

Especificaciones técnicas

SUBTITULADO AUTOMÁTICO DE PROGRAMAS DE INFORMATIVOS TERRITORIALES



1 Introducción

En Radio Televisión Española tenemos desde hace dos años la subtítulos automática de los informativos territoriales que se realizan en dos idiomas. En estos momentos se dispone de este servicio en los Centros Territoriales (CCTT) de Valencia, Islas Baleares, Galicia, País Vasco y Navarra y en el CPP de Cataluña para programas de corte informativo en catalán.

El presente expediente contempla la contratación de un servicio de subtítulos automático para los informativos territoriales de los 5 CCTT antes descritos y también para programas de corte informativo del Centro de Producción de Cataluña que se realizan simultáneamente en catalán y castellano. Completando este servicio también queremos ampliarlo con un subtítulo en diferido que sincronice y mejore el subtítulo anterior.

RTVE desea contratar este servicio de forma que esté basado en aprendizaje automático (Machine Learning), nuevas tecnologías de procesamiento de lenguaje natural en tiempo real, reconocimiento del habla y corrección, y que esté apoyado en procesos en la cloud (nube) pública o privada.

2 Objeto del contrato

El objeto del contrato del lote 1 constará de dos servicios, el primero es mejorar la accesibilidad de los Informativos Territoriales bilingües mediante el subtítulo automático y en directo de los programas emitidos por los Centros Territoriales de la Comunidad Valenciana, Islas Baleares, Galicia, País Vasco y Navarra para procesar indistintamente contenidos bilingües en castellano y valenciano, castellano y mallorquín, castellano y gallego y castellano y euskera de forma que, en cada momento, el subtítulo generado sea en la misma lengua que se está hablando. El segundo servicio consistirá en la sincronización y mejora del subtítulo para poder incluirlo junto con el vídeo correspondiente en la página web de RTVE y la generación de metadatos para la incorporación de éstos al archivo de RTVE.

Un segundo lote contempla la contratación de los dos servicios anteriormente descritos de subtítulo automático, en este caso en catalán, para programas de corte informativo producidos en el Centro de Producción de Programas de RTVE en Cataluña, procesando indistintamente el contenido en catalán y castellano.

Por último, el proyecto incluye un tercer lote para el análisis de calidad de las soluciones contratadas que, además de comprobar el buen funcionamiento de las mismas y que los niveles de calidad se mantengan y mejoren con el tiempo, sirva para retroalimentar los sistemas ayudando a su aprendizaje y mejorando la calidad de los servicios.

Dado el propósito del Lote 3, el adjudicatario de este lote no podrá ser el mismo que para los Lotes 1 y/o 2 ni tener relación comercial o de otro tipo con la/s empresas que realicen el servicio de subtítulo.

3 Lote 1. Subtítulo automático bilingüe para CCTT

3.1 Servicio subtítulo en directo

El servicio consistirá en llevar a cabo el subtítulo automático de los Informativos territoriales que se emiten en las 5 comunidades autónomas de Valencia, Islas Baleares, Galicia, País Vasco y Navarra.

El ofertante deberá prestar sus servicios todos los días laborables de lunes a viernes, durante las horas en las que se emitan informativos desde los Centros objeto del contrato, Comunidad Valenciana, Islas Baleares, Galicia, Navarra y País Vasco. Serán dos informativos por día, coincidentes en horario en todos los Centros y con una duración máxima de 45 minutos en total, cinco días a la semana durante 12 meses de servicio, prorrogables otros 12 más.

Las duraciones de cada informativo y sus horas de emisión dependerán de la programación de TVE en cada momento. A modo de ejemplo, en estos momentos se emite un primer informativo a las 14:00h con una duración de 12 minutos y el segundo, de 25 minutos de duración, comienza a las 15:50h. Durante el contrato, la distribución de minutos entre los dos informativos y las horas de emisión de estos podrán variar según las necesidades de RTVE, debiendo, en ese caso el adjudicatario de este lote, ajustar el servicio a los nuevos horarios.

Para ello, RTVE entregará al adjudicatario de este lote la señal de audio de los programas informativos en el propio Centro Territorial donde se produzcan. El sistema de subtitulación contratado deberá realizar el procesamiento necesario para la generación de los subtítulos de forma automática, en tiempo real y con reconocimiento del idioma hablado. Todo el procesado se llevará a cabo en la cloud. Los subtítulos generados para todos los Centros se entregarán en el CPP de Torrespaña (Madrid) para su inserción en la emisión del programa.

El adjudicatario dará soporte con sus propios recursos técnicos y humanos. Como el servicio está funcionando actualmente, es necesario garantizar la continuidad de éste, por lo que el nuevo sistema deberá estar en funcionamiento en todos los CCTT el 1 de agosto de 2024. El adjudicatario de este lote dispondrá de un plazo de un mes desde la firma del contrato para instalar en los cinco CCTT objeto de este expediente y en Torrespaña los equipos necesarios para dar el servicio, realizar configuraciones y ajustes en el sistema y probar que el funcionamiento global de la solución sea el que se espera.

Si en el periodo de vigencia del contrato se detectara, por parte de los usuarios de CRTVE o en los controles de calidad llevados a cabo por el adjudicatario del lote 3, fallos reiterados en la transcripción de alguna palabra concreta referida a personalidades, lugares geográficos o entidades que afecten de forma sustancial a la comprensión de lo hablado, el adjudicatario de este lote se comprometerá a realizar las acciones oportunas para la rápida subsanación de estos fallos. Si no se consigue resolver, RTVE lo considerará un fallo endémico, y provocará la suspensión temporal del contrato.

