Código contractual	SE188151HO2023049

Especificaciones técnicas del contrato de servicios de investigación y desarrollo en tecnologías de computación en la nube para el desarrollo del proyecto "Aprendizaje automático federado para redes eléctricas de distribución".

Por el hecho de presentar una oferta, los licitadores participantes aceptan las especificaciones técnicas aquí establecidas, que se consideran los requisitos mínimos obligatorios.

Toda propuesta que no cumpla los requisitos mínimos establecidos en las presentes especificaciones técnicas quedará excluida del procedimiento de contratación.

1. Objeto del contrato o necesidad del poder adjudicador

El objeto de este contrato es la prestación de servicios de Investigación y Desarrollo en tecnologías de computación en la nube, en el marco del Proyecto "Aprendizaje automático federado para redes eléctricas de distribución", conforme al Real Decreto 959/2022, de 15 de noviembre, y a la Resolución de la titular de la Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales.

Las tareas de Investigación y Desarrollo a realizar se enmarcan en las siguientes actividades principales del Proyecto Coordinado, detalladas en el Anexo II del Real Decreto 959/2022, de 15 de noviembre:

ACTIVIDADES DEL PROYECTO			
Cod.	**Actividad**	**Resumen trabajo a realizar**	**Lote**
T1.1.1.1	Identificación de casos de uso y compartición de datos	Se analizarán los posibles casos de uso para desarrollar el Concepto del proyecto, incluyendo una red de distribución rural con gran penetración de energías renovables y una red de distribución con alta presencia de cargadores EV bidireccionales. Se identificarán y adquirirán fuentes de datos que puedan alimentar los algoritmos de mejora de la operación y planificación de actuaciones en las redes de distribución. El formato de los datos adquiridos (CSV, JSON, etc.) y la resolución serán analizados y tratados cuando sea necesario. Las fuentes de datos incluyen, entre otros, medidores inteligentes, concentradores, analizadores de red, SCADA, PMU, fuentes de datos fotovoltaicas, fuentes de datos de vehículos eléctricos y fuentes de datos de sistemas de almacenamiento. Si algunos de los datos requeridos no están disponibles, se generarán datos sintéticos.	1

T1.1.1.2	Adquisición de datos	Se adquirirán aquellos datos identificados en la Tarea 1.1.1.1 que servirán de entrada a los algoritmos de mejora de la operación y planificación de los casos de uso seleccionados en la Tarea 1.1.1.1. Esta tarea también comprende la identificación de protocolos de comunicación y datos además del desarrollo de los adaptadores de datos necesarios para implementar como base de acceso a la plataforma digital. Para la adquisición de estos datos se instalarán los equipos de captación necesarios o se reconfigurarán los existentes para recuperar la información mediante protocolos de comunicaciones tales como IEC 60870-5-103, IEC 61850, DLMS y Modbus. Los datos se enviarán mediante la red de comunicaciones disponible y en un protocolo estándar (IEC 60870-5-104) hacia los servidores del operador de la red, donde se instalará el software necesario para procesar la información de la red, incluyendo el filtrado, normalización y formateado de los datos, y enviarla hacia los servidores cloud.	2
T1.1.2	Preparación de datos. Protocolos y conectores	Esta tarea contempla la persistencia y securización de los datos adquiridos por la plataforma digital, de forma que puedan estar disponibles para el resto de procesos.	3
T1.2	Preparación y armonización de los datos	Esta tarea se centra en la comprensión y el preprocesamiento de los datos. Son pasos fundamentales para asegurar la calidad de los datos (distinguir información relevante y no relevante, datos que pueden ser redundantes, ruidosos o poco fiables). La comprensión de los datos (el paso inmediato antes de la preparación de los datos) se logra a través del análisis basado en estadísticas descriptivas y visualización. Luego, los datos se preparan mediante la limpieza de datos, la selección de características y las transformaciones. Las eventuales anomalías se tratarán mediante el reconocimiento de patrones y técnicas de aprendizaje automático (p. ej., técnicas basadas en la densidad, detección de valores atípicos basada en el análisis de agregados, técnicas de conjunto). En función de la anomalía detectada, se realizará automáticamente una corrección. Esto permitirá separar los conjuntos de datos en los conjuntos de datos de entrenamiento y prueba. Además, esta tarea contempla el desarrollo e implantación de un modelo de datos coherente con los datos identificados a ingestar y también compatible con las ontologías y modelos de datos horizontales para energía usados en Espacios de Datos, que permitan un flujo fácil de estos datos en aquellos sistemas en caso de ser necesario. Este modelo de datos servirá asimismo como idioma común para los servicios de Inteligencia Artificial, de modo que puedan abstraerse del	3

