

# E7.1 PLAN DE TRABAJO Y PLAN DE CALIDAD

KG4LLM - SERVICIOS PARA EL ENRIQUECIMIENTO DE MODELOS DE LENGUAJE CON GRAFOS DE CONOCIMIENTO (SER-21/23 OTT)

## Resumen

Este entregable incluye la versión definitiva del plan de trabajo elaborado a partir de la planificación descrita en la propuesta original de este proyecto. Incluye el plan de calidad, que define los procedimientos para asegurar la calidad de los entregables mediante el uso de plantillas, revisión por pares y gestión de los documentos. Finalmente, este entregable describe los procedimientos de comunicación y el proceso para gestionar los riesgos que potencialmente afecten el proyecto.

José Manuel Gómez Pérez

30 de Diciembre de 2023  
Expert.ai Language Technology Research Lab

## Historia de revisions

Revision	Date	Description	Author (Organisation)
0.1	26/12/2023	Primer draft completo	José Manuel Gómez
0.2	29/12/2023	Revisión, corrección y extensión	José Manuel Gómez
1.0	30/12/2023	Versión actualizada lista para enviar	José Manuel Gómez

## Contents

1	Introducción.....	4
2	Planificación Inicial.....	4
2.1	Paquetes de trabajo .....	4
2.2	Hitos del proyecto y cronograma.....	4
2.3	Listado de entregables .....	7
3	Plan de Calidad.....	7
3.1	Directrices de comunicación .....	8
3.1.1	Intercambio de documentos .....	8
3.1.2	Convención de denominación y numeración de documentos.....	9
3.1.3	Herramienta Colaborativa (MS Teams).....	10
3.2	Procedimiento de generación de informes.....	10
3.2.1	Entregables y documentos.....	10
3.2.2	Agenda y actas de las reuniones. ....	11
3.2.3	Reporte de gestión del proyecto .....	11
3.3	Gestión del Riesgo .....	12
3.3.1	Definiciones .....	12
3.3.2	Clasificación del riesgo.....	13
3.3.3	Organización y responsabilidades de la gestión de riesgos .....	13
3.3.4	Proceso de gestión del riesgo .....	13
3.3.5	Plan de Contingencia .....	16
3.3.6	Procedimiento de monitorización del riesgo .....	16
3.3.7	Tabla de riesgo.....	17
Apéndice A:	Descripción detallada del trabajo .....	18

## 1 Introducción

Este entregable incluye la planificación inicial del proyecto y el plan de calidad que garantiza que los resultados del proyecto sean de la más alta calidad y se ajusten a los objetivos del proyecto. El entregable ha sido generado como parte de la tarea T7.1 Progreso del proyecto, gestión del riesgo y aseguramiento de la calidad.

En la planificación inicial se incluye la duración, fecha de inicio y fin, los paquetes de trabajos en los que se agrupa las diferentes actividades y tareas, los hitos, el cronograma y los entregables. El plan de calidad, por otra parte, describe las directrices de comunicación entre el equipo de desarrollo y la dirección del proyecto, definiendo claramente los formatos y plantillas para los documentos que se generan, su denominación y el sistema que se usa para gestionar y compartir los documentos. El plan de calidad también incluye los procedimientos para la gestión de los entregables, la generación de agendas y actas para las reuniones, y los reportes de gestión del proyecto. Además, se describe el proceso de gestión del riesgo en el proyecto desde la identificación del riesgo, evaluación y planes de mitigación.

## 2 Planificación Inicial

La planificación del proyecto se deriva de la planificación que se presentó en el documento Calidad Técnica de la Oferta (sobre 2 de la licitación SER-2123-OTT). Esta planificación inicia en el mes de diciembre de 2023 y finaliza en el mes de junio de 2025 para un total de 19 meses.

