de texto en función de los segmentos

tro de un vocabulario amplio, sin límites.

El departamento I+D de la BBC decidió desarrollar estas tecnologías de extracción de metadatos automáticos de una manera que permitiera el procesamiento a gran escala y se optó por la construcción de una plataforma basada en la nube que permitiera trabajar con grandes archivos de forma rápida y barata. Se trata de introducir colecciones de medios de comunicación y procesarlas para extraer metadatos útiles. Este tipo de flujo de trabajo se presta muy bien a una moderna infraestructura de nube. Mientras que tradicionalmente se podría estar limitado por la cantidad de equipos disponibles para ejecutar el proceso, utilizando una infraestructura de nube las posibilidades son interminables. Una gran ventaja de estos recursos es que se pueden comprar por horas o ampliar de forma masiva para satisfacer las demandas de uso. La plataforma cloud computing elegida fue Amazon AWS.

Actualmente "COMA" se está utilizando de forma experimental en algunas salas de prensa, donde es capaz de tomar automáticamente el audio del videoclip, identificar los diferentes oradores y proporcionar una transcripción de la grabación de prensa. Después añade los metadatos importantes para usuarios dentro de la base de datos.

De cara al futuro inmediato y previo a su lanzamiento, se ofrece la posibilidad de recibir actualizaciones e incluso probar la plataforma a cualquier institución cultural, empresa de medios de comunicación o entidad con un considerable volumen de archivos sonoros a través de la siguiente dirección: <http://comma.io/> ■

cisiones informadas sobre el contenido y los altavoces, estas herramientas pueden ayudar eficazmente a rellenar los huecos que faltan en metadatos de nuestro archivo". COMMA está basado en un conjunto de algoritmos de procesamiento de voz llamado Kiwi. Se empezó a trabajar con ellos en 2011 a través de un primer experimento con un gran archivo de radio-discurso, el fondo de la BBC World Service, que en ese momento tenía muy pocos metada-

detectados y en el etiquetado automatizado con identificadores DBpedia (gran base de datos extraídos de Wikipedia).

Además, tuvieron que tener muy en cuenta importantes factores. El enfoque debía ser lo más eficiente posible, al trabajar para usuarios con alta demanda de recuperación de información como son los profesionales de los medios de comunicación. Por otro lado, ante la gran variedad de temas de los contenidos, el algoritmo debía ser capaz de operar den-