

TLP: WHITE



Textualización de Grabaciones

Premios Energetic 2021

Índice

1	INTRODUCCIÓN	1
2	TEXTUALIZACIÓN DE GRABACIONES.....	2
2.1	OBJETIVOS	3
2.2	IMPACTO	3
2.3	DESCRIPCIÓN Y FUNCIONALIDADES.....	4
2.4	FASES DEL PROYECTO.....	5

1 INTRODUCCIÓN

La transición ecológica y la digitalización son dos realidades incuestionables que están condicionando el presente y van a condicionar el futuro, obligando a reorientar las políticas públicas y la acción de las Administraciones.

En esta línea, se está desarrollando el Plan estratégico Justicia 2030, para la transformación del servicio público de Justicia en los próximos años, con el objetivo de consolidar los derechos y garantías de los ciudadanos, promover una mayor eficiencia del servicio público y, garantizar el acceso a la Justicia, en todo el territorio.

En este nuevo marco, aprovechando el avance del desarrollo digital de los sistemas de gestión procesal junto con el potencial de la aplicación de la inteligencia artificial contribuye al desarrollo de nuevas soluciones que van a apoyar la mejora en la gestión y conocimiento, además del impacto en los servicios ofrecidos a los ciudadanos a través de una justicia más abierta, más eficiente, ágil y eficaz.

La ley dispone que las actuaciones orales en vistas, audiencias y comparecencias celebradas ante los jueces o magistrados o, en su caso, ante los letrados de la Administración de Justicia, se deben registrar en soporte apto para la grabación y reproducción del sonido y la imagen y que no pueden transcribirse.

El hecho de que las actuaciones judiciales estén documentadas únicamente a través del archivo que contiene la grabación audiovisual de las mismas, ha supuesto que distintos operadores reclamen con insistencia la posibilidad de disponer de una transcripción de lo acontecido en ellas.

Ello ha llevado a plantearse la necesidad de facilitar y agilizar el trabajo diario de los Órganos Judiciales y Fiscalías mediante la transformación en texto de cualquier actuación procesal documentada en formato audio o audiovisual: declaración grabada, videoconferencia, etc.

El proyecto tiene por objeto la implantación y entrenamiento de un sistema que permita localizar rápidamente cualquier información relativa a las declaraciones y manifestaciones producidas durante la celebración de una actuación que se haya documentado mediante la grabación audiovisual. Esto permitiría a los operadores acceder a declaraciones o manifestaciones concretas realizadas durante la actuación, sin tener que hacer la visualización completa del acto. Este sistema, al reflejar en un texto todo lo acontecido y manifestado, permite buscar con más facilidad la información que se desea utilizar.

En el siguiente enlace se puede acceder a un vídeo explicativo de Textualización de Grabaciones <https://youtu.be/YRwdqPvsxdk>

2 TEXTUALIZACIÓN DE GRABACIONES

La evolución en técnicas de procesamiento de lenguaje natural y de su aplicación práctica mediante costes razonables, ha permitido emprender el análisis del procesamiento de determinadas fuentes de información “no estructurada”, como las grabaciones de vistas, de modo que pueda aportar valor para realizar búsquedas, interrelación de información, indexación, volcados a elementos estructurados, etc.

En los últimos tres años se han celebrado más de 200.000 vistas judiciales al año, lo que supone más de 50.000 horas de grabaciones. Para la extracción de la información contenida en la grabación de una vista, se precisa el visionado completo o parcial por parte de los jueces y magistrados, sin más ayuda que desplazarse de manera manual por la reproducción y dedicar un tiempo costoso en esta tarea.

Para agilizar su trabajo, el Ministerio de Justicia ha desarrollado esta solución basada en técnicas de aprendizaje neuronal e integrada con los sistemas de grabación de las Salas de Vistas, que realiza el procesamiento de archivos de vídeo o audio con la finalidad de extraer el texto correspondiente de las intervenciones orales que se realicen en juicios, vistas y comparecencias grabadas, de forma automatizada.

La puesta en marcha de este sistema en el ámbito de la justicia supone un hito de gran importancia, ya que es la primera vez que se aplican tecnologías disruptivas como instrumento de apoyo a la actividad judicial.

Proyecto que permite la **generación automática del texto tomando como origen los videos grabados durante las vistas y las declaraciones**, permitiendo el desarrollo de nuevos servicios y funcionalidades.

Integrado con Visor Horus

- Basado en Inteligencia Artificial**
Un sistema que presenta componentes de aprendizaje o entrenamiento que va perfilando la precisión de la textualización.
- Permite un ahorro de tiempo**
En la búsqueda y localización de contenido a partir de palabras clave sobre el texto de las intervenciones habladas.
- Ofrece una mayor agilidad**
En la gestión de la información contenida en las vistas y el acceso directo a las intervenciones de los intervinientes.