Por cada día en que el servicio esté parado, el adjudicatario será penalizado con las cantidades que se enuncian en el Pliego de cláusulas administrativas particulares

3.1.1 Descripción de los requerimientos técnicos y de calidad

3.1.1.1 Características principales del sistema de subtulado

RTVE proporcionará al adjudicatario un acceso al audio de las emisiones sin retardo en cada Centro donde se produce la señal. Para acceder a su sistema, el adjudicatario deberá instalar el equipamiento y software que fuera necesario y asumir los costes relacionados con dicha instalación.

RTVE realizará la entrega de audio digital embebido en la señal HD-SDI en banda base a través del interfaz **audio digital AES3**.

El sistema ofertado deberá cumplir las siguientes características:

- trabajar solo sobre el sonido en directo del programa, ya que no contará con ayudas de sistemas de información previa, como guiones de las noticias, escaletas o minutados.
- trabajar con tecnología cloud, con posibilidad de ser instalado en cloud pública o en cloud privada.

- disponer de elementos potentes y ágiles de procesado, para minimizar retardos entre el vídeo y el subtítulo.
- poder incorporar un listado de palabras indebidas a evitar donde RTVE decidirá por qué palabras se sustituyen.
- incorporar una lista de sustituciones para la presentación en el subtítulo de términos concretos con la que poder interactuar tanto el adjudicatario como RTVE para correcciones específicas.
- proporcionar un fichero de texto con la transcripción literal de la voz a texto en formato XML.
- cumplir la norma sobre subtítulo UNE 153010/2012 o cualquiera que la sustituya y/o actualice, así como del código de buenas prácticas del CESyA (Centro Español del Subtitulado y la Audiodescripción) para los subtítulos generados.
- ser bilingüe, distinguiendo entre los idiomas castellano y valenciano, castellano y mallorquín, castellano y gallego, castellano y euskera, generando los subtítulos en el mismo idioma en el que se esté hablando.
- incorporar un reconocedor de sintonías que facilite el reconocimiento de cambio de idioma en el caso del CT de Navarra.
- entrega en Torrespaña del subtítulo en directo bilingüe conformado en una trama DVB TSoIP cumpliendo con lo especificado en la norma TR101290.
- la entrega de la trama TSoIP será en formato multicast con una IP y puerto acordado entre ambas partes.
- los flujos de subtítulos se entregarán en una única trama de transporte TSoIP. Cada componente de subtítulos DVB-Sub (mapa de bits según la EN300743) estará asociado a un servicio de la trama de transporte con su correspondiente PCR. El nombre de cada uno de los servicios que se indicarán en la tabla SDT del TS lo establecerá RTVE.
- la trama de transporte únicamente estará compuesta por las tablas PSI/SI y los componentes que sean estrictamente necesarios. No se admitirán tramas de transporte con tablas o componentes adicionales a los previstos.
- la resolución de los subtítulos y todo lo referente a su formato (tipo de letra, tamaño, caracteres por línea, etc.) será configurable y definido por RTVE para que sea homogéneo con el resto de los subtítulos emitidos por sus cadenas. Las especificaciones podrán variar si RTVE lo considera necesario.
- la tasa binaria máxima del PID de subtítulos será configurable a determinar por RTVE.
- la oferta incluirá el servicio y soporte del Sistema de subtítulo por doce meses a contar desde que el sistema cumpla los requisitos solicitados en este expediente y haya sido validado por RTVE.
- asumir la entrega de los trabajos en el/los nuevo/s formato/s de subtítulo en el caso de que CRTVE cambiara el sistema actual.
- disponer de sistemas de alarma en caso de que se produzca algún fallo. En caso de mal funcionamiento, la alarma no sólo debe avisar al adjudicatario sino a RTVE a través de un correo electrónico, además de poder desconectar el equipo de subtítulo y garantizar la recuperación del servicio con un plazo máximo de 2 días. Para ello, el adjudicatario generará un grupo de Microsoft Teams con acceso para RTVE donde se guarden todas las alarmas generadas.

3.1.1.2 Otros requerimientos del servicio

- El nivel de aciertos será igual o superior al de los programas subtítulos en directo, de forma no automática.
- El retardo de aparición de los subtítulos en la TDT será como máximo de 8 segundos con respecto a la señal sonora de televisión.
- Cuando se produzca un cambio de idioma el tiempo máximo de la ventana para que se detecte y reproduzca el subtítulo en el nuevo idioma será de 5 segundos y no deberá penalizar el retardo máximo definido en el punto anterior.
- Dada la criticidad del retraso entre el audio y los correspondientes subtítulos en la emisión de la señal, el adjudicatario deberá proporcionar un procedimiento que permita medir el tiempo transcurrido desde la entrega de la señal de audio en origen por parte de RTVE hasta salida del equipo del adjudicatario, instalado en Torrespaña, donde se recibe el streaming del subtítulo.