		origen de los mismos. Para ello, además del modelo de datos común se desarrollará una interfaz de acceso estándar.	
T1.3	Conexión y compartición en Espacios de datos	Esta tarea se encarga de los mecanismos que hay que poner en funcionamiento para habilitar la compartición bidireccional de datos entre la plataforma digital del proyecto y posible Espacios de Datos donde se decida federarse. Esto se consigue mediante el despliegue de conectores para Espacios de Datos, que permitan verificar la identidad digital de la plataforma, hacer el registro de los metadatos que describan los datasets disponibles y alinearse al modelo de datos horizontal del Espacio de Datos que habilite el descubrimiento. Además se estudiarán e implementarán las políticas necesarias que aseguren la soberanía y condiciones de intercambio de datos para aquellos propiedad de los socios del proyecto.	3
T1.4.1	Orquestación y despliegue de servicios en contenedores	Esta tarea definirá las estrategias y tecnologías necesarias para conseguir el despliegue efectivo de la plataforma. Debido a sus numerosas ventajas en escalabilidad, estabilidad y facilidad de despliegue los servicios serán desplegados en contenedores. Asegurando así su aislamiento y desacoplándolos del resto de servicios. Docker es una de las tecnologías de “containerización” más extendidas actualmente, aunque se estudiará la viabilidad de posibles alternativas (containerd, cri-o, podman, etc). Para orquestar tal arquitectura se estudiará el estado del arte y se seleccionará la mejor tecnología de orquestación atendiendo a distintas métricas: soporte de la comunidad, rendimiento o apertura del código, entre otras. Una de las posibles tecnologías a evaluar será Kubernetes que actualmente ostenta el estado del arte y es un estándar de facto en cuanto a tecnologías de orquestación.	3
T1.4.2	Despliegue de modelos cloud en el edge	El uso de los modelos desarrollados en el cloud, a partir de la información recogida de instalaciones reales, requiere que los propios dispositivos de captación actúen como procesadores Edge, que permitan la validación de los modelos, con la comparación entre los datos esperados y los recolectados, y después el reentrenamiento de los modelos para mantenerlos ajustados a las variaciones que puedan existir en los elementos físicos. Para ello, se diseñarán e implementarán los mecanismos de volcado y ejecución de los modelos en los dispositivos remotos, así como los mecanismos de comunicación y control.	2
T4.1	Machine Learning como servicio (MLaaS)	El desarrollo de una capa adicional conocida como MLaaS permite abstraer la complejidad a la hora de construir y desplegar servicios de IA. De este modo, se eliminan barreras como la curva de aprendizaje para reducir drásticamente el tiempo de acceso al mercado en productos basados en IA. En	3

		esta tarea se recogerá el progreso obtenido y se garantizará un despliegue semi-automático de los componentes necesarios para ejecutar los servicios de IA, siempre dentro del marco de entornos basados en Kubernetes. Para ello, se estudiará la utilización de los framework Kubeflow y KubeEdge, que garantiza una operación sencilla, portable y escalable sin importar la infraestructura subyacente, tanto en plataformas en la nube como en dispositivos edge.	
T4.3	Integración de servicios IA	Esta tarea se encarga de (1) asegurar que todos los algoritmos de operación y planificación estén correctamente desplegados en contenedores (2) que expongan todas las APIs necesarias para su interacción y (3) puedan orquestarse en la plataforma digital, conectándose con las fuentes de datos y produciendo los resultados esperados. Esta tarea velará asimismo por la integración de la plataforma del proyectos con los Espacios de Datos externos, a través de los conectores y componentes desarrollados en la tarea T1.3. Se asegurará la bidireccionalidad de esta relación, pudiendo usarse datos provenientes de dichos Espacios de Datos externos y también indexando parte de los datos generados en la plataforma para ser compartida en este tipo de entornos, valorando tanto su monetización como posible intercambio.	3
T4.4	Gestión del ciclo de vida de un proceso de AI (MLOps)	Como complemento y extensión directa de la T4.1, esta tarea engloba la operación y monitorización continua del ciclo de vida de un servicio de ML. Para ello, conecta directamente con la adquisición de datos desde el Espacio de Datos, el entrenamiento y gestión de modelos, despliegue o puesta en producción de los mismos y posterior monitorización continua durante su ejecución. En este punto, se utilizarán plataformas de gestión de modelos como MLFlow, de datasets como DVC, de ejecución de modelos como KServe o TensorFlow Serving o de monitorización como Evidently AI o WhyLabs (herramientas a definir durante la ejecución de la tarea).	3