FUNCIONALIDADES:

 <p>GENERACIÓN DE TEXTO AUTOMÁTICA</p> <p>Muestra el texto correspondiente a las intervenciones grabadas en el vídeo con su duración.</p>	 <p>BÚSQUEDA DE TEXTO LIBRE</p> <p>A través de un campo de texto libre se puede localizar contenido en la grabación escribiendo la palabra clave.</p>	 <p>CONSULTA DE INTERVINIENTE</p> <p>Permite visualizar a los intervinientes en los diálogos de la textualización, identificados mediante un número.</p>
 <p>VISUALIZACIÓN DE TEXTUALIZACIÓN</p> <p>Muestra la línea de tiempo pudiendo navegar digitalmente por las intervenciones y también permite buscar palabras clave.</p>	 <p>DESCARGA DE TEXTUALIZACIÓN</p> <p>Permite obtener el registro de la textualización de la vista en un documento editable de texto.</p>	 <p>CREACIÓN DE ETIQUETAS</p> <p>Posibilidad de generar etiquetas y visualizar en qué intervalos de tiempo aparece dicha palabra/etiqueta.</p>

Imagen real del sistema de Textualización de Vistas

2.1 Objetivos

El **objetivo** de esta aplicación es agilizar el trabajo de los jueces y magistrados en la revisión de las grabaciones de las vistas y declaraciones, evitando que tengan que realizar el recorrido del video sin más apoyo que la memoria de haber asistido de un modo personal.

Los objetivos específicos son:

- ▶ Dotar a las aplicaciones de consulta de los expedientes judiciales de funciones que faciliten el uso de medios audiovisuales y su organización.
- ▶ Extraer datos de soportes que, aun siendo digitales, no permiten un procesamiento computacional sencillo de la información contenida.
- ▶ Implementar las bases para la carga de información no estructurada a sistemas de interpretación de datos masivos.
- ▶ Proporcionar una primera herramienta, que, mediante el uso de técnicas de Inteligencia Artificial, abra las puertas a la explotación de estas nuevas ayudas tecnológicas, y, permita evaluar su potencial redundando en la agilización de los sistemas Judiciales.

2.2 Impacto

Textualización de Grabaciones aporta importantes beneficios que mejorarán con su uso:

- ▶ **Mejora en la productividad:** supone un ahorro estimado del 60% del tiempo que los usuarios tardan en localizar un extracto específico de una grabación y realizar el análisis de las vistas celebradas. El usuario lo está explotando, no solo para la revisión de los vídeos, sino como apoyo de la información extraída para la redacción de dictámenes y conclusiones.
- ▶ **Eficiencia:** facilita a los jueces y magistrados el entendimiento jurídico del procedimiento global, mejorando, por ende, el hallazgo de resoluciones jurídicamente motivadas en un menor tiempo.
- ▶ **Fiabilidad:** este sistema presenta un nivel de acierto superior al 80% en la textualización de las grabaciones. No obstante, se espera que a medida que se vaya avanzando en la extensión del sistema al resto de órganos judiciales se irá reentrenando la herramienta, lo cual supondrá un aumento progresivo de la fiabilidad.
- ▶ La **integración con la herramienta Visor Horus** permite que el conjunto formado por la grabación, las marcas introducidas por el personal del órgano judicial y el fichero de subtítulos, esté disponible cuando los usuarios accedan a los expedientes judiciales electrónicos a través del Visor Horus, incluyendo las funcionalidades de búsquedas y otras auxiliares para copiar y pegar el texto procesado del contenido de la grabación.

- La **capacidad de aprendizaje por reentrenamiento**, permitirá la ampliación y mejora de prestaciones de esta herramienta. En este sentido, se ha de tener en cuenta que el lenguaje no está compuesto por un conjunto de frases, palabras y expresiones uniformes, sino que hay distintos lenguajes adaptados a rasgos culturales concretos como expresiones, acentos etc, así como acepciones o conceptos jurídicos, que el sistema tiene que aprender de forma previa para incrementar sus prestaciones.
- **Sostenibilidad y eficiencia energética:** Textualización de grabaciones permite ahorrar energía eléctrica gracias al ahorro de tiempo que supone para los usuarios el poder visualizar directamente la parte de la grabación de la vista que necesitan para el desempeño de su trabajo, así como poder entender fácilmente lo que se dice en dicha grabación gracias a la transcripción de voz a texto. En caso contrario, tendrían que tener sus equipos informáticos encendidos más tiempo para realizar la búsqueda de forma manual y también necesitarían emplear más tiempo en entender lo que se dice en la grabación con el consiguiente gasto de energía eléctrica que ello supone.

2.3 Descripción y Funcionalidades

Textualización de Grabaciones, entrenado con lenguaje jurídico, permite, de manera diferida, procesar los vídeos grabados en las vistas y declaraciones judiciales para extraer en formato de texto la información que se alberga en ellos. Esta información se anexa al vídeo original, permitiendo la búsqueda y navegación a través del video seleccionado, a través del visor de expedientes judiciales Horus o con múltiples reproductores.