- La densidad de subtítulo deberá ser superior al 80% de todo lo hablado en el programa.
- Los subtítulos tendrán un tiempo de permanencia en pantalla mínimo de 1 segundo y máximo de 7 segundos.
- Los subtítulos deberán ubicarse en la posición de la pantalla que menos afecte al contenido de la imagen. Como norma general irán en la zona inferior de la pantalla, en formato de dos líneas, respetando siempre la norma UNE 153010 o cualquiera que la sustituya y/o actualice. En los programas con 'faldones informativos' la/s línea/s de subtítulo se colocará por encima de ellos para no sobrescribirlos.
- Para medir el rendimiento del sistema, se utilizará la tasa de error de palabra (WER). La tasa de error de palabras se puede calcular como:

$$\text{WER (\%)} = \frac{I + B + S}{N} 100$$

dónde:

- I son los errores de inserción, cuando el sistema da como salida una palabra que no existe en la transcripción de referencia.
 - B son los errores de borrado, cuando el sistema no da como salida una palabra que existe en la transcripción de referencia.
 - S son los errores de sustitución, cuando el sistema da como salida una palabra distinta a la que existe en la transcripción de referencia.
 - N es el número de palabras en la transcripción de referencia.
- Se espera que, en el momento de la adjudicación, para el subtítulo en castellano, el sistema, (sin entrenamiento previo con nuestros programas) tenga una tasa de error WER menor que el 10%.
 - Para los subtítulos en el resto de lenguas, se espera que, en el momento de la adjudicación, el sistema (sin entrenamiento previo con nuestros programas) tenga una tasa de error WER menor que:
 - el 20% para el valenciano, mallorquín y gallego.
 - El 10% para el euskera.
- Para el cálculo de las medidas de tasa de error WER no se tendrán en cuenta los errores cometidos durante la ventana del cambio de idioma ni cuando haya errores del reconocedor del idioma.
- El no cumplimiento de los baremos de calidad que se establecen en el presente pliego, en cuanto a porcentaje de aciertos, conllevará las penalizaciones estipuladas en el pliego de cláusulas administrativas particulares.
 - De la misma manera, si los niveles de calidad disminuyeran durante los 12 meses de servicio se aplicarán las penalidades establecidas en el pliego de cláusulas administrativas particulares.

3.1.1.3 Conexiones en los CCTT y Torrespaña

RTVE proporcionará acceso a Internet con velocidad suficiente para la gestión del servicio en cada Centro Territorial.

La instalación y el mantenimiento del software, hardware y medios de conexión será a cargo de la empresa adjudicataria, que velará por su funcionamiento correcto y su utilización exclusiva para los fines que se le encomiendan.

3.1.1.4 Equipos en las instalaciones de RTVE

RTVE facilitará al adjudicatario el espacio para albergar los equipos necesarios tanto en los Centros Territoriales como en Torrespaña, en salas debidamente refrigeradas y con alimentación. El equipamiento instalado en cada sede no podrá superar el tamaño disponible en dos unidades de rack.

Además, en el caso de que este equipamiento fuera de tipo informático, cada equipo deberá poseer:

- Licencia actualizada del sistema operativo que esté utilizando.

- Sistema operativo que posea un sistema vigente de actualizaciones de seguridad.
- Gestión de actualizaciones de seguridad aplicables al menos una vez al mes.
- Software antivirus con licencia vigente y actualizado periódicamente, no menos de una vez por semana.
- Todo el software instalado será de tipo libre o con su correspondiente licencia legal.

Las averías de hardware y/o software en los equipos del adjudicatario deberán ser cubiertas en cualquier caso por la propia empresa adjudicataria y a la máxima celeridad posible. No se podrán justificar retrasos, fallos o errores en la entrega por este motivo.

RTVE se reserva el derecho de auditar los equipos del ofertante, presentes en sus instalaciones con el fin de verificar el cumplimiento de estos apartados y, sobre todo, la seguridad informática. Aun así, tanto la gestión de las licencias, como los gastos derivados de ellas, son responsabilidad exclusiva del adjudicatario.

3.2 Servicio de sincronización de los subtítulos y generación de metadatos

Se propone la realización de una prueba de concepto (PoC) que consistirá en la corrección y sincronización de los subtítulos obtenidos del informativo en directo, para que puedan ser subidos a la página web de RTVE. También se realizará la generación de los metadatos del audio recibido y previo a su análisis básico, el adjudicatario deberá transcribir el audio en el idioma en el que se hable y traducirlo al español. Sobre los archivos de audio se requerirán los siguientes procesos:

- Transcripción de audio a texto.
- Puntuación y capitalización del texto transcrito.
- Segmentación por hablantes (diarización).
- Reconocimiento de entidades nombradas (NERC) o extracción de entidades. Entendiendo como tal la tarea que supone identificar y clasificar en categorías predefinidas como personas, organizaciones y lugares.
- Extracción de palabras clave.
- Clasificación automática de contenido, entendida como la asignación de una o varias categorías en función de los conceptos expresados en el texto resultante de la transcripción automática.
- Resúmenes automáticos.

Se deberá recibir, para cada audio analizado, un archivo XML con el resultado del análisis de la media, en el formato y las características definidas por RTVE para su ingesta en el Gestor Documental, aceptándose que este formato puede variar en función de nuevos requerimientos mientras dure el contrato, con la nomenclatura adecuada que permita asociar al fichero de media al que corresponda y con el código de tiempo asociado al momento en que aparece cada elemento detectado en el audio original.