2. Actividades y funciones de los licitadores participantes

Los licitadores deben asumir las siguientes funciones, tal como se detallan a continuación:

Lote 1 ‘Identificación de casos de uso y compartición de datos’

Se analizarán los posibles casos de uso para desarrollar el Concepto del proyecto, incluyendo una red de distribución rural con gran penetración de energías renovables y una red de distribución con alta presencia de cargadores EV bidireccionales. Se identificarán y adquirirán fuentes de datos que puedan alimentar los algoritmos de mejora de la operación y planificación de actuaciones en las redes de distribución. El formato de los datos adquiridos (CSV, JSON, etc.) y la resolución serán analizados y tratados cuando sea necesario. Las fuentes de datos incluyen, entre otros, medidores inteligentes, concentradores, analizadores de red, SCADA,

PMU, fuentes de datos fotovoltaicas, fuentes de datos de vehículos eléctricos y fuentes de datos de sistemas de almacenamiento. Si algunos de los datos requeridos no están disponibles, se generarán datos sintéticos.

El lote 1 incluye la tarea T1.1.1.1.

Lote 2 'Adquisición de datos y despliegue de modelos cloud en el edge'

Se adquirirán aquellos datos identificados que servirán de entrada a los algoritmos de mejora de la operación y planificación de los casos de uso seleccionados. Para la adquisición de estos datos se instalarán los equipos de captación necesarios o se reconfigurarán los existentes para recuperar la información mediante protocolos de comunicaciones tales como IEC 60870-5-103, IEC 61850, DLMS y Modbus. Los datos se enviarán mediante la red de comunicaciones disponible y en un protocolo estándar (IEC 60870-5-104) hacia los servidores del operador de la red, donde se instalará el software necesario para procesar la información de la red, incluyendo el filtrado, normalización y formateado de los datos, y enviarla hacia los servidores cloud. Se diseñarán e implementarán los mecanismos de volcado y ejecución de los modelos en los dispositivos remotos, así como los mecanismos de comunicación y control, para explorar el despliegue de modelos cloud en el Edge.

El lote 2 incluye las tareas T1.1.1.2 y T1.4.2.

Lote 3 'Desarrollo de la plataforma digital diseñada y creada para el espacio de datos de energía'

Esta lote se enfoca en comprender y preprocesar los datos para asegurar su calidad, incluyendo análisis estadístico y visualización, limpieza, selección y transformación de datos, detección y corrección de anomalías y separación en conjuntos de entrenamiento y prueba. También se desarrollará un modelo de datos coherente con las ontologías y modelos de datos horizontales para energía usados en Espacios de Datos, y una interfaz de acceso estándar. Se desplegarán conectores para Espacios de Datos, se implementarán políticas para asegurar la soberanía y condiciones de intercambio de datos, y se definirán estrategias y tecnologías para el despliegue efectivo de la plataforma, incluyendo contenedores y orquestación con tecnologías como Kubernetes. Se desarrollará una capa adicional conocida como MLaaS para abstraer la complejidad de construir y desplegar servicios de IA, y se asegurará que todos los algoritmos de operación y planificación para redes de distribución estén correctamente desplegados y orquestados en la plataforma digital. También se integrará la plataforma del proyecto con Espacios de Datos externos, permitiendo la bidireccionalidad y posible intercambio de datos. También se incluirá operación y monitorización continua del ciclo de vida de un servicio de ML. Esto conecta directamente con la adquisición de datos desde el Espacio de Datos, el entrenamiento y gestión de modelos, despliegue o puesta en producción de los mismos y posterior monitorización continua durante su ejecución. En este punto, se utilizarán plataformas de gestión de modelos como MLFlow, de datasets como DVC, de ejecución de modelos como KServe o TensorFlow Serving o de monitorización como Evidently AI o WhyLabs (herramientas a definir durante la ejecución de la tarea).

