

PLIEGO DE CLÁUSULAS TÉCNICAS PARTICULARES QUE HA DE REGIR EL CONTRATO DE SERVICIOS INTEGRALES PARA LA CONFIGURACIÓN DE UNA PLATAFORMA MLOPS CON CASOS DE USO REALES EN EL MARCO DEL PROYECTO DE REDES INTELIGENTES PARA SERVICIOS DE 6G MEJORADOS (INES), DEL PROGRAMA UNICO I+D 6G 2023, EN EL MARCO DEL PLAN DE RECUPERACIÓN, TRANSFORMACIÓN Y RESILIENCIA – FINANCIADO POR LA UNIÓN EUROPEA – NEXT GENERATION EU, A ADJUDICAR POR PROCEDIMIENTO ABIERTO SIMPLIFICADO CON PLURALIDAD DE CRITERIOS.

EXPEDIENTE L17-2025

ÍNDICE

1.	OBJETO DEL CONTRATO	2
2.	DESCRIPCIÓN DE LOS TRABAJOS A REALIZAR Y CASO DE USO.....	4
3.	DESCRIPCIÓN DE LA PLATAFORMA.....	5
4.	CONDICIONES DE PRESTACIÓN DEL SERVICIO	11
4.1.	MEDIOS MATERIALES Y PERSONALES.....	11
4.2.	LUGAR DE PRESTACIÓN DEL SERVICIO Y DEDICACIÓN MÍNIMA.....	13
4.3.	DURACIÓN DEL CONTRATO	13
5.	DIRECCIÓN Y SEGUIMIENTO DEL CONTRATO. CONTROL DE LA CALIDAD DEL SERVICIO.	14
	ANEXO I: COORDINACIÓN EN MATERIA DE PREVENCIÓN DE RIESGOS LABORALES	15

PLIEGO DE CLÁUSULAS TÉCNICAS PARTICULARES QUE HA DE REGIR EL CONTRATO DE SERVICIOS INTEGRALES PARA LA CONFIGURACIÓN DE UNA PLATAFORMA MLOPS CON CASOS DE USO REALES EN EL MARCO DEL PROYECTO DE REDES INTELIGENTES PARA SERVICIOS DE 6G MEJORADOS (INES), DEL PROGRAMA UNICO I+D 6G 2023, EN EL MARCO DEL PLAN DE RECUPERACIÓN, TRANSFORMACIÓN Y RESILIENCIA – FINANCIADO POR LA UNIÓN EUROPEA – NEXT GENERATION EU, A ADJUDICAR POR PROCEDIMIENTO ABIERTO SIMPLIFICADO CON PLURALIDAD DE CRITERIOS.

EXPEDIENTE L17-2025

1. OBJETO DEL CONTRATO

La Fundación IMDEA Networks (en adelante, la Fundación) invierte en laboratorios y equipos de testeo de última generación para apoyar la investigación de vanguardia en redes de computación. Esto dota a la Fundación de la capacidad de transformar la investigación en productos y servicios de alto valor añadido.

El objeto del contrato es la prestación del servicio integral para la configuración de una plataforma MLOPS con casos de uso reales.

- Antecedentes:

El avance en tecnologías de Inteligencia Artificial, especialmente en modelos de lenguaje de gran escala (LLMs) y técnicas de aprendizaje automático (ML), está generando un cambio estructural en la forma en que se gestionan tanto los procesos administrativos como las actividades de investigación, particularmente en disciplinas como la ciencia de datos aplicada a redes de comunicaciones y protocolos. Estas tecnologías permiten automatizar tareas que anteriormente requerían una alta dedicación manual, además de aumentar significativamente la eficiencia, la fiabilidad y la capacidad de adaptación de los sistemas de trabajo.

En el ámbito administrativo, LLMs y modelos de ML ya permiten reforzar los controles tradicionales, como la doble validación manual, a través de verificaciones automáticas que detectan incoherencias, omisiones o desviaciones frente a requisitos normativos o procedimentales. Esto se traduce en una reducción de errores humanos, una mejora en la trazabilidad documental y una mayor alineación con las exigencias de cumplimiento. Al automatizar la clasificación, extracción y validación de datos en informes, formularios y justificantes, estas herramientas descargan al personal de tareas repetitivas y les permiten concentrarse en funciones de mayor valor añadido, como la supervisión estratégica o la planificación operativa, al margen de la propia tarea de validación de las justificaciones.

Sin embargo, el potencial de estas tecnologías va mucho más allá del ámbito puramente administrativo. En el contexto de la investigación en ciencia de datos, especialmente en áreas relacionadas con redes y protocolos de comunicación, los LLMs y los modelos de ML se posicionan como herramientas clave para acelerar el análisis de grandes volúmenes de datos, la identificación de patrones, la simulación de tráfico y la optimización de algoritmos. Por ejemplo, en el estudio de redes 5G/6G, los modelos pueden detectar anomalías de comportamiento en



paquetes, predecir congestiones, o incluso sugerir mejoras en el diseño de protocolos de control. La capacidad de generar conocimiento a partir de datos no estructurados —como logs, publicaciones, o resultados experimentales— también facilita una mayor integración entre teoría y práctica.