Si en el análisis de los resultados de esta PoC por parte de los usuarios de CRTVE o en los controles de calidad llevados a cabo por el adjudicatario del lote 3, se detectara que el WER del subtítulo no mejora un 3% respecto al WER del subtítulo en directo exigido en el punto 3.1, RTVE no continuará con este servicio y provocará la suspensión de esta parte del contrato.

4 Lote 2. Subtitulado automático CPP Cataluña

4.1 Servicio subtitulado en directo

El servicio consistirá en llevar a cabo el subtitulado automático de alguno/os de los programas de corte Informativo que se emiten desde el CPP de RTVE en Cataluña.

El adjudicatario deberá prestar sus servicios todos los días de lunes a sábado, durante las horas en las que se emitan los programas de corte informativo del CPP de Cataluña seleccionados. Los programas concretos a subtítular en Cataluña le serán comunicados al adjudicatario a la firma del contrato; serán bilingües, distinguiendo entre los idiomas castellano y catalán, generando los subtítulos en el mismo idioma en el que se esté hablando y tendrán una duración de 5 horas en total a la semana. Durante el contrato, los programas y las horas de emisión podrán variar según las necesidades de RTVE, debiendo, en ese caso, ajustar el servicio a los nuevos programas y horarios.

El adjudicatario tendrá un periodo de un mes desde la fecha de firma del contrato para la instalación y configuración de sus equipos. Pasados este periodo comenzarán los 12 meses de servicio estipulados previamente.

RTVE entregará al adjudicatario de este lote la señal de audio de los programas a subtítular en el propio CPP. El sistema de subtitulación contratado deberá realizar el procesamiento necesario para la generación de los subtítulos de forma automática y en tiempo real. Todo el procesado se llevará a cabo en la cloud. Los subtítulos generados se entregarán para su inserción en la señal de TV en el mismo CPP de Cataluña, utilizando para ello el sistema de subtulado de RTVE de la marca Anglatecnic. Será responsabilidad del adjudicatario de este lote 2 la adquisición e instalación de cualquier elemento necesario para la integración en dicho sistema.

Si en el periodo de vigencia del contrato se detectara, por parte de los usuarios de CRTVE o en los controles de calidad llevados a cabo por el adjudicatario del lote 3, fallos reiterados en la transcripción de alguna palabra concreta referida a personalidades, lugares geográficos o entidades que afecten de forma sustancial a la comprensión de lo hablado, el adjudicatario de este lote se comprometerá a realizar las acciones oportunas para la rápida subsanación de estos fallos. Si no se consigue resolver, RTVE lo considerará un fallo endémico, y provocará la suspensión temporal del contrato.

Por cada día en que el servicio esté parado, el adjudicatario será penalizado con las cantidades que se enuncian en el pliego de cláusulas administrativas particulares. Descripción de los requerimientos técnicos y de calidad

4.1.1.1 Características principales del sistema de subtulado

RTVE proporcionará al adjudicatario un acceso al audio de las emisiones sin retardo desde el Centro de Producción de Programas de Cataluña donde se produce la señal. Para acceder a su sistema, el adjudicatario deberá instalar el equipamiento y software que fuera necesario y asumir los costes relacionados con dicha instalación.

RTVE realizará la entrega de audio digital embebido en la señal HD-SDI en banda base a través del interfaz **audio digital AES3**.

El sistema ofertado deberá cumplir las siguientes características:

- trabajar solo sobre el sonido en directo del programa, ya que no contará con ayudas de sistemas de información previa, como guiones de las noticias, escaletas o minutos.
- trabajar con tecnología cloud, con posibilidad de ser instalado en cloud pública o en cloud privada.
- generar subtítulos en castellano y catalán.
- disponer de elementos potentes y ágiles de procesado, para minimizar retardos entre el vídeo y el subtulado.
- poder incorporar un listado de palabras indebidas a evitar donde RTVE decidirá por qué palabras se sustituyen.
- incorporar una lista de sustituciones para la presentación en el subtítulo de términos concretos con la que poder interactuar tanto el adjudicatario como RTVE para correcciones específicas.

- proporcionar un fichero de texto con la transcripción literal de la voz a texto en formato XML.
- cumplir la norma sobre subtulado UNE 153010/2012 o cualquiera que la sustituya y/o actualice, así como del código de buenas prácticas del CESyA (Centro Español del Subtulado y la Audiodescripción) para los subtítulos generados.
- los ficheros con el subtulado se entregarán en el CPP de Cataluña y se deberá hacer en IP en formato Newfor o en ITN.
- la oferta incluirá el servicio y soporte del Sistema de subtulado por doce meses a contar desde que el sistema cumpla los requisitos solicitados en este expediente y haya sido validado por RTVE.
- asumir la entrega de los trabajos en el/los nuevo/s formato/s de subtulado en el caso de que CRTVE cambiara el sistema actual.
- disponer de sistemas de alarma en caso de que se produzca algún fallo. En caso de mal funcionamiento, la alarma no sólo debe avisar al adjudicatario sino a RTVE a través de un correo electrónico, además de poder desconectar el equipo de subtulado y garantizar la recuperación del servicio con un plazo máximo de 2 días. Para ello, el adjudicatario generará un grupo de Microsoft Teams con acceso para RTVE donde se guarden todas las alarmas generadas.