El lote 3 incluye las tareas T1.1.2, T1.2, T1.3, T1.4.1, T4.1, T4.3 y T4.4.

Las ofertas presentadas por los licitadores participantes deberán abarcar todas las actividades y funciones mencionadas en el presente pliego de prescripciones técnicas y en las

condiciones particulares de los contratos públicos, todas ellas obligatorias para que una oferta sea admitida.

3. Requisitos técnicos generales obligatorios de la prestación y/o criterios de rendimiento o requisitos funcionales de la prestación

Con el fin de garantizar y asegurar unos altos estandares de calidad para los entregables, se describen a continuación los requisitos para cada uno de ellos en la siguiente tabla:

REQUISITOS DE ENTREGABLES POR LOTES					
Cod.	**Requisito**	**Descripción**	**Criterio de Aceptación**	**Entregable**	**Lote**
R1	La plataforma digital se desarrolla en base a principios FAIR y cumpliendo GDPR	Se cumple el objetivo O1 de desarrollar una primera versión plataforma digital preparada para el espacio de datos de energía y para integrar servicios IA, que es consistente con FAIR y GDPR y se documenta adecuadamente	Cumplimiento de principios FAIR (Findable, Accessible, Interoperable and Re-usable) y GDPR (General Protection Data Regulation) se refleja en el documento	E1	1,2 y 3
R2.1	La plataforma digital se desarrolla en base a principios FAIR y cumpliendo GDPR	Se cumple el objetivo O1 de desarrollar una plataforma digital preparada para el espacio de datos de energía y para integrar servicios IA, que es consistente con FAIR y GDPR y se documenta adecuadamente	Cumplimiento de principios FAIR (Findable, Accessible, Interoperable and Re-usable) y GDPR (General Protection Data Regulation) se refleja en la plataforma	E2.1	2 y 3
R2.2	La plataforma digital se desarrolla en base a principios FAIR y cumpliendo GDPR	Se cumple el objetivo O1 de desarrollar una plataforma digital preparada para el espacio de datos de energía y para integrar servicios IA, que es consistente con FAIR y GDPR	Cumplimiento de principios FAIR (Findable, Accessible, Interoperable and Re-usable) y GDPR (General Protection Data Regulation) se refleja en la plataforma	E2.2	2 y 3
R5.1	Plataforma digital integra algoritmos basados en IA	Se cumple el objetivo O4 de integrar servicios basados en IA en la plataforma y se documenta adecuadamente	Los KPI que evalúan la ejecución de dichos algoritmos son los esperados	E5.1	3
R5.2	Plataforma digital integra algoritmos basados en IA	Se cumple el objetivo O4 de integrar servicios basados en IA en la plataforma	Los KPI que evalúan la ejecución de dichos algoritmos desde la plataforma son los esperados	E5.2	3

4. Sistema de seguimiento y control de la ejecución del pliego de condiciones

El contratista deberá designar una persona encargada de dirigir la ejecución del contrato, que deberá garantizar la calidad del objeto del contrato y tratar directamente con el interlocutor

designado por el órgano de contratación cualquier cuestión relacionada con el normal desarrollo de las tareas aquí indicadas.

Las citadas personas deberán reunirse, al menos, cada mes para supervisar, controlar y tratar cualquier aspecto relacionado con el desarrollo del contrato a fin de garantizar que su ejecución se ajusta a lo establecido en el presente pliego.

Con el fin de garantizar y controlar la correcta ejecución de los contratos, a continuación, se detalla la lista de entregables definidos, su tipo, su criterio de aceptación y fecha de entrega, así como los lotes afectados.