Otro aspecto fundamental es su capacidad para adaptarse dinámicamente a entornos cambiantes. Tanto en la gestión de convocatorias de financiación como en entornos experimentales de redes, las condiciones, los requisitos y los marcos regulatorios evolucionan con rapidez. Los sistemas basados en IA pueden actualizarse de forma incremental y aprender de nuevas entradas sin necesidad de ser reprogramados desde cero, lo que les otorga una enorme ventaja en entornos sujetos a cambio constante. Esto resulta especialmente valioso cuando se trata de mantener la conformidad con las bases regulatorias de las ayudas, el seguimiento de hitos, o la actualización de métricas en proyectos colaborativos.

Además, la trazabilidad y auditabilidad que ofrecen los modelos de IA bien implementados resultan esenciales tanto para la rendición de cuentas en entornos administrativos, como para la reproducibilidad científica en el ámbito de la investigación. Gracias al registro automatizado de procesos, fuentes de datos, parámetros utilizados y resultados obtenidos, se facilita una revisión posterior más rigurosa y transparente, tanto desde el punto de vista de la gestión pública como del avance del conocimiento científico.

En resumen, la integración de LLMs y ML en procesos administrativos y científicos representa un salto cualitativo hacia una forma de trabajo más eficiente, inteligente y sostenible. Lejos de reemplazar al personal humano, estas herramientas actúan como asistentes inteligentes que amplían sus capacidades y reducen su carga operativa. En el ámbito administrativo, aportan robustez al cumplimiento normativo y a la eficiencia documental. En el entorno investigador, permiten explorar con mayor profundidad conjuntos de datos complejos, validar hipótesis y acelerar los ciclos de innovación. Su adopción no es una cuestión de moda tecnológica, sino una respuesta natural a la creciente complejidad de los datos, los procesos y las exigencias del entorno actual.

- **Objetivo y objeto de contratación:**

El servicio a contratar consiste en el diseño, despliegue y puesta en marcha de una plataforma de MLOps (Machine Learning Operations) orientada a gestionar de forma eficiente, segura y escalable todo el ciclo de vida de desarrollo de sistemas basados en aprendizaje automático e inteligencia artificial. Esta plataforma abarcará desde la integración de datos iniciales y el entrenamiento de modelos, hasta su validación, puesta en producción, monitorización continua y mejora iterativa a partir del feedback de los usuarios finales y el rendimiento en entornos reales.

El servicio incluye la adopción de mecanismos robustos de gobernanza de datos y modelos, que permitan establecer políticas claras de control de versiones, auditoría, trazabilidad, explicabilidad (explainability), reproducibilidad y cumplimiento normativo (compliance), asegurando que tanto los conjuntos de datos como los modelos sean gestionados de forma ética y conforme a los estándares vigentes en materia de protección de datos y seguridad.

Asimismo, se abordará la optimización de la infraestructura tecnológica incluyendo recursos de computación, almacenamiento, redes y escalabilidad en entornos locales (on premise), garantizando una correcta asignación de recursos según las necesidades de cada etapa del ciclo de vida. El servicio contempla también la configuración e integración de un orquestador adecuado

que permita gestionar el ciclo de vida de los modelos en producción, ejecutar pipelines automatizados y facilitar la integración continua (CI/CD) y el despliegue continuo (CD).

En conjunto, esta solución proporcionará una plataforma adaptable, modular y sostenible, capaz de acelerar la innovación, reducir los tiempos de desarrollo, y facilitar el mantenimiento y evolución de los modelos mediante un enfoque basado en la automatización, la colaboración entre equipos de Soporte Investigación e IT, y el aprovechamiento eficiente de los recursos técnicos disponibles.

Se requerirá el desarrollo de **al menos. un caso de uso real o prueba de concepto que explore todas o el mayor número posible de las capacidades de la plataforma.**

La prueba de concepto será desplegada en el contexto del área de cumplimiento normativo (compliance), con el fin de evaluar su eficacia en la mejora de los mecanismos de control interno.

Se desarrollarán **modelos para validar la elegibilidad de los gastos desde un punto de vista semántico, descargando al equipo de Soporte Investigación de tareas repetitivas, contribuyendo a una gestión más ágil y conforme con la normativa propia de cada una de las Ayudas**, incorporando controles automatizados adicionales, exhaustivos y de muy bajo coste operativo.

Al margen de automatizar las tareas, se pretende crear una infraestructura sólida para integrar inteligencia adaptativa en procesos administrativos críticos, minimizando riesgos y **mejorando la transparencia institucional, asegurando que los procesos se adapten de manera continua a las nuevas regulaciones**, manteniendo una gestión rigurosa y confiable conforme a los estándares legales y de cumplimiento establecidos y a un futuro incremento de exigencias en materia de cumplimiento. Estas soluciones permitirán también aplicar una inteligencia preventiva a los sistemas de control interno, mejorando su capacidad de detección temprana.