4.1.1.2 Otros requerimientos del servicio

- El nivel de aciertos será igual o superior al de los programas subtulados en directo, de forma no automática.
- El retardo de aparición de los subtítulos en la TDT será como máximo de 8 segundos con respecto a la señal sonora de televisión.
- Dada la criticidad del retraso entre el audio y los correspondientes subtítulos en la emisión de la señal, el adjudicatario deberá proporcionar un procedimiento que permita medir el tiempo transcurrido desde la entrega de la señal de audio en origen por parte de RTVE hasta salida del equipo del adjudicatario, instalado en el CPP de Cataluña, donde se recibe el streaming del subtulado.
- La densidad de subtulado deberá ser superior al 80% de todo lo hablado en el programa.
- Los subtítulos tendrán un tiempo de permanencia en pantalla mínimo de 1 segundo y máximo de 7 segundos.
- Los subtítulos deberán ubicarse en la posición de la pantalla que menos afecte al contenido de la imagen. Como norma general irán en la zona inferior de la pantalla, en formato de dos líneas, respetando siempre la norma UNE 153010 o cualquiera que la sustituya y/o actualice. En los programas con 'faldones informativos' la/s línea/s de subtulado se colocará por encima de ellos para no sobrescribirlos.
- Para medir el rendimiento del sistema, se utilizará la tasa de error de palabra (WER). La tasa de error de palabras se puede calcular como:

$$WER (\%) = \frac{I + B + S}{N} 100$$

dónde:

- I son los errores de inserción, cuando el sistema da como salida una palabra que no existe en la transcripción de referencia.
 - B son los errores de borrado, cuando el sistema no da como salida una palabra que existe en la transcripción de referencia.
 - S son los errores de sustitución, cuando el sistema da como salida una palabra distinta a la que existe en la transcripción de referencia.
 - N es el número de palabras en la transcripción de referencia.
- Se espera que, en el momento de la adjudicación, para el subtulado en castellano, el sistema, (sin entrenamiento previo con nuestros programas) tenga una tasa de error WER menor que el 10%.

- Para los subtítulos en catalán se espera que, en el momento de la adjudicación, el sistema (sin entrenamiento previo con nuestros programas) tenga una tasa de error WER menor que el 15%. Para el cálculo de las medidas de tasa de error WER no se tendrán en cuenta los errores cometidos durante la ventana del cambio de idioma ni cuando haya errores del reconocedor del idioma.
- El no cumplimiento de los baremos de calidad que se establecen en el presente pliego, en cuanto a porcentaje de aciertos, conllevará las penalizaciones estipuladas en el Pliego de Condiciones Generales.
- De la misma manera, si los niveles de calidad disminuyeran durante los 12 meses de servicio se aplicarán las penalidades establecidas en el pliego de cláusulas administrativas particulares.

4.1.1.3 Conexiones en el CPP de RTVE en Cataluña

RTVE proporcionará acceso a Internet con velocidad suficiente para la gestión del servicio en el CPP de Cataluña.

La instalación y el mantenimiento del software, hardware y medios de conexión será a cargo de la empresa adjudicataria, que velará por su funcionamiento correcto y su utilización exclusiva para los fines que se le encomiendan.

4.1.1.4 Equipos en las instalaciones de RTVE

RTVE facilitará al adjudicatario el espacio para albergar los equipos en el CPP Cataluña, en salas debidamente refrigeradas y con alimentación. El equipamiento instalado, no podrá superar el tamaño disponible en dos unidades de rack.

Además, en el caso de que este equipamiento fuera de tipo informático, cada equipo deberá poseer:

- Licencia actualizada del sistema operativo que esté utilizando.
- Sistema operativo que posea un sistema vigente de actualizaciones de seguridad.
- Gestión de actualizaciones de seguridad aplicables al menos una vez al mes.
- Software antivirus con licencia vigente y actualizado periódicamente, no menos de una vez por semana.
- Todo el software instalado será de tipo libre o con su correspondiente licencia legal.

Las averías de hardware y/o software en los equipos del adjudicatario deberán cubrirse por la empresa adjudicataria y con la máxima celeridad posible. No se podrán justificar retrasos, fallos o errores en la entrega por este motivo.

RTVE se reserva el derecho de auditar los equipos del ofertante, presentes en sus instalaciones con el fin de verificar el cumplimiento de estos apartados y, sobre todo, la seguridad informática. Aun así, tanto la gestión de las licencias, como los gastos derivados de ellas, son responsabilidad exclusiva del adjudicatario.

4.2 Servicio de sincronización de los subtítulos y generación de metadatos

Se propone la realización de una prueba de concepto (PoC) que consistirá en la corrección y sincronización de los subtítulos obtenidos del informativo en directo, para que puedan ser subidos a la página web de RTVE. También se realizará la generación de los metadatos del audio recibido y previo a su análisis básico, el adjudicatario deberá transcribir el audio en el idioma en el que se hable y la parte que esté en catalán traducirlo al español. Sobre los archivos de audio se requerirán los siguientes procesos:

- Transcripción de audio a texto.
- Puntuación y capitalización del texto transcrito.

- Segmentación por hablantes (diarización).
- Reconocimiento de entidades nombradas (NERC) o extracción de entidades. Entendiendo como tal la tarea que supone identificar y clasificar en categorías predefinidas como personas, organizaciones y lugares.
- Extracción de palabras clave.
- Clasificación automática de contenido, entendida como la asignación de una o varias categorías en función de los conceptos expresados en el texto resultante de la transcripción automática.
- Resúmenes automáticos.