Se asume que M1 corresponde al mes en que se resuelve la licitación.

LISTA DE ENTREGABLES						
Cod.	**Entregable**	**Descripción**	**Clasificación**	**Criterio de Aceptación**	**Fecha**	**Lote**
E1	Primera versión de la plataforma digital	Documento que describe la primera versión de la plataforma digital para datos de energía	Documento	El documento refleja el proceso de creación de la plataforma y refleja que está operativa	M12	1, 2 y 3
E2.1	Reporte de la versión final de la plataforma digital	Documento que describe la versión final de la plataforma digital para datos de energía	Documento	El documento refleja el proceso de creación de la plataforma y refleja que está operativa	M19	2 y 3
E2.2	Prototipo de versión final de la plataforma digital	Prototipo de Versión final de la plataforma digital para datos y aprendizaje federado	Prototipo	Plataforma digital para datos de energía operativa	M19	2 y 3
E5.1	Reporte de Integración de servicios en la plataforma	Documento que describe el proceso de integración de los servicios basados en IA en la plataforma digital	Documento	El documento refleja cómo se accede a algunos de los algoritmos desde la plataforma y que el servicio basado en este algoritmo se ejecuta de forma satisfactoria	M21	3
E5.2	Prototipo de plataforma digital con servicios integrados	Prototipo de la plataforma digital integradora de datos de energía y servicios	Prototipo	Plataforma digital para datos de energía e integradora de servicios operativa	M21	3

5. Resultados previstos que debe proporcionar el adjudicatario

Las especificaciones técnicas propuestas por la empresa licitadora en su oferta se convertirán en condiciones obligatorias que deberán cumplirse durante la ejecución del contrato si la empresa licitadora resulta adjudicataria del mismo.

RESULTADOS PREVISTOS		
T1.1.1.1	**Identificación de casos de uso y compartición de datos**	M1-M12
Resultados previstos		
Documento que describe detalladamente los trabajos realizados en la tarea y que forma parte de E1		
T1.1.1.2	**Adquisición de datos**	M1-M12
Resultados previstos		
Documento que describe detalladamente los trabajos realizados en la tarea y que formará parte de E1		
T1.1.2	**Preparación de datos. Protocolos y conectores**	M1-M12
Resultados previstos		
• Documento que describe detalladamente los trabajos realizados en la tarea y que forma parte de E1 • Datos que permiten desarrollar los algoritmos de operación y planificación		
T1.2	**Preparación y armonización de los datos**	M1-M12
Resultados previstos		
• Documento que describe detalladamente los trabajos realizados en la tarea y que forma parte de E1		
T1.3	**Conexión y compartición en Espacios de datos**	M1-M19
Resultados previstos		
• Documento que describe detalladamente los trabajos realizados en la tarea y que forma parte de E2.1 • Elementos que forman parte de E2.2		
T1.4.1	**Orquestación y despliegue de servicios en contenedores**	M1-M19
Resultados previstos		
• Documento que describe detalladamente los trabajos realizados en la tarea y que forma parte de E2.1 • Elementos que forman parte de E2.2		
T1.4.2	**Despliegue de modelos cloud en el edge**	M1-M19
Resultados previstos		
• Documento que describe detalladamente los trabajos realizados en la tarea y que forma parte de E2.1 • Elementos que forman parte de E2.2		
T4.1	**Machine Learning como servicio (MLaaS)**	M1-M21
Resultados previstos		
• Documento que describe detalladamente los trabajos realizados en la tarea y que forma parte de E5.1 • Elementos que forman parte de E5.2		
T4.3	**Integración de servicios IA**	M1-M21
Resultados previstos		
• Documento que describe detalladamente los trabajos realizados en la tarea y que forma parte de E5.1 • Elementos que forman parte de E5.2		

T4.4	Gestión del ciclo de vida de un proceso de AI (MLOps)	M4-M21
Resultados previstos		
• Documento que describe detalladamente los trabajos realizados en la tarea y que forma parte de E5.1 • Elementos que forman parte de E5.2		

Responsable del contrato
Mònica Aragüés Peñalba, profesora agregada de UPC

Barcelona a 19 de mayo de 2023