Para poder validar las capacidades abajo descritas que no hayan sido utilizados durante el caso de uso o prueba de concepto, se desarrollarán escenarios que simulen la ejecución de dichos procesos o funcionalidades de las capas.

Asimismo, **se valorarán aplicaciones desarrolladas por el Licitador y cedidas a la Fundación**, que exploten las capacidades de la Plataforma y puedan contribuir a la automatización de los procesos administrativos. Por mencionar algunos a título ilustrativo, pero no exhaustivo: validación de facturas, contratos, elaboración de pliegos técnicos y administrativos, registro automático de datos en sitios web gubernamentales desprovistos de APIs usando técnicas de Robotics Process Automation. También revisión y valoración de ofertas en procesos de licitación y con contratos menores, evaluación inteligente de documentación curricular, análisis de perfiles profesionales, valoración automática de CV, cartas de motivación y referencias, etc. En estos últimos casos, la IA sólo asistirá, estando la capacidad de resolver en manos del Departamento correspondiente de Soporte a Investigación.

2. DESCRIPCIÓN DE LOS TRABAJOS A REALIZAR Y CASO DE USO

El alcance de los servicios objeto de este contrato comprende:

1. Servicio de diseño técnico y funcional de la Plataforma.

2. Servicio de configuración, instalación y puesta en marcha de la Plataforma y de desarrollo del caso de uso.
3. Servicio de transferencia de tecnología y conocimiento.
4. Servicio de Mantenimiento de la Plataforma y soporte al equipo de IT y a los usuarios finales de Soporte a Investigación.
5. Definición de la estrategia de gobernanza del dato para el proyecto.

Los licitadores deberán presentar una descripción detallada de dichos servicios y los entregables de los mismos, así como un cronograma de trabajo.

CASO DE USO

La Fundación tramita anualmente más de 50 líneas de financiación activas, procedentes de convocatorias regionales, nacionales, europeas e internacionales. Este panorama de gran complejidad normativa exige un conocimiento profundo de las bases reguladoras y una revisión exhaustiva de la elegibilidad de los gastos. Actualmente, estas tareas se ejecutan de forma manual con gran inversión de tiempo, y con riesgo de interpretaciones erróneas que pueden derivar en el reintegro del gasto.

Para reducir dichos riesgos y aumentar la eficiencia, la prueba de concepto incluirá el desarrollo de modelos de inteligencia artificial (machine learning) que analicen automáticamente los justificantes de gasto y su contexto contable, cotejándolos con las reglas de elegibilidad extraídas de las bases reguladoras. Los modelos podrán, además, proponer asignaciones contables adecuadas, detectar posibles incoherencias semánticas y validar si un gasto se ajusta a los criterios establecidos por el ente financiador. Esta automatización supondrá un control adicional y una mayor garantía de cumplimiento.

Como elemento diferenciador, se desarrollará también un chatbot especializado en normativa de gastos y justificaciones, integrado en el sistema de gestión, que permitirá a los usuarios realizar consultas interactivas sobre la naturaleza de un gasto, su posible elegibilidad y requisitos asociados, descargará significativamente al equipo de Soporte a Investigación administrativo, permitiendo que focalicen sus esfuerzos en tareas de mayor valor añadido, al tiempo que acelerará los tiempos de respuesta a los Investigadores, aumentando las productividad en toda la cadena.

El chatbot utilizará técnicas de procesamiento del lenguaje natural (NLP) y se nutrirá de la documentación histórica de ayudas, resoluciones, consultas a entidades financiadoras y normativa vigente, funcionando como asistente digital.

Además, el sistema proporcionará trazabilidad completa de los procesos de validación, conservará las decisiones tomadas y permitirá ajustes y reentrenamientos automáticos de los modelos conforme se detecten nuevos patrones o cambios normativos, asegurando así un sistema adaptable y vivo.

Esta iniciativa permitirá una gestión más ágil, segura y transparente, alineada con los principios de buena gobernanza y rendición de cuentas que demanda la normativa nacional y europea en el uso de fondos públicos. Además, sentará las bases para una futura extensión modular a otros procesos administrativos de la Fundación.

3. DESCRIPCIÓN DE LA PLATAFORMA

La solución permitirá industrializar el ciclo de vida de los modelos de aprendizaje automático, desde la ingesta de datos hasta la monitorización y el reaprendizaje automático, garantizando trazabilidad, reproducibilidad y cumplimiento normativo: RGPD, Esquema Nacional de Seguridad, AI Act, etc.

En líneas generales:

- Deberá eliminar o reducir al mínimo la intervención manual y acelerar la implementación de nuevos modelos.
- Garantizar la reproducibilidad, la gobernanza y el control de versiones de extremo a extremo para los datos y los modelos.
- Mejorar la calidad, la solidez y la equidad de los modelos mediante la validación y la explicabilidad automatizadas.
- Establecer un ciclo de retroalimentación / feedback continuo (ingesta, etiquetado automático, reentrenamiento) para hacer crecer y perfeccionar los conjuntos de datos y las inferencias.
- Integrar controles de privacidad y cumplimiento (privacidad diferencial, cifrado, coordinación del DPO) a lo largo del ciclo de vida del modelo.

