

Especificaciones técnicas

INTELIGENCIA ARTIFICIAL APLICADA AL ANÁLISIS DE ODS EN RTVE



1 Introducción

El 25 de septiembre de 2015 bajo el nombre “Transformar nuestro mundo la Agenda 2030 para el Desarrollo Sostenible”, (ODS) los líderes mundiales adoptaron un conjunto de objetivos globales para erradicar la pobreza, proteger el planeta y asegurar la prosperidad para todos como parte de una nueva agenda de desarrollo sostenible.

Cada uno de los 17 objetivos tienen metas específicas que deberían alcanzarse en esos próximos 15 años. Los estados y sus dirigentes se deben comprometer, con todas las herramientas disponibles a su alcance, a conseguir estas metas.

Compromiso clave, también adquirido por la Corporación de Radio Televisión Española (CRTVE), es implicar de forma urgente a la sociedad en estos objetivos haciendo difusión de los mismos y concienciando de su responsabilidad en la exigencia de su cumplimiento a todos los niveles, local, nacional y global, haciendo valer que la sociedad es parte esencial de la solución y a su vez su principal beneficiaria.

Se busca, por tanto, un servicio que nos permita disponer de un sistema de monitorización y cuantificación que nos dé una perspectiva de hasta qué punto se difunde el objetivo marcado.

Durante el último año se ha trabajado en una PoC, amparada por el expediente: S-02784-20200504 “Servicio de análisis de difusión de los Objetivos de Desarrollo Sostenible”, de la cual se han obtenido unos resultados totalmente satisfactorios. De esta forma, con el presente expediente se busca la contratación de un servicio de análisis de difusión de los 17 ODS, como se especificará en el Lote 1.

Por otro lado, y con el objetivo de garantizar que el servicio cumple los requisitos de calidad mínimos especificados por CRTVE, en el Lote 2 se solicita un servicio de control de calidad externo mediante el que se verifique de forma independiente y aseptica los resultados aportados por la empresa adjudicataria del Lote 1.

Dado el propósito del lote 2, el adjudicatario de este lote no podrá ser el mismo que para el lote 1 ni de ninguna empresa que tengan relación comercial o de otro tipo con la empresa que realiza el sistema de monitorización y cuantificación.

2 Objeto del contrato

Mediante este expediente se quiere disponer de una herramienta que permita, en un entorno controlado, monitorizar, clasificar y cuantificar de forma automatizada el tiempo dedicado, en los programas de Televisión Española (TVE) seleccionados, a cada uno de los 17 ODS.

Al inicio del servicio, la herramienta debe integrar operativamente dos ODS: 5 “Igualdad de género” y 13 “Lucha contra el cambio climático”. A partir de ese momento, se incorporará progresivamente al sistema un nuevo ODS cada 4 semanas, hasta completar los 17.

3 Lote 1. Servicio de análisis y clasificación ODS

El oferente al Lote 1 deberá ofertar un servicio automatizado de clasificación de texto basado en Inteligencia Artificial (IA) que proporcione una cuantificación del tiempo que se emplea, en los contenidos señalados por CRTVE, en tratar temas relativos a ODS, utilizando como fuente de datos el subtítulo de éstos.

Para ello, se pondrá el foco en programas informativos, de los cuales la CRTVE seleccionará aquellos sobre los que se aplicará el servicio, y serán comunicados al adjudicatario en el momento de la adjudicación. El volumen de contenidos a analizar será de 65 horas semanales. Será potestad de la CRTVE durante el periodo de contrato, cambiar el programa o programas sobre los que se aplicará el servicio sin ningún coste añadido.

CRTVE proporcionará mensualmente al adjudicatario de este Lote 1 los subtítulos de los programas a analizar mediante ficheros con formato estándar de subtitulado (SRT, VTT). La forma de acceso a estos ficheros será manual y consensuada con el adjudicatario al inicio del proyecto. Sin embargo, está previsto automatizar estos envíos, por lo que el adjudicatario deberá asumir en el periodo de duración del contrato y sin coste para CRTVE un posible cambio en el procedimiento.

El adjudicatario será el responsable de elaborar una ontología por cada uno de los ODS que refleje los conceptos más importantes del dominio y la relación entre ellos. En este proceso deberá participar personal experto en cada materia que será previamente propuesto a CRTVE para su validación. El anexo I contiene una relación de los perfiles que se consideran adecuados según el ODS de que se trate.

