

Especificaciones técnicas

**CONTROL DE CALIDAD Y ANÁLISIS
MEDIANTE IA DE LOS PROGRAMAS DE RNE**



1 Introducción

En el presente expediente se propone la contratación de un servicio de control de calidad mediante el que se verifiquen de forma independiente y aséptica los resultados aportados por la herramienta contratada por RTVE, mediante el expediente S-04505-2024 Inteligencia artificial aplicada al análisis de los programas de RNE, (lote 1). El objetivo de esta nueva contratación, es mantener la calidad de dichos resultados, así como optimización del sistema a través de la evaluación y validación de la media analizada.

Para tal propósito, el adjudicatario de este expediente no podrá ser el mismo que el adjudicatario que resultó en el Lote 1 del pliego S-04505-2024 “Inteligencia artificial aplicada al análisis de los programas de RNE” ni tratarse de otra empresa que tenga relación comercial o de otro tipo con la empresa adjudicataria del servicio recogido en dicho Lote .

2 Lote único. Control de calidad y análisis mediante IA de los programas de RNE

El adjudicatario deberá realizar un control de calidad sobre los resultados de la media procesada y metadatada por el sistema del adjudicatario del Lote 1 del pliego S-04505-2024 “Inteligencia artificial aplicada al análisis de los programas de RNE”.

Para ello, se le habilitará un espacio de trabajo en la plataforma de gestión descrita en el pliego anteriormente mencionado donde podrá localizar y descargar aquellos segmentos de texto elegidos para revisión y al que podrá acceder mediante las credenciales asignadas al perfil “Validador”.

La evaluación consistirá en un análisis de los resultados presentados por el automatismo del sistema del Lote 1 del pliego S-04505-2024 en las categorías y métricas especificadas. Este análisis se realizará cada semana durante la fase de servicio del Lote 1, siguiendo la metodología descrita en el Anexo I de este pliego técnico y se distribuirá de la siguiente manera:

- Sobre una o varias catas con un total de 20 minutos de contenido de la programación de RNE tratada. Se estima 1,5 horas de trabajo por cada 20 minutos de contenido.
 - El análisis a realizar sería: cálculo de WER.
 - Cálculo de porcentaje de acierto en segmentación de hablantes.
 - Cálculo de porcentaje de acierto en segmentación de noticias.
- Sobre 1 hora de contenido de la programación de RNE tratada, el cálculo del resto de parámetros. Se estima 1,5 horas de trabajo por cada hora de contenido.
 - El análisis a realizar sería: clasificación de hablantes por género, así como la comprobación de que se ha realizado la clasificación de noticias y la segmentación semántica y detectado de forma coherente entidades, palabras clave y signos de puntuación y capitalización.

Las reglas para la selección de las catas serán comunicadas al adjudicatario de este expediente por parte de RTVE al comienzo del proyecto. No obstante, y en función de la marcha del proyecto, RTVE podrá variar los programas o segmentos a analizar, manteniendo siempre el mismo número de minutos. En ningún caso se incluirán catas con datos especialmente protegidos ni de colectivos vulnerables.

El adjudicatario de este expediente deberá realizar las siguientes acciones:

- Seleccionar y descargar desde la plataforma de gestión el texto transcrito correspondiente a las catas de media que han de ser analizadas.
- Corregir cada parámetro a validar sobre una copia del texto original.

- Entregar el archivo original y el archivo corregido a RTVE y al adjudicatario del Lote 1 del pliego S-04505-2024 Inteligencia artificial aplicada al análisis de los programas de RNE una vez finalizado el proceso de corrección y validación, y con el fin de contribuir a la continua actualización y mejora del sistema de análisis de programas de RNE .
- En un plazo de 4 días laborables al finalizar cada semana analizada se adjuntará la siguiente documentación:
 - **Excel de datos semanales.** En un formato a indicar por RTVE una vez adjudicado el expediente, se presentará de forma incremental los resultados de la última semana analizada.
 - **Un informe** por cada una de las muestras analizadas con la siguiente estructura:
 - Resumen de datos: Un cuadro en que se refleje de forma individual el recuento numérico de cada uno de los parámetros descritos.
 - Análisis de la transcripción: Un documento de texto con el contenido de la transcripción sobre el cual se marcarán de forma distinguible con un código de colores, los distintos errores especificados en los puntos anteriores.
 - Comentarios / consideraciones: Un análisis subjetivo acerca de las impresiones derivadas del análisis realizado.
- 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.
- Aunque las entregas de los informes se solicitan una vez por semana, cualquier incidencia ya sea en el servicio o incumplimiento de los márgenes de calidad establecidos deberá reportarse a RTVE en el mismo día del análisis. RTVE podrá requerir el envío de los resultados con periodicidad inferior a la semana si se dan unos niveles de calidad críticos y así lo considera.

Dado que la actividad del adjudicatario está supeditada a la del Lote 1 del pliego S-04505-2024 Inteligencia artificial aplicada al análisis de los programas de RNE, en caso de que por alguna circunstancia se produzca una interrupción en su actividad, los trabajos recogidos en el presente expediente, además de las causas indicadas en la cláusula 18ª del Pliego de Cláusulas Administrativas Particulares, 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 de este expediente quedará interrumpida. En función de las circunstancias que originaran este paro, podría dar lugar a un cambio en la fecha de finalización del contrato.
- La cancelación por cualquier causa del contrato al adjudicatario del Lote 1 del pliego S-04505-2024 Inteligencia artificial aplicada al análisis de los programas de RNE daría lugar a la cancelación del contrato de este expediente.