### 2.1 Paquetes de trabajo

No se prevé ninguna modificación a nivel de los paquetes de trabajos (ver Ilustración 1) ni de las tareas originalmente planteadas. La descripción detallada de cada paquete de trabajo y sus tareas se presenta en el Apéndice A:

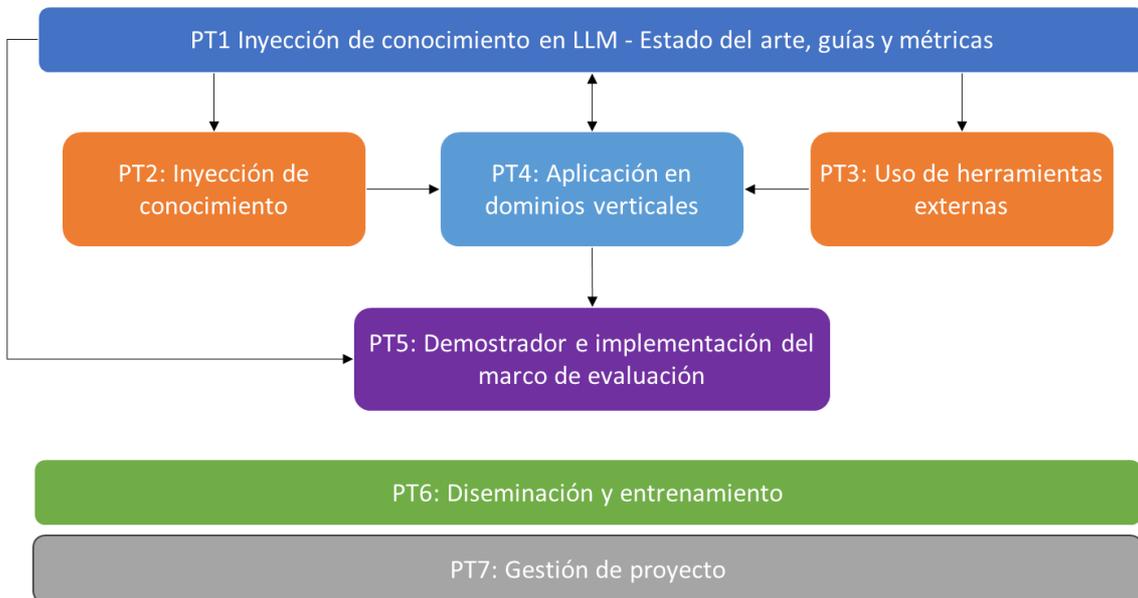


Ilustración 1. Estructura de paquetes de trabajo. Las flechas indican que la salida del PT de origen se usa en el PT de destino. Los PT 6 y 7 son transversales a todo el proyecto.

### 2.2 Hitos del proyecto y cronograma

La planificación de las tareas a lo largo del cronograma de trabajo tampoco sufre modificación alguna (ver Ilustración 2) ni los hitos del proyecto (ver Tabla 1 Ilustración 1)

*Tabla 1 Hitos y plazos parciales del programa de trabajo.*

<b>Hito</b>	<b>Descripción</b>	<b>Mes</b>	<b>Fecha</b>	<b>Fase</b>
1	Plan de trabajo inicial, incluyendo plan de calidad.	M1	31/12/2023	Inicio
2	Documento de estado de la cuestión de métodos para la inyección de conocimiento, guías con los enfoques para la inyección de conocimiento y métricas de evaluación de inyección de conocimiento y de medición de la adaptación del modelo de lenguaje.	M3	28/02/2024	Construcción 1
3	Versión inicial de los métodos.	M7	30/06/2024	Construcción 2 (MVP)
4	Versión final de los métodos y aplicación a los LLM para dominios verticales.	M13	31/12/2024	Construcción 3
5	Demostrador y marco de evaluación resultante de la aplicación de los LLM a los dominios verticales	M19	30/06/2025	Construcción 4