Esta solución ofrece a los usuarios las siguientes **funcionalidades**:

- **Capacidad de aprendizaje por reentrenamiento:** la Textualización de Grabaciones es un sistema de inteligencia artificial dotado de capacidad de aprendizaje, denominado Machine Learning. Ello permite nuevos procesos de reentrenamiento para obtener sucesivas mejoras en la fiabilidad de las textualizaciones.
- **Búsqueda:** permite añadir búsquedas de texto libre sobre el contenido del vídeo con el fin de ir a momentos determinados en los que se dijo la palabra que se quiere localizar dentro de la grabación.
- **Descarga:** permite obtener el registro de la textualización de la vista en un documento de texto editable.
- **Consulta de los intervinientes en los diálogos:** el sistema permite identificar a los intervinientes que aparecen en el vídeo y posibilita la localización de los momentos determinados en los que un interviniente dice una determinada palabra. Esta identificación se ha demostrado poco eficiente por los sistemas de Inteligencia Artificial de esta herramienta y deberá ser completado por marcas del personal de auxilio o, en un futuro, mejorando las instalaciones de las Salas de Vista con separación de señales acústicas en función del micrófono de grabación.

- Visualización de la línea de tiempo de la grabación:** muestra una barra con la línea de tiempo pudiendo navegar ágilmente por las declaraciones del interviniente que se haya seleccionado. Además, también permite buscar palabras clave para localizarlas en esta línea de tiempo del vídeo.
- Mostrar y/o ocultar marcas y textualizaciones:** nos permite visualizar, ocultar y/o abrir una nueva ventana las marcas establecidas en la grabación, por ejemplo, visualizar las marcas únicamente de un interviniente.
- Crear etiquetas:** el sistema permite generar etiquetas asociadas a momentos concretos de la grabación. Al pulsar sobre una de las etiquetas, aparecerán tantas marcas como veces se haya asociado la misma durante la grabación. Al pulsar sobre una de las marcas, se reproducirá el vídeo en el instante que aparece ese término.

La siguiente imagen muestra un ejemplo real de como los usuarios visualizan el sistema de Textualización de Grabaciones:



Imagen real del sistema de Textualización de Vistas

2.4 Fases del proyecto

Para llevar a cabo la elección del sistema más idóneo se evaluaron diferentes soluciones tecnológicas de textualización de vídeos mediante sistemas de Inteligencia Artificial.

Realizado un primer análisis, se consideraron tres soluciones para realizar las pruebas que permitieran seleccionar la tecnología más adecuada a las necesidades demandadas en el ámbito de la Justicia. En concreto se valoró la precisión (cálculo mediante un algoritmo que alinea las palabras y compara su coincidencia), tiempo de ejecución (cálculo según tiempos de logs y fechas de los ficheros de subtítulos), estabilidad, identificación y detección de hablantes e integración con otras aplicaciones del Ministerio de Justicia.

Teniendo en cuenta los aspectos anteriores, se seleccionó la opción que por fiabilidad, implicación y pruebas realizadas obtuvo los mejores resultados. Esta solución se basa en un

sistema de red neuronal y procesamiento fonético propio, y se ha integrado con la herramienta de visionado de expedientes judiciales electrónicos Horus.

El proceso de implantación se está realizando de forma progresiva en todos los órdenes jurisdiccionales y todas las instancias del ámbito territorial en los que es competente el Ministerio de Justicia, incluyendo las Fiscalías (CCAA de Castilla la Mancha, Castilla y León, Islas Baleares, Murcia, Extremadura, ciudades autónomas Ceuta y Melilla y órganos centrales).

Textualización de Grabaciones es una realidad en 17 provincias y en las dos ciudades autónomas (Cuenca, Ávila, Soria, Palencia, Segovia, Guadalajara, Ciudad Real, Burgos, Valladolid, Albacete, Toledo, León, Salamanca, Ceuta, Melilla, Zamora, Cáceres, Badajoz y Murcia, así como los órganos centrales) con la previsión de finalizar el despliegue en el último territorio (Baleares), **en octubre de 2021**. La muestra del calado de este proyecto son los datos que arroja. Hasta el 31 de julio de 2021 **se han textualizado con éxito 96.329 grabaciones de vistas judiciales**.



Imagen mapa España donde se encuentra disponible la Textualización de Vistas

Además, las comunidades autónomas con competencias transferidas en materia de justicia, en el ámbito del Comité Técnico Estatal de la Administración Judicial Electrónica (CTEAJE), han mostrado su interés en incorporar esta herramienta a sus sistemas de grabación. Al respecto se están realizando actuaciones con Madrid, Cataluña, Canarias, Valencia, Galicia, Asturias y La Rioja. Destacar que **el 11 de marzo se inició una experiencia piloto en la**

Comunidad de Madrid con éxito, habiéndose textualizado a fecha 31 de mayo 282 grabaciones. Y se está trabajando en otros pilotos en las CCAA de Canarias, Galicia.

Textualización de Grabaciones es un proyecto puntero dentro de la Administración de Justicia, cuyas prestaciones y capacidades se irán ampliando y perfeccionando a medida que se vaya extendiendo su uso en casos reales y se detecten nuevas necesidades y mejoras posibles. En este sentido, se está trabajando para que la solución sea capaz de textualizar las intervenciones realizadas en las lenguas cooficiales de España y otros idiomas.