Además, estas ontologías deberán ser actualizadas basándose en las correcciones realizadas por el adjudicatario del Lote 2 a lo largo de la vida del proyecto. Una vez elaboradas, y superados los criterios de calidad mínimos establecidos, el adjudicatario deberá entregar a CRTVE un archivo, en el formato que se defina una vez adjudicado el lote, con el contenido de dichas ontologías. Esta entrega se realizará al finalizar el proyecto y en cualquier momento a lo largo de la vida del mismo en que CRTVE lo considere necesario.

CRTVE dispondrá de una o varias bases de datos con un contenido de aproximadamente 300 horas de informativos, en el que está garantizado que se incluyen noticias relacionadas con todos los ODS y que servirá para comprobar el buen funcionamiento del sistema en cada nueva incorporación. Para ello, el adjudicatario deberá analizar los contenidos de una de estas bases de datos cuando la ontología de un ODS esté finalizada (es decir justo antes de la entrada en servicio) y a los tres meses de comenzar el servicio con el fin de garantizar que se cumplen los objetivos de calidad definidos en el punto 3.2.

Dado que, tanto el sistema de monitorización y cuantificación como la ontología sobre los ODS 5 y 13 contemplados en este Lote 1 deberán estar totalmente operativos al comienzo del servicio, la CRTVE podrá requerir a todos los oferentes una demostración del funcionamiento de los sistemas ofertados para estos ODS. Con una antelación máxima de 7 días respecto a la fecha fijada para la prueba, CRTVE enviará a los oferentes un comunicado con información acerca del día, hora, procedimiento etc, así como un archivo SRT o VTT de idénticas características a los que se utilizarán el día de la prueba, para ello deben indicar en sus ofertas técnicas la información de contacto más conveniente.

El día de la prueba y para la realización de la misma se proporcionarán los archivos SRT o VTT con los subtítulos correspondientes a 10 horas de programas informativos. Este material será accesible al oferente mediante un servidor ftp o correo electrónico al inicio de la misma. Los oferentes dispondrán de un tiempo de 15 minutos para:

- La agrupación de subtítulos en segmentos que podrán ser frases o noticias, dejándose a criterio del oferente una u otra opción, siempre y cuando explique el criterio adoptado.
- El análisis y clasificación de los contenidos.
- La devolución de resultados obtenidos, por correo electrónico, a la dirección que será facilitada antes del comienzo de la prueba.

Los resultados se recogerán en un archivo Excel donde se rellenen los siguientes campos por columnas:

- Texto con los contenidos de los subtítulos agrupados por segmentos.
- Duración de cada segmento en milisegundos.
- Resultados de la clasificación de los segmentos en función de los ODS 5 y 13.

Para que la prueba sea superada el sistema deberá alcanzar, como mínimo, un 90% de precisión.

3.1 Requisitos funcionales del sistema

Todo el sistema estará operativo en la nube, ajena a CRTVE y cumpliendo con la normativa de la UE, y deberá poder realizar las siguientes funciones:

- Procesar el texto de los programas proporcionados por CRTVE en formato SRT o VTT.
- Agrupar los subtítulos en estructuras más completas, bien frases o bien noticias, dejándose a criterio del oferente una u otra opción, siempre y cuando explique el criterio adoptado.
- Disponer de una base de datos que contenga al inicio del proyecto las ontologías de los ODS 5 y 13, y a la que se irán incorporando las del resto de ODS según se vayan integrando al proyecto, hasta el total de 17 ontologías.
- Para cada uno de los informativos a analizar, localizar, identificar y etiquetar aquellas frases o noticias que traten temas relacionados con cada ODS en base a las ontologías descritas en el punto anterior.
- Cuantificar la duración de las frases o noticias identificadas.

Los resultados obtenidos deberán estar disponibles para CRTVE en una plataforma web accesible mediante un navegador web común (Chrome, Firefox) y con acceso protegido mediante usuario y contraseña en los siguientes términos:

- Acceso mínimo de 4 usuarios simultáneos.
- Presentación de resultados en un formato amable, de fácil acceso y previamente acordado con la CRTVE, no dándose por válido dicho formato hasta que el responsable del proyecto nombrado por la CRTVE dé el visto bueno al mismo.
- Acceso directo a resultados por cada ODS. Los Excel generados en el proceso inmediato de validación automático, descritos en párrafos siguientes, también estarán disponibles en esta plataforma.
- Presentación de un dashboard que permita:
 - Filtrar los resultados por:
 - Canal.
 - Programa.
 - Franja horaria.
 - ODS
 - Selección de varios ODS.
 - Presentar los resultados numéricamente distinguiendo por:
 - Tiempo total de horas analizadas.
 - Tiempo total dedicado a cada ODS/varios ODS.
 - Precisión del sistema sobre los archivos validados por el Lote 2.
 - Presentar gráficamente los siguientes valores:
 - Tiempo diario dedicado a cada ODS.
 - Porcentaje del tiempo diario dedicado a cada ODS respecto a la duración total del programa.
 - Tiempo mensual dedicado a cada ODS.
 - Porcentaje del tiempo mensual dedicado a cada ODS respecto a la duración total del programa.
 - Las 20 palabras clave más utilizadas por cada ODS seleccionado.
 - Ranking de las 5 palabras clave más utilizadas por cada ODS seleccionado.
- El acceso a esta plataforma deberá estar disponible en horario laboral de lunes a viernes 8:00 a 20:00h.

3.2 Requisitos no funcionales del sistema

El sistema encargado de proporcionar el servicio deberá como mínimo:

- Tener disponible en la plataforma la clasificación de los contenidos candidatos en menos de 24 horas a contar desde la entrega por CRTVE de los ficheros SRT o VTT.
- Permitir la escalabilidad horizontal y vertical.

El adjudicatario de este Lote 1 deberá proporcionar una serie de entregables al adjudicatario del Lote 2 con información suficiente que le permita realizar el proceso de validación y seguimiento de calidad del servicio. Estos entregables tendrán un formato Excel y contendrán los resultados obtenidos por el sistema en una serie de programas que CRTVE comunicará al comienzo del servicio y que se irán variando a lo largo de la vida del proyecto.

La distribución de estos entregables será de la siguiente manera:

- Excel tipo 1 (10h de transcripción).
 - Contenido: Muestra seleccionada de la/las bases de datos mencionadas en el punto 3.
 - Periodicidad: Este documento se generará cada vez que se incorpore un ODS al servicio y a los tres meses de haberse incorporado un ODS, hasta completar la integración de todos los ODS, excepto para los ODS 5 y 13 (uno por cada objetivo) que se generarán al inicio del servicio.
 - Envío: Cada vez que se genere el documento
- Excel tipo 2 (10h de transcripción por semana).
 - Contenido: CRTVE comunicará al comienzo del servicio el contenido.
 - Periodicidad: Este documento se generará una vez al mes durante la duración del expediente separando el contenido referido a cada semana en pestañas diferentes.
 - Envío: un único envío en los 5 primeros días naturales de cada mes, por mes vencido.
- Excel tipo 3 (10h de transcripción por semana).
 - Contenido: CRTVE comunicará al comienzo del servicio el contenido.
 - Periodicidad: Este documento se generará una vez al mes durante la duración del expediente separando el contenido referido a cada semana en pestañas diferentes.
 - Envío: un único envío en los 5 primeros días naturales de cada mes, por mes vencido
- Cada documento deberá incluir los siguientes campos, todos ellos rellenos con la información obtenida del análisis:
 - Fecha.
 - Franja horaria o título del programa.
 - Texto (subtítulos agrupados en frases o noticias).
 - Duración en milisegundos del texto.
 - Palabras clave identificadas (si procede).
 - ODS detectado (incluye un selector desplegable con la opción de los 17 ODS o campo en blanco).
 - Resultado de la validación ODS.
 - Comentarios (este último campo a rellenar en texto).

Para que la incorporación de un nuevo ODS se considere apta el análisis realizado por el adjudicatario del Lote 2 deberá alcanzar los siguientes objetivos sobre una validación en un Excel de tipo 1:

- Mínimo 70% de acierto en la primera prueba, a realizar en la primera incorporación.
- Mínimo 90% de acierto en una segunda prueba a realizar en un plazo de tres meses desde la consecución del objetivo anterior.

En el caso de los ODS 5 y 13 el objetivo será 90% de acierto al comienzo del servicio.

Además, el adjudicatario del Lote 1 deberá generar mensualmente un archivo Excel con el contenido analizado para CRTVE, con el fin de monetizar el tiempo empleado en hablar de cada ODS, que indique:

- Texto de la intervención.
- Fecha.
- Time code inicio.
- Time code final.
- Programa al que corresponde.
- Cadena donde se ha emitido.

3.3 Servicio

Desde la firma del contrato el adjudicatario dispondrá de un máximo de 4 semanas para configurar el sistema y dar comienzo al servicio. Desde ese momento estará disponible el análisis de los ODS 5 y 13 y a partir de este punto se desarrollará la ontología e incorporará al servicio un nuevo ODS cada 4 semanas como máximo. El orden de incorporación será comunicado por CRTVE una vez adjudicado el proyecto.