Los Licitadores deberán presentar una definición a alto nivel de la arquitectura de la Plataforma, que contemple las siguientes capas o planos:

Capa o plano de ingesta y preprocesamiento

- Soportará múltiples tecnologías, dispositivos y mecanismos de captura de información, garantizando la correcta integración y procesamiento de datos provenientes de múltiples fuentes. Deberá disponer de conectores para: Excel, csv, json, MSSQL/PostgreSQL, PDF, imágenes.
- Realizará la extracción automatizada de datos públicos desde sitios web para alimentar modelos de análisis inteligente en aprendizaje o en inferencia.
- Permitirá la conexión a distintos orígenes de datos, tanto internos como externos a la solución mediante interfaces de acceso basadas en estándares abiertos o de amplia aceptación como, por ejemplo: API REST, JDBC, ODBC o tecnología equivalente.
- Deberá ser capaz de integrar y gestionar datos estructurados, no estructurados y semiestructurados.
- Proporcionará una solución para la construcción y ejecución de procesos ETL/ELT desde la Capa de adquisición hasta la Capa de Almacenamiento o Explotación.
- Entre los pasos de preprocesamiento figurarán: Escalado, normalización, agrupación/bucketization, PCA, vectorización de texto a embeddings, redimensionamiento de imágenes, ingeniería de características/feature engineering
- Incorporar datos incompletos, permitiendo estimar y completar datos ausentes utilizando métodos estadísticos o modelos predictivos, para mantener la integridad de los datos sin eliminar registros incompletos.
- Para los modelos de aprendizaje supervisado, posibilitará el etiquetado automático.
- Mostrará un gráfico de las diferentes versiones de preprocesamiento o similar que permita interpretar mejor el comportamiento de dicho preprocesamiento a través del orquestador (ver Capa o plano de ejecución y orquestación).

Capa o plano de almacenamiento

- Se encargará de proveer la funcionalidad de persistencia del dato, apoyándose sobre la plataforma de Cloud computing OpenNebula.

- La plataforma incluirá capacidades almacenamiento tipo DataLake, abajo descrita, para almacenar grandes volúmenes de datos en formato nativo, estructurados, semi-estructurados o sin ninguna estructura, permitiendo que los datos almacenados en el mismo sean provisionados hacia otras capas de la Plataforma.
- La arquitectura del Data Lake permitirá el procesamiento de datos tanto en streaming como batch.
- Los licitadores deberán incluir en su oferta un sistema de copia de seguridad y recuperación adecuado al objeto del contrato, en el caso de que los mecanismos proporcionados por OpenNebula sean insuficientes para una recuperación con el grado de resolución apropiado para garantizar la persistencia de una versión concreta de componentes individualizados de la Plataforma, así como restaurar la Plataforma o servicios concretos a partir de las copias realizadas.
- Organizará el DataLake en las siguientes capas lógicas.
 - o Capa cruda (raw layer): Donde se almacenarán los datos brutos, procedentes de múltiples fuentes en su formato nativo, sin ningún procesamiento (raw data). Esta capa actúa como repositorio de datos históricos en su estado original.
 - o Capa depurada o curada (curated layer): Donde se almacenan los datos que han sido preprocesados, limpiados, validados y estructurados. Esta capa facilita su posterior análisis, integración y explotación por parte de sistemas intermedios en el pipeline.
 - o Capa de consumo (consumption layer): Donde se alojan los datos preparados y organizados específicamente para su uso por aplicaciones analíticas, visualizaciones, modelos de machine learning u otros servicios de explotación de datos, orientados a usuarios finales.

Capa o plano de orquestación y ejecución

- Coordinará y automatizará las tareas del ciclo de vida del modelo.
- Gestionará la ejecución de pipelines complejos (ETL, entrenamiento, validación, inferencia).
- Integrará y optimizará los recursos computacionales, almacenamiento y redes, contando con el **autoescalado nativo de Kubernetes y supervisión de la producción**.
- Gestionará las dependencias entre tareas, programación de tareas y tolerancia a errores.

Capa o plano de entrenamiento y ajuste de modelos

- Dispondrá de cargadores o generadores de conjuntos de datos.
- Permitirá el ajuste automático o selección heurística de hiper-parámetros.
- Capturará todos metadatos de entrenamiento, permitiendo la trazabilidad de decisiones y registro de artefactos, métricas y explicaciones.
- Posibilitará la optimización del modelo basada en la curva ROC y el área bajo la curva (AUC), evaluando el equilibrio entre sensibilidad y especificidad, para clases desbalanceadas, permitiendo automatizar el ajuste de hiper-parámetros y la selección de modelos con mejor capacidad discriminativa.
- Proveerá técnicas de data augmentation para mejorar la generalización del modelo y reducir el sobreajuste (overfitting).
- Ofrecerá diferentes estrategias para el despliegue automático óptimo del proceso de entrenamiento en el cluster, aprovechando al máximo el paralelismo y las GPUs disponibles.
- Facilitará la transferencia de know-how, poda/pruning de la red neuronal y destilación/distillation de conocimiento.