Se deberá recibir, para cada audio analizado, un archivo XML con el resultado del análisis de la media, en el formato y las características definidas por RTVE para su ingesta en el Gestor Documental, aceptándose que este formato puede variar en función de nuevos requerimientos mientras dure el contrato, con la nomenclatura adecuada que permita asociar al fichero de media al que corresponda y con el código de tiempo asociado al momento en que aparece cada elemento detectado en el audio original.

Si en el periodo de vigencia del contrato se detectara, por parte de los usuarios de CRTVE o en los controles de calidad llevados a cabo por el adjudicatario del lote 3, que el WER del subtulado no mejora al menos un 3% respecto al WER del subtulado en directo exigido en el punto 4.1, RTVE no continuará con este servicio y provocará la suspensión de esta parte del contrato.

5 Lote 3. Control de calidad y retroalimentación para la mejora del sistema

El adjudicatario de este lote deberá realizar el control de calidad del subtulado generado por los adjudicatarios de los Lotes 1 y 2 de este expediente. El servicio descrito a continuación es común a todos los Centros.

5.1 Muestras

CRTVE proporcionará al adjudicatario una muestra semanal por cada uno de los Centros, de cinco minutos de duración, con el contenido audiovisual y el subtulado correspondientes a una parte de los programas informativos del Centro que se tiene que analizar, en total se analizarán muestras de 6 Centros. El adjudicatario deberá disponer de personal con conocimientos demostrables, con un nivel mínimo de B2, en las siguientes lenguas: castellano, valenciano, balear, gallego, euskera y catalán.

5.2 Metodología - Evaluación de transcripción

Se realizarán controles de calidad tanto del funcionamiento del reconocedor de idioma como de la transcripción. Por cada muestra se considerarán los siguientes parámetros de calidad:

- **WER.** El WER (acrónimo del inglés Word Error Rate) es un cálculo estandarizado que mide la fidelidad con que un sistema automático transcribe una locución. Consiste en contabilizar errores en términos de palabras incorrectas (Sustitución), palabras ausentes (Inserción) y palabras sobrantes (Borrados). El total de errores obtenidos se pone en relación con el número total de palabras (N) presentes en una transcripción 100% correcta aplicando la siguiente fórmula:

$$\text{WER} = \frac{S+I+B}{N}$$

- o No se tendrán en cuenta para el cálculo de WER aquellas frases en las que el reconocedor de idioma haya tenido un mal funcionamiento y como consecuencia de ello no estén transcritas en el idioma que les correspondería.

- o No se tendrán en cuenta las transcripciones correspondientes a los cinco segundos siguientes a un cambio de idioma.
- **Precisión en la transcripción.** En complemento al valor obtenido de WER, se han establecido una serie de parámetros adicionales que se consideran con impacto en la calidad final del resultado. Estos valores son:
 - o Puntuación incorrecta (P).
 - o Formato incorrecto de números (F).

$$\text{Precisión transcripción} = \frac{N - (S + I + B + P + F)}{N}$$

Quedarían fuera del análisis descrito para ambos parámetros:

- aquellos fragmentos que incluyan canciones o conversaciones en segundo plano.
- transcripciones de hablantes en otra lengua distinta de las dos lenguas habladas en cada Centro Territorial.
- **Precisión del reconocedor de idioma.** Se trata del porcentaje de acierto del sistema que detecta en qué idioma se está hablando en cada momento. Es la relación entre el número de cambios erróneos y el número total de cambios, calculado mediante la siguiente fórmula:
 - o Cambios erróneos (E): entendidos como aquellas situaciones en las que el reconocedor detecta un cambio de idioma que no se ha producido realmente.
 - o Cambios ausentes (A): son aquellas ocasiones en las que sí se ha producido un cambio de idioma, pero el reconocedor continúa detectando el idioma que se hablaba previamente.
 - o Número total de cambios (N): es el número total de cambios de idioma que se producen realmente en la muestra.

$$\text{Precisión del reconocedor (\%)} = \frac{E + A}{N} \cdot 100$$

Quedarían fuera del análisis descrito para estos tres parámetros:

- aquellos fragmentos que incluyan canciones o conversaciones en segundo plano,
- transcripciones de hablantes en otras lenguas distintas de las dos lenguas habladas en cada Centro Territorial.
- **Otras medidas.** Se han identificado algunos parámetros adicionales que, aunque no computarán en los cálculos descritos de WER y Precisión, deberán tenerse en consideración y reflejarlos por tanto en los análisis e informes.
 - o Cambios de hablante. Se marcarán de forma visual los cambios que se vayan sucediendo en la muestra. Se deberá aportar un porcentaje de cambios de hablante detectados por el sistema de subtítulo.
 - o Canciones.
 - o Conversaciones en segundo plano.

5.3 Metodología – Medida de Sincronía

Se realizará un seguimiento del retardo que se produce en la aparición de subtítulos en pantalla respecto al audio. Este dato se presentará en segundos con dos decimales para aumentar la precisión.

La medición del retardo en la aparición de subtítulos se realizará según lo especificado en el Anexo C de la Norma UNE 153010:2012:

“Para medir el retardo de los subtítulos de un programa en concreto (en nuestro caso cada muestra) conviene analizar una muestra aleatoria y representativa del total de los subtítulos de dicho programa. Aunque se recomienda que la muestra sea aleatoria, conviene incluir subtítulos que pertenezcan a cambios de hablante o a intervenciones cortas y/o solapadas, que suelen conllevar mayor retraso.