4 Lote 2. Control Calidad Servicio Análisis ODS

El adjudicatario de este lote deberá realizar el control de calidad de los fragmentos clasificados por el sistema del adjudicatario del Lote 1, comprobando que pertenecen al ODS especificado, señalando los mal clasificados y los no detectados por la automatización.

No se exige personal especializado para realizar el servicio en las materias a tratar para esta primera revisión, pero dado que en algunos casos no es trivial clasificar si una noticia habla o no de un ODS específico, el adjudicatario de este lote deberá contar con un servicio de consultoría con profesionales con conocimientos en las materias escogidas que estarán relacionadas con la pobreza, el combate al cambio climático, la educación, la igualdad de la mujer, la defensa del medio ambiente o el diseño de nuestras ciudades. Se estima que este servicio requerirá de una dedicación de 1 hora al mes. Se puede encontrar una descripción de los perfiles requeridos para esta tarea en el Anexo I del presente documento. El personal experto que participe en cada materia será previamente propuesto a CRTVE para su validación.

El adjudicatario de este Lote 2 recibirá del adjudicatario del Lote 1 una entrega cada vez que un nuevo ODS se incorpore al sistema y a los tres meses de su incorporación, con el siguiente contenido:

- Un archivo Excel “tipo 1” con 10h de transcripción.

Además, el adjudicatario del Lote 2 recibirá del adjudicatario del Lote 1 una entrega mensual con el siguiente contenido:

- Un archivo Excel “tipo 2” con 10h de transcripción semanales, hasta un máximo de 50 horas mensuales.
- Un archivo Excel “tipo 3” con 10h de transcripción semanales, hasta un máximo de 50 horas mensuales.

Todos los Excel incluirán el texto fragmentado y etiquetado con los siguientes campos informativos:

- Fecha.
- Franja horaria.
- Texto (subtítulos agrupados en frases o noticias).
- Duración en milisegundos del texto.
- Palabras clave identificadas (si procede).
- ODS detectado (En este proceso deberá participar personal experto en cada materia que será previamente propuesto a CRTVE para su validación.).
- Resultado de la validación ODS.
- Comentarios (este último campo a rellenar en texto).

El adjudicatario del Lote 2 deberá revisar que la transcripción de la columna “Texto” se corresponda con el objetivo marcado en la columna de “ODS detectado”. Si el valor de la columna “ODS detectado” no se corresponde con el objetivo de desarrollo del que se habla, se deberá modificar haciendo uso de un desplegable donde aparecerán las posibles opciones. Al realizar un cambio en dicho campo, el valor de la celda “Resultado de la validación ODS” se modificará automáticamente y pasará a un estado de “NOK”, “Error” o similar.

Se deberán añadir las anotaciones que se consideren oportunas cuando se detecte un fallo en la validación para la mejora del sistema en el campo “Comentarios”.

La metodología a seguir por cada uno de los Excel de validación es la siguiente:

- Excel tipo 1 (10h de contenido): revisión exhaustiva únicamente del nuevo ODS incorporado. De esta manera, la primera revisión se realizará sobre los ODS 5 y 13 simultáneamente, mientras que las posteriores serán sobre cada nuevo ODS, lo que comprende un total de 32 revisiones exhaustivas hasta completar la implementación de los 17 ODS.
- Excel tipo 2 (10h de contenido): revisión de manera exhaustiva de todos los ODS incluidos en etapas previas y que han pasado con anterioridad por la validación del Excel tipo 1.
- Excel tipo 3 (10h de contenido): revisión ligera de los ODS descritos en el Excel tipo 2.

Las revisiones exhaustiva y ligera mencionadas en el párrafo anterior se entienden de la siguiente manera:

- Revisión exhaustiva: revisión individual de cada registro/fragmento del texto.
 - Pertenencia o no al ODS bajo validación.
 - Añadir comentarios.
- Revisión ligera: revisión únicamente de los registros etiquetados como ODS por el sistema.
 - Pertenencia o no al ODS bajo validación.
 - Añadir comentarios.
 - Revisión utilizando el automatismo de búsqueda en Excel mediante términos clave relacionados con el ODS bajo estudio. A discreción del evaluador.