3 Coordinación con RTVE

La empresa adjudicataria nombrará 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.

4 Formato de la oferta

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

Propuesta técnica:

- Relación de instalaciones, medios técnicos y humanos.
- Organización del trabajo y sistema de aseguramiento de la calidad del servicio.
- Cumplimiento y mejoras de las especificaciones técnicas requeridas.
- Planificación de tiempos, lo más detallada posible de los plazos de configuración del sistema, pruebas y puesta en servicio.
- Todos los ítems explicitados en la oferta económica deberán estar incluidos y desarrollados en la oferta técnica para su valoración.

Anexo I- Niveles de calidad en el metadato: métricas.

A continuación, se describen las métricas que se utilizarán para la evaluación de cada una de las funcionalidades solicitadas en el presente expediente.

- Cálculo de porcentaje de acierto en segmentación de **Transcripción de audios a texto**: la transcripción de audio a texto se calculará mediante la tasa de error por palabras (WER en su acrónimo en inglés) que se obtiene comparando la transcripción automática con la transcripción real. Responde la siguiente fórmula matemática:

$$WER = \frac{N_{INS} + N_{BOR} + N_{SUS}}{N}$$

Donde:

- N_{INS} es el número de inserciones o palabras reconocidas que no se han pronunciado.
 - N_{BOR} es el número de borrados o palabras no reconocidas pero que se han pronunciado.
 - N_{SUS} es el número de sustituciones o palabras reconocidas erróneamente.
 - N el número total de palabras a reconocer.
- **Puntuación y capitalización**: se centrará en puntos y mayúsculas.
 - **Segmentación por hablantes y por noticias**: en este caso, la exactitud se entiende como la proporción entre los fragmentos correctamente segmentados y el número total de fragmentos existentes.

$$Exactitud = \frac{C}{R}$$

Donde:

- C es el número de fragmentos segmentados correctamente.
 - R es el total de fragmentos existentes.
- **Segmentación de hablantes por género**: En este caso, la exactitud se entiende como la proporción entre los fragmentos correctamente asignados a voces de mujer/hombre y el número total de fragmentos de cada una de las voces.

$$Exactitud = \frac{C_m}{R_m}$$

Donde:

- C_m es el número de fragmentos de voces de mujer detectados correctamente.
- R_m es el total de fragmentos de voces de mujer existentes.

$$Exactitud = \frac{C_h}{R_h}$$

Donde:

- C_h es el número de fragmentos de voces de hombre detectados correctamente.
- R_h es el total de fragmentos de voces de hombre existentes.

- **Reconocimiento de entidades**

El sistema debería ser capaz de identificar al menos, los siguientes tipos de entidades:

- **Personas (PER):** nombres de personas individuales y nombres de familias, nombres que hacen referencia a grupos de personas que se integran en una organización formal (madridistas, atléticos), personajes de ficción, gentilicios (europeos, italianos, madrileños), iniciales y pseudónimos, títulos, tratamientos o cargos cuando no están asociados a un nombre propio (Presidente del gobierno, Conde de Ahumada, Primer ministro).
- **Organizaciones (ORG):** partidos políticos, instituciones públicas, administraciones públicas, Organizaciones no gubernamentales (ONG's), organizaciones internacionales, organizaciones militares, organizaciones religiosas, compañías, equipos deportivos, instituciones educativas y de investigación, grupos de música, empresas de entretenimiento y de radio y televisión.
- **Localizaciones (LOC):** topónimos de ciudades, países, provincias, regiones, continentes, masas de agua y formaciones geológicas.
- **Eventos (EVT):** conferencias, ruedas de prensa, comparencias, conciertos, festivales, periodos de tiempo (Semana Santa, Navidad, Carnaval, etc.), guerras, batallas, desastres provocados por la acción del hombre, desastres naturales, fenómenos atmosféricos, juicios, etc.
- **Productos (PRO):** nombres de productos electrónicos, vehículos, armamento, servicios web, libros, software, películas, programas de televisión, periódicos y periódicos digitales, obras de arte, nombres de documentos legales, tratados, iniciativas o programas, etc.
- **Fechas (DAT):** menciones a años o periodos de años (1977, 1936-1939), décadas (los ochenta, los 70), fechas concretas (25 de abril, 14 de julio, 25/12/1992) menciones a un siglo o periodo de siglos (expresados con números romanos o con referencias del tipo "Siglo de oro", Renacimiento, etc.) o referencias a fechas emblemáticas como 11-S, 23F.
- **Instalaciones (FAC):** edificios y estructuras desarrolladas por el ser humano: estadios, parques, museos, teatros, hoteles, aeropuertos, puertos, estaciones de tren, iglesias, calles, autopistas, puentes, túneles, aparcamientos y otros edificios, equipamientos o servicios urbanos o no.
- **Extracción palabras clave,** será requerido, lo que implica demostrar que la funcionalidad está disponible y que los términos extraídos son relevantes en el fragmento de texto.
- **Clasificación del contenido:** en este caso, la exactitud se entiende como la proporción entre los contenidos correctamente clasificados y el número total de contenidos a clasificar.

$$Exactitud = \frac{C}{R}$$

Donde:

- C es el número de contenidos clasificados correctamente.
- R es el total de contenidos clasificados.
- **Segmentación semántica** será requerido, lo que implica demostrar que la funcionalidad está disponible y que los resultados obtenidos aportan valor documental.