La muestra se divide en dos grupos de tamaños diferentes. El primer grupo de subtítulos se usará para medir el retardo en la aparición de la primera palabra de los subtítulos (retardo de inicio), mientras que el segundo grupo servirá para medir el retardo en la aparición de la última palabra (retardo de fin).

La proporción entre los dos grupos de subtítulos de la muestra es de 2 a 1. De cada 3 subtítulos elegidos en la muestra, 2 se usarán para medir el retardo de inicio y 1 se usará para medir el retardo de fin. Por lo tanto, si la muestra es de un 15%, habrá un 10% de subtítulos para medir retardo de inicio y un 5% para medir retardo de fin.

El resultado final de retardo se obtiene haciendo una media entre los dos retardos parciales:

$$\text{retardo} = \frac{\text{retardo de inicio} + \text{retardo de fin}}{2}$$

Para determinar el retardo de inicio de un subtítulo i se calcula la diferencia entre el momento de aparición de la primera palabra en el subtítulo (t_{pps}) y el momento en el que se dice esa palabra en el discurso original (t_{ppo}):

$$\text{retardo de inicio}_i = t_{pps} - t_{ppo}$$

El retardo de inicio total se calcula como la media de los retardos de inicio de los subtítulos elegidos en el primer 10% (considerando que ese 10% de subtítulos tiene n elementos):

$$\text{retardo de inicio} = \frac{\sum_i^n \text{retardo de inicio}_i}{n}$$

Para determinar el retardo de fin de un subtítulo j se calcula la diferencia entre el momento de aparición de la última palabra en el subtítulo (t_{ups}) y el momento en que se dice esa última palabra en el discurso original (t_{upo}):

$$\text{retardo de fin}_j = t_{ups} - t_{upo}$$

El retardo de fin total se calcula como la media de los retardos de fin de los subtítulos elegidos en el segundo 5% (considerando que ese 5% de subtítulos tiene m elementos):

$$\text{retardo de fin} = \frac{\sum_i^m \text{retardo de fin}_j}{m}$$

5.4 Documentación

Al finalizar cada semana el adjudicatario de este lote entregará a la CRTVE la siguiente documentación:

- **Excel de datos semanales.** En un formato a indicar por CRTVE una vez adjudicado el expediente, se presentará de forma incremental los resultados de la última semana analizada.
- **Un resumen gráfico semanal.** Unas gráficas donde se detalle la evolución del WER en cada centro territorial, así como el resto de parámetros de interés, que serán especificados una vez adjudicado el expediente en formato Power BI o similar.

- **Excel de fallos recurrentes.** Indicando qué tipos de fallos se repiten reiterativamente y en qué cantidad en un formato que se indicará una vez adjudicado el expediente.
- Un informe por cada una de las muestras analizadas con la siguiente estructura:
 - **Resumen de datos.** Un cuadro donde se refleje individualmente el recuento numérico de cada parámetro descrito, así como los valores generales obtenidos en WER, precisión y retardo de subtítulos.
 - **Análisis de la transcripción.** Un documento de texto con la transcripción sobre la que se marcarán distinguiblemente con un código de colores especificado por RTVE una vez adjudicado el expediente, los distintos errores descritos en los puntos anteriores.
 - **Comentarios / consideraciones.** Un análisis subjetivo acerca de las impresiones derivadas del análisis realizado indicando los errores más relevantes en función de la gravedad o la reiteración de los errores.

Por último, al término del proyecto se incluirá un informe final donde se incluya un resumen de lo analizado y se explique la evolución de la calidad.

5.5 Características del servicio

Dado que la actividad del Lote 3 está supeditada a la del Lote 1 y el Lote 2, en caso de que por alguna circunstancia se produzca una interrupción en su actividad, los trabajos recogidos en el presente Lote 3, además de las causas indicadas en la cláusula 18ª del Pliego de condiciones generales, se verán afectados de la siguiente manera:

- Paro inferior a una semana: conllevará una redistribución de las catas a analizar.
- Paro superior a una semana: la actividad del Lote 3 quedará interrumpida. En función de las circunstancias que originaran este paro, podría dar lugar a un cambio en la fecha de finalización del contrato o a la cancelación del mismo.
- La cancelación del contrato al adjudicatario del Lote 1 o del Lote 2 si no alcanza los niveles de calidad exigidos en la prueba de aceptación, cancelaría la parte correspondiente del contrato de este Lote 3.

5.6 Centros territoriales

En este punto se presenta el desglose por ítems y la descripción del servicio a prestar por cada uno.

Control calidad CT Valencia

Control de calidad en los términos descritos en el punto 5 del presente documento a realizar sobre el subtítulo de informativos territoriales en Valencia en los idiomas: castellano / valenciano.

Control calidad CT Galicia

Control de calidad en los términos descritos en el punto 5 del presente documento a realizar sobre el subtítulo de informativos territoriales en Galicia en los idiomas: castellano / gallego.

Control calidad CT Baleares

Control de calidad en los términos descritos en el punto 5 del presente documento a realizar sobre el subtítulo de informativos territoriales en Islas Baleares en los idiomas: castellano / balear.