A partir de la recepción de los archivos enviados por parte del adjudicatario del Lote 1, el adjudicatario del Lote 2 deberá proporcionar los siguientes entregables:

- Los resultados del Excel tipo 1 se enviarán en el plazo de una semana natural, corregidos y añadiendo al nombre del archivo original “_validado”, al adjudicatario del Lote 1 y a la persona designada por CRTVE como coordinador.
- Los resultados de los Excel tipo 2 y tipo 3, que serán recibidos con el contenido dividido por semana, se enviarán uno por semana, corregidos y añadiendo al nombre del archivo original “_validado”, al adjudicatario del Lote 1 y a la persona designada por CRTVE como coordinador.
- Mensualmente se deberá presentar un informe que muestre un resumen de los aciertos y los fallos del sistema, así como los cálculos de precisión alcanzados por cada ODS.

Hasta ahora, el personal de CRTVE ha realizado el control de calidad descrito. Se estima que el tiempo necesario para realizar la validación de un conjunto de entregables es de la siguiente manera:

- Evaluación exhaustiva: 2 horas de análisis por cada 10 horas de contenido.
- Evaluación ligera: 1 hora de análisis por cada 10 horas de contenido.

Los licitadores del Lote 2 deberán presentar una oferta acorde a esta estimación de tiempos.

Adicionalmente, deberá incluir en su oferta 100 horas de análisis opcional, de la que CRTVE podrá disponer a lo largo de la vida del proyecto en caso de necesidad.

5 Coordinación con CRTVE

Las dos empresas adjudicatarias nombrarán un responsable del servicio dentro de su equipo, que será el coordinador. Este coordinador será el interlocutor único con el personal de CRTVE para la coordinación y priorización de tareas. Así mismo, desde CRTVE también habrá otro coordinador que será su único contacto para cualquier duda o problema que pueda surgir.

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

6 Formato de la oferta

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

Propuesta técnica:

- Relación de proyectos de similares características, medios técnicos y humanos.
- Diagrama de bloques con la arquitectura de la solución ofertada.
- Compromiso de calidad del servicio.
- 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.

ANEXO I

A continuación, se describen los perfiles de los expertos por ODS y conocimientos necesarios:

ODS 6 “Agua limpia y saneamiento”, ODS 12 “Producción y consumo responsables”, ODS 13 “Acción por el clima”, ODS 14 “Vida submarina”, ODS 15 “Vida de ecosistemas terrestres”:

- Titulados en carreras relacionadas con Ciencias Sociales y Jurídicas, Biología, Ingeniería, o experiencia demostrable en:
 - Cambio climático.
 - Biodiversidad.
 - Océanos y mares.
 - ONG y asociaciones especializadas en medio ambiente.
 - Instituciones como PNUMA, Agencia Europea del Medio Ambiente, Ministerio Transición Ecológica...

ODS 1 “Fin de la pobreza”, ODS 2 “Hambre cero”, ODS 4 “Educación de calidad” y ODS 10 “Reducción de las desigualdades”, ODS 3 “Salud y bienestar”, ODS 5 “Igualdad de género”:

- Titulados en carreras relacionadas con Ciencias Sociales y Jurídicas, Ciencias Biológicas y de la salud o experiencia demostrable en:
 - ODS 1, ODS 2, ODS 4 y ODS 10:
 - Pobreza.
 - Derechos Humanos.
 - Sociología.
 - ONG y asociaciones.
 - Instituciones como ONU, FAO, UNICEF, ACNUR...
 - ODS 3:
 - ONG y asociaciones, especialmente enfocadas en temas de salud (ej. AECC).
 - Instituciones como OMS.
 - ODS 5:
 - Expertas en género.
 - ONG y asociaciones, especialmente enfocadas en mujeres.
 - Instituciones como ONU Mujeres, Ministerio de Igualdad, Instituto de las Mujeres.

ODS 7 “Energía asequible y no contaminante”, ODS 8 “Trabajo decente y crecimiento económico”, ODS 9 “Industria, innovación e infraestructura”, ODS 11 “Ciudades y comunidades sostenibles”:

- Titulados en carreras relacionadas con Ciencias Sociales y Jurídicas, Economía, Ingeniería y Arquitectura, o experiencia demostrable en:
 - Derechos laborales.
 - Innovación.
 - Asociaciones y ONG Instituciones como la OIT, Ministerio de Trabajo.

ODS 16 “Paz, justicia e instituciones sólidas”, ODS 17 “Alianza para lograr los objetivos”:

- Titulados en carreras relacionadas con Ciencias Sociales y Jurídicas, o experiencia demostrable en temas relacionados con:
 - Derechos Humanos.
 - Leyes y justicia.
 - Instituciones como la ONU, ACNUR, OMC...