Paquetes de trabajo y tareas	Inicio	F. Const. 1			Fase Construcción 2				Fase de Construcción 3					Fase de Construcción 4					
	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19
<b>PT1 Inyección de conocimiento en LLM - Estado del arte, guías y métricas</b>																			
T1.1 Análisis actualizado del estado del arte			H2																
T1.2 Definición de guías con los enfoques propuestos para la inyección de conocimiento			H2																
T1.3 Diseño de una metodología de evaluación			H2																
T1.4 Selección de grafos de conocimiento estructurado y APIs																			
<b>PT2 Inyección de conocimiento procedente de grafos de conocimiento</b>																			
T2.1 Desarrollo de métodos para la extracción de subgrafos de conocimiento							H3						H4						
T2.2 Desarrollo de métodos para la adaptación de LLM mediante conocimiento estructurado							H3						H4						
T2.3 Inyección de estrategias de resolución de problemas							H3						H4						
T2.4 Desarrollo de métodos para la inyección de conocimiento durante preentrenamiento							H3						H4						
<b>PT3 Uso de herramientas externas</b>																			
T3.1 Desarrollo de métodos para habilitar el uso de herramientas externas							H3						H4						
T3.2 Desarrollo de métodos de recuperación de información factual							H3						H4						
T3.3 Desarrollo de métodos basados en propiedades no funcionales							H3						H4						
<b>PT4 Aplicación</b>																			
T4.1 Definición y análisis de dominios de aplicación y casos de uso																			
T4.2 Diseño de secuencias de pensamiento orientadas a resolución de problemas																			
T4.3 Adaptación a dominio y despliegue de LLM													H4						
T4.4 Extracción de grafos de conocimiento a partir de texto usando LLM													H4						
<b>PT5 Demostrador y marco de evaluación</b>																			
T5.1 Desarrollo del marco de evaluación de factualidad y sesgo																			
T5.2 Diseño, desarrollo y despliegue del demostrador																			H5
<b>PT6 Diseminación y entrenamiento</b>																			
T6.1 Web del proyecto, enlazada con la Plataforma INESData																			
T6.2 Publicaciones científicas																			
T6.3 Organización de eventos (hackathons, tutoriales, seminarios)																			
T6.4 Colaboración con otras iniciativas Europeas y en el marco de INESData																			
<b>PT7 Gestión de proyecto</b>																			
T7.1 Progreso del proyecto, gestión del riesgo y aseguramiento de la calidad	H1																		
T7.2 Gestión administrativa y financiera																			

Ilustración 2. Diagrama Gantt de KG4LLM. En verde se indica los meses en los que las tareas correspondientes generan entregables para los hitos del proyecto.

## 2.3 Listado de entregables

Como se mencionó al principio de este documento, los entregables y su fecha de entrega se mantienen con respecto a la propuesta.

No	Título del entregable	PT	Tipo <sup>1</sup>	Entrega
E7.1	Plan de trabajo inicial y plan de calidad	7	R	M1
E4.1	Definición y análisis de dominios de aplicación y casos de uso	4	R	M1
E6.1	Web del proyecto	6	S	M3
E1.1	Documento de estado de la cuestión, guías para la inyección de conocimiento en LLM y métricas	1	R	M3
E2.1v1	Métodos de inyección de conocimiento en LLM	2	S+R	M7
E3.1v1	Métodos de uso de herramientas externas por LLM	3	S+R	M7
E5.1	Marco de evaluación y recursos asociados	5	S+D+R	M7
E5.2	Diseño del demostrador	5	R	M10
E7.2v1	Informe periódico de la gestión del proyecto	7	R	M12
E2.1v2	Métodos de inyección de conocimiento en LLM	2	S+R	M13
E3.1v2	Métodos de uso de herramientas externas por LLM	3	S+R	M13
E4.2	LLM adaptados a dominio	4	M+R	M13
E4.3	Métodos de extracción de grafos de conocimiento a partir de texto mediante LLM	4	S+R+G	M13
E5.3	Demostrador y resultados del marco de evaluación	5	S+R	M19
E6.2	Informe de publicaciones, actividades de diseminación y colaboración	6	R	M19
E7.2v2	Informe periódico de la gestión del proyecto	7	R	M19