Capa o plano de explotación o inferencia

- Facilitará el consumo de los datos almacenados en la Plataforma DataLake, así como los metadatos.
- Proveerá herramientas de análisis e informes de Inteligencia de Negocio/Business Intelligence, visualización de datos, generación de cuadros de mando/dashboards e informes de analítica avanzada que haga uso de los modelos de IAs y de algoritmos desarrollados por la Fundación.
- Proveerá firmas de modelos estables explicitando sus características: dimensiones, tipos de datos de entrada y salida, formato, etc., facilitando la trazabilidad de los modelos,
- Proporcionará soporte para despliegues: pruebas A/B, implementaciones canary e híbridas.
- Permitirá aplicar técnicas de cuantización/quantization, para el despliegue de modelos simplificados, posibilitando una inferencia más rápida y menor consumo de energía, siempre que la respuesta del modelo cuantizado sea lo suficientemente precisa.
- Ofrecerá diferentes estrategias para el despliegue automático óptimo del proceso de inferencia en el cluster, aprovechando al máximo el paralelismo y las GPUs disponibles.
- Proveerá herramientas para el análisis de hipótesis (what-if tools).

Integrará herramientas SHAP, LIME o equivalentes para ayudar a interpretar cómo las características de entrada individuales influyen en las predicciones de un modelo, con el fin de validar modelos, detectar sesgos y cumplir con la normativa, haciendo que las decisiones de IA sean más transparentes y fiables.

Capa o plano de feedback

- El sistema incorporará una capa de observabilidad y retroalimentación/feedback capaz de detectar de forma automática deriva de datos y modelos (en base al cálculo automático de métricas de clasificación y regresión), degradación de rendimiento y sesgos, correlacionar eventos y generar alertas inteligentes. Dichas alertas podrán activar respuestas automatizadas como pipelines de reentrenamiento, conforme a reglas configurables, con mecanismos de interrupción controlada/circuit breakers y controles para evitar oscilaciones indeseadas.
- Se exigirá visibilidad de extremo a extremo mediante métricas, registros y trazas que faciliten el diagnóstico y la toma de decisiones.
- Facilitará también la canalización de retroalimentación/feedback por parte de los usuarios a través de algún mecanismo de interacción con ellos y a través de consultas a las Bases de Datos de los sistemas transaccionales: IdiNET (ERP), D-Space (producción científica), etc.

Capa o plano de gobierno

- Englobará todas las funcionalidades relacionadas con el gobierno activo de los datos y modelos:
 - o Permitirá catalogar la información de todas las fuentes de datos de la capa de almacenamiento mediante metadatos o atributos, permitiendo acciones de perfilado de los datos, linaje, versionado, clasificación y establecimiento de ontologías, para posibilitar auditorías para cumplir con el RGPD, , Esquema Nacional de Seguridad, AI Act, etc.
 - o Ofrecerá la trazabilidad de datos desde la fuente hasta el modelo entrenado.

- Posibilitará el registro automático de artefactos: modelos, conjuntos de datos/datasets, scripts, parámetros, métricas, logs, outputs, con su correspondiente versionado.
- Realizará la validación formal previa al despliegue, con criterios cuantitativos (métricas mínimas) y cualitativos (auditoría).
- Alertará de la expiración de modelos y realizará control de vigencia.
- En la medida de lo posible, definirá reglas de calidad sobre las diferentes entidades gobernadas para mostrar de forma intuitiva el resultado de dichas reglas, así como para alterar el comportamiento de procesos en función de las mismas.
- Indexado y búsqueda: La Plataforma deberá proporcionar capacidades de búsqueda avanzada en la información, tanto estructurada como no estructurada.

Como se ha enunciado en el apartado 2 del presente pliego, la oferta del Licitador deberá incluir una propuesta de Estrategia de Gobernanza del Dato para el proyecto, que deberá describir, a alto nivel, el tratamiento de los diferentes tipos de datos (por fuentes, dispositivos, categorías, etc.), abordando con carácter ejemplificativo y no exhaustivo:

- Establecimiento de Roles y responsabilidades: alineadas con las indicaciones contenidas en este documento.
- Política de calidad de datos: normas para garantizar la precisión, integridad y consistencia de los datos. Establecer procesos para evaluar y monitorizar de forma continua la calidad de los datos y procesos de gestión asociados.
- Estándares de datos: definir estándares para la nomenclatura, formato y almacenamiento de datos.
- Gestión de ciclo de vida del dato: definiendo política de retención y eliminación de datos y el control sobre la modificación y creación de archivos.
- Seguridad: alineado con las especificaciones en materia de seguridad contenidas en el presente documento, se definirán los mecanismos para proteger los datos contra acceso no autorizados, brechas de seguridad y cumplimiento normativo.
- Capacitación y concienciación.
- Monitorización y auditoría: estableciendo las medidas y su periodicidad
- Cumplimiento normativo: RGPD, Esquema Nacional de Seguridad, AI Act, etc.