Control calidad CT País Vasco

Control de calidad en los términos descritos en el punto 5 del presente documento a realizar sobre el subtítulo de informativos territoriales en País Vasco en los idiomas: castellano / euskera.

Control calidad CT Navarra

Control de calidad en los términos descritos en el punto 5 del presente documento a realizar sobre el subtítulo de informativos territoriales en Navarra en los idiomas: castellano / euskera.

Control calidad CPP Cataluña

Control de calidad en los términos descritos en el punto 5 del presente documento a realizar sobre el subtítulo de programas informativos en Cataluña en los idiomas: castellano / catalán.

6 Prueba previa a la adjudicación

A la vista de las ofertas recibidas y para su mejor selección, la CRTVE requerirá a los licitantes del Lote 1 y Lote 2, una demostración del sistema ofertado, en las instalaciones de CRTVE o en otro lugar que previamente se determine, sin ningún coste adicional, donde deberá quedar demostrada la idoneidad de la solución presentada.

La convocatoria se enviará a los oferentes a través de correo electrónico con un plazo mínimo de **cinco días** naturales hasta la fecha de celebración de la prueba y deberán responder a ella en el plazo de **tres días** laborables. Una vez cumplido este plazo, si el oferente no ha respondido, la oferta quedará desestimada. Para ello, los oferentes deberán indicar en su oferta dos cuentas de correo electrónico de las personas de contacto.

La prueba propuesta consistirá en lo siguiente:

Lote 1:

El día señalado para la prueba, CRTVE proporcionará a los oferentes cada hora 10 minutos de un informativo territorial de cada uno de los siguientes CCTT: Valencia, Baleares, Galicia y Navarra o País Vasco, dado que ambos comparten idioma, total 40 minutos con una duración total de la prueba de cuatro horas. Los oferentes deberán generar y enviar los subtítulos a CRTVE en formato SRT como máximo 15 minutos después de haber recibido cada archivo en los mismos términos que se describen para el servicio del Lote 1.

Lote 2:

CRTVE proporcionará a los oferentes a este Lote 10 minutos de un programa de corte informativo en catalán. El oferente deberá generar y enviar los subtítulos a CRTVE en formato SRT, 15 minutos después de recibir el archivo en los mismos términos descritos en el Lote 2.

El sistema tendrá que cumplir con los valores WER previamente indicados en los puntos 3.1.1.2 y 4.1.1.2.

7 Coordinación con RTVE

Los ofertantes a los tres lotes nombrarán un responsable del Servicio dentro de su equipo, que será el Coordinador. Este Coordinador será el interlocutor único con el personal de RTVE para la coordinación y priorización de tareas. Así mismo, desde RTVE también habrá otro coordinador que será su único contacto para cualquier duda o problema que pueda surgir.

Se establecerán reuniones ordinarias quincenales para seguimiento del proyecto. Será labor de los coordinadores la convocatoria de reuniones extraordinarias si se considerara necesario.

Algunas de las actividades que realizará el Coordinador serán:

- Planificación y asignación de tareas.

- Supervisión de actividades.
- Supervisión del cumplimiento de los trabajos.
- Gestión de prioridades.
- Mejora de las bases de datos de conocimiento.
- Optimización de procesos.
- Asistencia a las reuniones de coordinación.
- Elaboración de informes en base a métricas que permitan medir el rendimiento, la efectividad y la calidad del servicio.

8 Formato de la oferta técnica

La oferta técnica presentada tendrá la estructura que se indica a continuación.

- Relación de instalaciones, medios técnicos y humanos. En el caso del Lote 3 incluirá información de los recursos utilizados para cada lengua con la certificación que acredite debidamente a dichos recursos.
- Relación pormenorizada de todos los trabajos actuales o anteriores relacionados con:
 - Lote 1: subtítulo automático bilingüe y en directo en los idiomas explicitados generando los subtítulos correspondientes en los mismos idiomas que se hable especificando fecha y duración del contrato, organismo o empresa para la que se ha trabajado, idiomas en los que se han generado los subtítulos en cada caso y horas de subtítulo para todos los casos.
 - Lote 2: subtítulo automático bilingüe y en directo en castellano y catalán especificando fecha y duración del contrato, organismo o empresa para la que se ha trabajado y horas de subtítulo en cada caso.
 - Lote 3: análisis de calidad en automatizaciones relacionadas con PLN con las métricas y en los idiomas propuestos, especificando fecha y duración del contrato, organismo o empresa para la que se ha trabajado e idiomas en los que se ha realizado el control de calidad.
- Organización del trabajo y sistema de aseguramiento de la calidad del servicio.
- Cumplimiento y mejoras de las especificaciones técnicas requeridas.
- Para los lotes 1 y 2, planificación de tiempos, lo más detallada posible de los plazos dedicados a cada tarea necesaria para la puesta en servicio como configuración del sistema, pruebas, etc teniendo en cuenta la fecha solicitada en el PCG para la entrada en servicio. En el caso del lote 1 se especificará la planificación para cada Centro.
- Marca y modelo de los equipos a instalar en RTVE.
- Ajuste, configuración y pruebas del sistema.
- Documentación descriptiva de las características del sistema, incluyendo información del equipamiento necesario para el servicio según lo establecido en estas especificaciones técnicas que estará instalado en RTVE incluyendo marca, modelo y licencias de software instalado en el mismo.