## 3 Plan de Calidad

El Plan de Calidad se centra en la gestión integral del proyecto, estableciendo directrices y procedimientos cruciales para garantizar la calidad de los resultados del proyecto. Este plan de calidad incluye Directrices de Comunicación que describen las estrategias clave para facilitar una comunicación eficiente, abarcando por tanto mecanismos de comunicación y de intercambio de documentos, nombramiento y numeración de los mismos y el uso de herramientas colaborativas para gestionar la documentación y la comunicación. Estas directrices sientan las bases para una colaboración transparente y un intercambio de información fluido entre los miembros del equipo, esenciales para el progreso sin contratiempos del proyecto.

Por otro lado, el plan de calidad cubre también la documentación y presentación de resultados del proyecto de forma estandarizada. Para ello se ha definido plantillas útiles para la creación de entregables y otros documentos, agendas de las reuniones y las actas correspondientes, y los informes de gestión del proyecto.

---

<sup>1</sup> R: Reporte, S: Software, D: Dataset, M: Modelos, G: Grafos de conocimiento.

El plan de calidad también incluye la Gestión de Riesgos que describe el proceso desde la identificación de los riesgos, su clasificación y los planes de mitigación. El Plan de Calidad asegura que el proyecto esté respaldado por sólidos protocolos de comunicación, generación de informes y gestión de riesgos, fundamentales para su éxito continuado.

### 3.1 Directrices de comunicación

Esta sección contiene un conjunto de mejores prácticas a seguir para facilitar el proceso de comunicación. La comunicación se realizará preferentemente a través de correo electrónico y el chat de la plataforma MS Teams.

Se ha creado una lista de correo para facilitar la circulación de información a todos los miembros del proyecto. La dirección utilizada es [kg4llm@expert.ai](mailto:kg4llm@expert.ai).

Las listas de correo se han creado con los contactos del proyecto, aunque se pueden actualizar según sea necesario en cualquier momento:

El asunto del mensaje comenzará con el acrónimo del proyecto KG4LLM (Knowledge graphs for Large Language Models) y opcionalmente el paquete de trabajo en cuestión, por ejemplo "PT1". Esto será de gran ayuda para identificar y clasificar fácilmente los mensajes.

Los correos electrónicos serán respondidos en un plazo máximo de cuatro días después de la recepción del correo original.

Se va a crear un equipo **Inesdata\_KG4LLM** en MS Teams para el proyecto donde se podrá tener conversaciones más fluidas sobre temas determinados que requieran una iteración online.

#### 3.1.1 Intercambio de documentos

Todos los documentos de texto intercambiados dentro del proyecto deben observar las siguientes reglas:

- Formato \*.docx/doc (Word o equivalente).
- Seguimiento de cambios activado.
- Una vez que el documento final haya pasado la revisión interna por pares, el gestor del proyecto de expert.ai que presente el documento a la dirección del proyecto generará el archivo PDF, debidamente protegido.
- Los archivos adjuntos no deben enviarse por correo electrónico sino colocarse en el repositorio del proyecto (MS Teams). Luego, la persona que ha subido el documento avisará al resto del equipo a través de correo electrónico, anunciando dónde se puede recuperar el documento.
- Se ha organizado una estructura lógica del repositorio para facilitar la recuperación de todos los documentos. Todos los participantes del proyecto seguirán utilizando esta estructura y crearán nuevos directorios de la misma forma lógica cuando sea necesario.
- Las presentaciones utilizarán el formato \*.pptx/ppt (o equivalente) según plantilla disponible en el repositorio.
- Todos los documentos para entregar a la dirección del proyecto, incluidas las presentaciones y los entregables finales, utilizarán únicamente el formato PDF, estarán debidamente protegidos y autorizarán únicamente su impresión, ninguna edición, ninguna copia y ninguna anotación.