Capa o plano de administración

- Dotará de herramientas para la monitorización y el control de las actividades de los usuarios.
- Permitirá configurar alertas de forma que se pueda notificar a los administradores cuando ciertos indicadores superen los umbrales configurados.
- Permitirá acceder a los logs de la plataforma y permitir el almacenaje y búsqueda de los mismos.
- Permitirá en colaboración con la capa de orquestación y ejecución la distribución de forma óptima y eficiente de las cargas de trabajo de “ejecución larga”, como las generadas por los procesos de aprendizaje, que requieren de un alto tiempo de disponibilidad de recursos del clúster, tanto de forma manual, como de forma automática.
- Permitirá la administración, configuración y despliegue de los módulos y servicios desde un panel unificado, accesible de manera gráfica.

Capa o plano de seguridad

- Proveerá los medios para la gestión segura de secretos, aplicación y verificación de políticas.

- Permitirá la anonimización/anonimization de los datos antes de su persistencia.
- Incorporará herramientas de control de accesos y enmascaramientos u ofuscación de datos. Además del control de acceso al dato, debe controlarse desde esta capa la representación del dato que se le entrega a la identidad que lo consulta, pudiendo definir si la información está enmascarada, oculta o cifrada.
- La Plataforma presentará garantías técnicas de protección y de seguridad de acceso a la información tanto en el almacenamiento como en su transmisión, y estará alineada con los requerimientos del Esquema Nacional de Seguridad.
- La Plataforma deberá asegurar el acceso, integridad, disponibilidad, autenticidad, confidencialidad, trazabilidad y conservación de los datos, información y servicios utilizados en medios electrónicos que sean objeto del procedimiento de contratación.

Capa o plano de interoperabilidad

Deberá facilitar la comunicación, interacción y compatibilidad entre diferentes sistemas, aplicaciones y dispositivos. Esta capa permite el intercambio de datos y la funcionalidad conjunta de diversos componentes del sistema y de la interacción con terceros.

Al tratarse de una plataforma MLOps (y LLMOps), se deberá abordar la integración completa con la plataforma de desarrollo colaborativo de la Fundación, GitLab, permitiendo una gestión centralizada, trazable y automatizada del ciclo de vida de las aplicaciones.

Esta integración deberá habilitar la ejecución de procesos de integración y entrega continuas (CI/CD) mediante la configuración de GitLab Runners desplegados sobre la propia infraestructura de la plataforma, garantizando así la coherencia entre los flujos de trabajo de desarrollo, validación y puesta en producción de modelos y otros artefactos.

Se asegurará la compatibilidad con los entornos de contenedores y herramientas de orquestación. También se documentará y validará la integración mediante la demostración de los flujos completos de CI/CD funcional sobre la plataforma desarrollados para el caso de uso, incluyendo tests automatizados, control de versiones, validación de modelos y su posterior despliegue en entorno controlado.

Las características de la Plataforma de base HW, SW de base (OS) que soportará la Plataforma MLOps, LLMOps está constituida por:

- 1 Management Server DELL PowerEdge R660
- 3 Nodos del cluster, cada uno de ellos: DELL 760 con 128 cores, 512 GB de RAM y 4 × NVIDIA L4 Tensor Core GPUs
- 1 unidad de almacenamiento: DELL PowerVault ME5024 con 5 × 3.84 TB SAS SSDs (Read-Intensive), con 8 × 25G iSCSI SFP28 (4 por controlador)
- 2 switches Extreme Networks 7520-48Y-8C con 48 × SFP28 (1/10/25G) y 8 × QSFP28 uplinks (40G/100G)
- Todas las conexiones entre componentes en base al estándar 25 Gigabit ethernet
- Cluster Software Stack: OpenNebula Latest Enterprise Edition: 7.0.0 "Phoenix" EE

El Licitador deberá emplear, para el desarrollo e implementación de la plataforma, herramientas de código abierto (open source) de amplia difusión, madurez y respaldo por parte de la comunidad técnica y científica, garantizando así la interoperabilidad, escalabilidad y sostenibilidad de la solución. La selección de dichas herramientas deberá responder a criterios técnicos y funcionales contrastados, evitando dependencias propietarias innecesarias y facilitando la transparencia, trazabilidad y audibilidad de los procesos. Asimismo, se exigirá que las herramientas seleccionadas cuenten con

actualizaciones activas y documentación técnica suficiente para su mantenimiento, integración y evolución en el marco de la solución ofertada.

Se podrán considerar herramientas propietarias de fabricantes de reconocido prestigio, siempre que estén basadas en estándares abiertos y en tecnologías de código abierto (open source) ampliamente aceptadas. No obstante, su inclusión deberá justificarse mediante una mejora sustancial y demostrable en términos de productividad, eficiencia en el desarrollo de nuevos modelos y casos de uso, así como en la reducción de costes operativos o de mantenimiento. Estas mejoras deberán estar respaldadas por métricas objetivas y comparativas, que acrediten que los beneficios obtenidos superan con claridad los costes asociados a las licencias.