- Los informes de avance del proyecto cuentan con plantillas específicas que están disponibles en el directorio “Templates” en el repositorio de ficheros del proyecto en Teams.

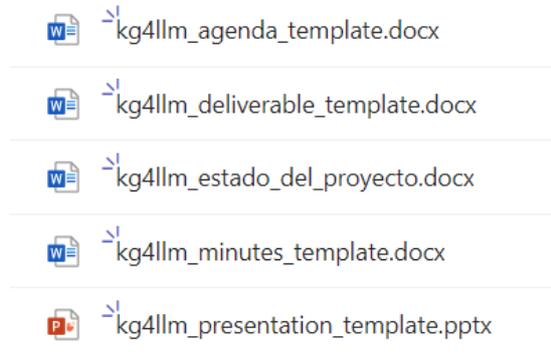


Ilustración 3. Plantillas definidas para la documentación del proyecto.

- Los entregables, informes breves de hitos intermedios y documentos deberán seguir el formato y estilos indicados en la plantilla disponible en la sección correspondiente del repositorio KG4LLM en MS Teams. Estas plantillas pueden evolucionar según las necesidades del proyecto.

### 3.1.2 Convención de denominación y numeración de documentos

Los entregables se etiquetan de la siguiente manera:

- R: Reporte
- S: Software
- D: Dataset
- M: Modelos
- G: Grafos de conocimiento.

Con respecto a la confidencialidad de los entregables y otros documentos, incluidas las presentaciones, se consideran los siguientes dos niveles de seguridad:

- PU: Uso Público. Sin restricciones de acceso (en formato PDF seguro).
- CO: Confidencial Sólo para diseminación interna y a la dirección del proyecto.

Para facilitar la navegación y el almacenamiento común en diferentes plataformas y SO, no se permiten espacios ni puntos en los nombres de los documentos y, en su lugar, se utilizará el carácter de subrayado “\_”.

Todos estos documentos serán nombrados y numerados según las siguientes reglas, con el fin de facilitar una rápida identificación e indexación:

KG4LLM\_<enum>\_<ename>\_<seg>\_v<ver>.pdf

Todos los nombres de los documentos comienzan con el acrónimo "KG4LLM" para facilitar la identificación con otros documentos del proyecto y generar conciencia sobre el proyecto entre un número de personas que descargarán los documentos del sitio web público. Los campos <enum> representan el código del entregable, <ename> el nombre del entregable tal como se

indica en el plan de entregables, <seg> se completa con las siglas relacionadas con el nivel de difusión (por ejemplo, co = confidencial, pu = público) y < ver> la versión del documento.