4. CONDICIONES DE PRESTACIÓN DEL SERVICIO

4.1. MEDIOS MATERIALES Y PERSONALES

Los licitadores deberán comprometerse a dedicar o adscribir a la ejecución del contrato los medios personales y materiales suficientes para la correcta prestación del servicio y a respetar las medidas de seguridad en las instalaciones de la Fundación.

El equipo de trabajo estará formado por cuantos técnicos sean necesarios para la realización de los trabajos, y según la planificación que se fije.

El Adjudicatario será únicamente quien ejerza de modo efectivo, real y continuo, el poder inherente de dirección a todo empresario. Es responsabilidad de la empresa adjudicataria impartir todas las órdenes, criterios de realización del trabajo y directrices a sus trabajadores/as, siendo la Fundación del todo ajena a estas relaciones laborales y absteniéndose, en todo caso, de incidir en las mismas. La adjudicataria asumirá la negociación y pago de salarios, concesión de permisos, licencias y vacaciones, las sustituciones en caso de baja o ausencia, las obligaciones legales en materia de seguridad social, las obligaciones en materia de prevención de riesgos laborales, el ejercicio de la potestad disciplinaria, así como cuantos derechos y obligaciones se deriven de la relación contractual entre empleado y empleador. Corresponde asimismo a la empresa adjudicataria de forma exclusiva la vigilancia del horario de trabajo de los trabajadores, las posibles licencias horarias o permisos o cualquier otra manifestación de las facultades del empleador.

El Adjudicatario se compromete a retribuir adecuadamente al personal que asigne al presente servicio asumiendo de forma directa y no trasladable a la Fundación el coste de cualquier mejora en las condiciones de trabajo o en las retribuciones de dicho personal, tanto si dichas mejoras son consecuencia de decretos, convenios colectivos, pactos, o acuerdos de cualquier índole que puedan conllevar o conlleven equiparación a otras situaciones, como si lo fuera por la integración de los centros a otro organismo de la Administración Pública, de modo que en ningún caso podrán repercutir dichas modificaciones en un incremento del precio de adjudicación.

La empresa adjudicataria se asegurará que todos los trabajos incluidos dentro del alcance de los servicios objeto de la presente contratación, se ejecuten a través de los miembros del equipo mínimo de trabajo con solvencia técnica o cualificación suficiente para ello, ajustándose a los siguientes parámetros mínimos:

Jefe de Proyecto

Encargado de la planificación, seguimiento y coordinación global del proyecto. Asegura el cumplimiento de plazos, presupuesto y calidad, y actúa como interlocutor principal con el cliente. También se ocupa de la gestión de riesgos y la alineación funcional de todos los entregables.

Se requiere una experiencia acreditada de al menos cinco (5) años en el desempeño de tareas de diseño y dirección de proyectos TIC, de los cuales al menos tres (3) habrán sido en proyectos relacionados con la implementación de plataformas de analítica avanzada de datos.

Arquitecto MLOps

Responsable de definir la arquitectura técnica de la solución, asegurando su alineación con los principios MLOps (escalabilidad, trazabilidad, reproducibilidad, seguridad) y con la infraestructura y diseñar la integración de todos los componentes del ciclo de vida ML.

Se requiere una experiencia acreditada de al menos cinco (5) años en el desempeño de tareas como Arquitecto de soluciones IT, de los cuales al menos tres (3) habrán sido en proyectos relacionados con la implementación de plataformas de analítica avanzada de datos.

Ingeniero DevOps

Especialista en despliegue de infraestructura sobre Kubernetes y automatización de pipelines CI/CD. Responsable de asegurar la escalabilidad, reproducibilidad y observabilidad del entorno, así como la gestión de recursos GPU y la integración continua.

Se requiere una experiencia acreditada de al menos cinco (5) años en el desempeño de tareas de despliegue de infraestructuras en proyectos TIC, de los cuales al menos tres (3) habrán sido en proyectos relacionados con la implementación de plataformas de analítica avanzada de datos.

Ingeniero de Datos / Científico de datos

Encargado de diseñar e implementar pipelines de ingesta y transformación de datos. Garantizará la calidad, trazabilidad y versionado de los datos, además de construir y mantener las features desarrolladas para el entrenamiento e inferencia.

Se requiere una experiencia acreditada de al menos cinco (5) años en el desempeño de tareas de diseño e implementación de pipelines ETL, de los cuales al menos tres (3) habrán sido en proyectos relacionados con la implementación de plataformas de analítica avanzada de datos.

Ingeniero de aprendizaje de máquina / Científico de datos

Responsable del entrenamiento, evaluación y trazabilidad de modelos. Implementará experimentación reproducible, optimización de hiperparámetros y estrategias de despliegue seguro. También participará en el diseño del sistema de feedback y reaprendizaje automático, así como en el desarrollo de procesos para el gobierno del dato.

Se requiere una experiencia acreditada de al menos cinco (5) años en el desempeño de tareas de entrenamiento, evaluación y trazabilidad de modelos.

Ingeniero de integración & automatización del aseguramiento de calidad

Perfil de soporte transversal para el desarrollo de servicios backend (APIs, conectores, etc.), la integración con la plataforma OpenNebula y otros servicios externos, y la automatización de pruebas.

Se requiere una experiencia acreditada de al menos cinco (5) años en el desempeño de tareas de integración & automatización del aseguramiento de calidad.

La acreditación de los nombres y cualificación profesional de los miembros del equipo de trabajo que se hubiese comprometido a dedicar o adscribir, se llevará a cabo por el licitador que haya presentado la oferta económicamente más ventajosa, una vez calificada ésta como tal, y previo requerimiento del órgano de contratación.

La autorización de cambios en el personal adscrito al servicio, ya sean puntuales o no, y a propuesta del contratista, requerirá de las siguientes condiciones:

- Justificación escrita del cambio.
- Presentación de posibles candidatos, deberán tener una cualificación técnica igual o superior a la de la persona que se pretende sustituir.
- Aceptación de alguno de los candidatos por parte de la Fundación.

4.2. LUGAR DE PRESTACIÓN DEL SERVICIO Y DEDICACIÓN MÍNIMA

El licitador elaborará una oferta y un plan de trabajo en el que se informe del número de horas presenciales que considera necesarias para cubrir las necesidades expuestas, y el correcto funcionamiento de todas las instalaciones, así como de la cualificación profesional de los trabajadores que prestarán servicio.

El licitador nombrará un responsable del servicio. Éste se responsabilizará del cumplimiento del plan de trabajo, del buen funcionamiento del servicio, con reuniones periódicas con el responsable que designe la Fundación, además de informar de manera puntual a la Fundación de cualquier aspecto relevante.

Es necesario que la ejecución de los servicios objeto del contrato tenga en cuenta las normas, modelos, documentación, codificación, estándares y estructuras definidas en el área de IT de la Fundación. Toda la documentación presentada deberá estar en inglés.

El horario de prestación del servicio IT y de nuestro Equipo de Soporte a Investigación que participará en el caso de uso del proyecto será de 9:00 a 18:00, de lunes a viernes, ambos inclusive, según el calendario laboral de Madrid.

4.3. DURACIÓN DEL CONTRATO

Atendiendo a los requisitos establecidos en la normativa de concesión de las subvenciones en el marco del Plan de Recuperación, Transformación y Resiliencia el objetivo en términos de plazo del proyecto es terminar la ejecución antes del 31 de diciembre del 2025.

El mantenimiento de estos servicios estará incluido en la oferta.

5. DIRECCIÓN Y SEGUIMIENTO DEL CONTRATO. CONTROL DE LA CALIDAD DEL SERVICIO.

Para verificar el cumplimiento de los preceptos del contrato objeto de esta contratación se establece un sistema de control de calidad de la prestación del servicio basado en la designación de un Encargado de Seguimiento, que velará por el cumplimiento del contrato pudiendo comprobar al personal asignado por la empresa y su trabajo en todo momento. En caso de incumplimiento lo hará saber a la empresa Adjudicataria.

La Dirección, si lo estima conveniente, podrá requerir a la empresa adjudicataria que subsane en el plazo máximo de tres días los defectos, insuficiencias técnicas, errores materiales, omisiones o infracciones de preceptos reglamentarios que les sean imputables de acuerdo con las conclusiones del informe a que se refiere el párrafo anterior. Si transcurrido ese plazo las deficiencias no han sido corregidas, la Dirección podrá proceder con las penalizaciones que correspondan.

ANEXO I: COORDINACIÓN EN MATERIA DE PREVENCIÓN DE RIESGOS LABORALES

La finalidad de este protocolo es establecer un acuerdo de colaboración en materia de Prevención de Riesgos entre La Fundación y sus Empresas Colaboradoras, para el desarrollo de las obligaciones establecidas en el artículo 24 "Coordinación de Actividades Empresariales" de la LPRL y en el RD 171/2004 que lo desarrolla.

Dentro de la coordinación de actividades, se establece el sistema de acreditación y control, en materia de Prevención Riesgos, de los trabajadores pertenecientes a la entidad adjudicataria y sus subcontratas, que van a desarrollar los servicios contratados por la Fundación en el centro de trabajo propio de esta entidad.

La acreditación de los trabajadores es consecuencia de la planificación de la prevención resultante de la evaluación de riesgos de los trabajos contratados, que realizará la entidad adjudicataria. Como consecuencia de la misma, se acreditará si cada trabajador es "apto para su trabajo habitual" o, si fuera necesario, que es "apto para trabajos que impliquen un riesgo especial", indicando claramente cual es este riesgo especial. Además, será utilizada para conocer los nombres de las personas designadas por la entidad adjudicataria como responsables en materia de prevención de riesgos durante la ejecución de los servicios contratados por La Fundación.

La entidad adjudicataria evaluará y conservará toda la documentación necesaria para, en cualquier momento, garantizar y demostrar las aptitudes acreditadas para cada trabajador, independientemente de que éstos pertenezcan directamente a la entidad adjudicataria o a cualquiera de sus subcontratas.