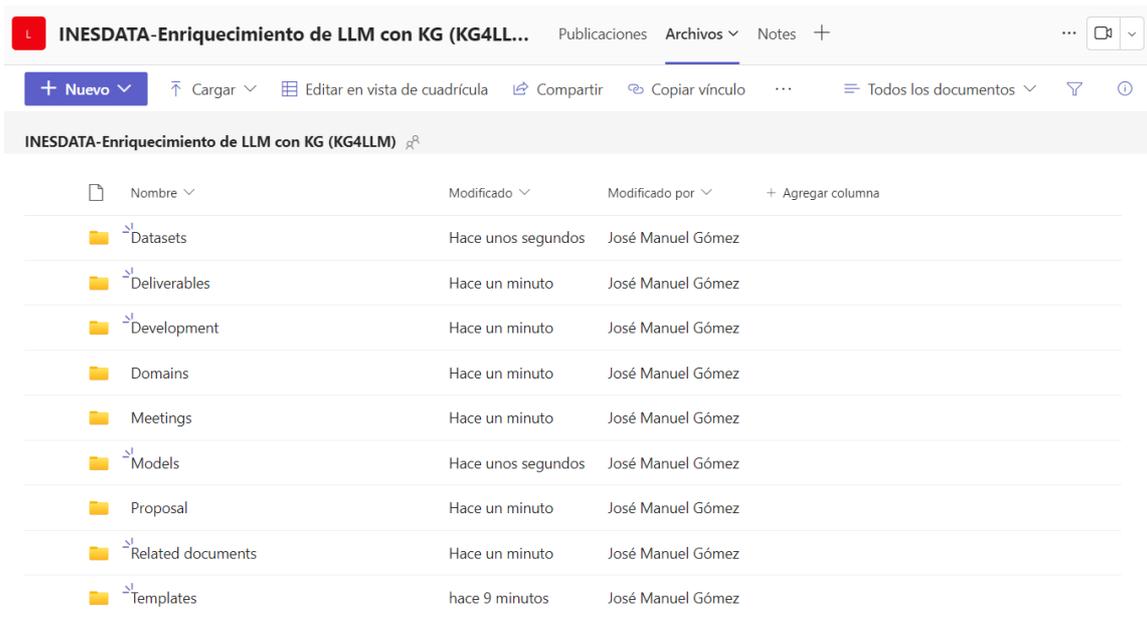
Las versiones 0.X indicarán que el documento aún es un borrador no aprobado por los revisores internos. El documento oficial que se enviará a la dirección de proyecto llevará la numeración v1.0. Las revisiones adicionales o nuevas ediciones de un entregable utilizarán el siguiente formato: v1.X, vY.X.

Por ejemplo, este entregable se presenta en un fichero llamado KG4LLM\_E7\_1\_Planificación\_y\_plan\_de calidad\_co\_v1.0.pdf

Para facilitar el trabajo/uso y localización de los documentos, todos los documentos se publicarán en el repositorio de MS Team.

### 3.1.3 Herramienta Colaborativa (MS Teams)

En el proyecto se usará MS Teams como repositorio de documentos y herramienta de comunicación por medio del chat. En la pestaña archivos del equipo del proyecto se tiene acceso a los ficheros gestionados en el proyecto.



Nombre	Modificado	Modificado por	+ Agregar columna
Datasets	Hace unos segundos	José Manuel Gómez	
Deliverables	Hace un minuto	José Manuel Gómez	
Development	Hace un minuto	José Manuel Gómez	
Domains	Hace un minuto	José Manuel Gómez	
Meetings	Hace un minuto	José Manuel Gómez	
Models	Hace unos segundos	José Manuel Gómez	
Proposal	Hace un minuto	José Manuel Gómez	
Related documents	Hace un minuto	José Manuel Gómez	
Templates	hace 9 minutos	José Manuel Gómez	

Ilustración 4. Repositorio de ficheros en MS Teams

## 3.2 Procedimiento de generación de informes

### 3.2.1 Entregables y documentos

Cualquier entregable o documento, incluidas las presentaciones, seguirá los procedimientos aquí definidos. La responsabilidad final de la calidad de los entregables reside en el equipo de revisión por pares que debe verificar la calidad de todos los entregables (sin incluir los informes de progreso periódicos), antes de la entrega final a la dirección del proyecto.

Los entregables estarán sujetos a una evaluación de pares. Al menos un experto de expert.ai que no participe en la tarea asociada al entregable lo revisará y dará feedback para la generación de la versión final.

El procedimiento de evaluación por pares y el envío del entregable a la dirección del proyecto se describe a continuación:

- El responsable de un entregable debe enviar (o cargar en el repositorio) una versión preliminar (borrador final, es decir, un documento completo y legible) del entregable al gestor del proyecto de expert.ai 2 semanas antes de la fecha límite.
- El gestor del proyecto enviará el entregable al experto para su revisión por pares. El experto enviará los comentarios en un plazo no mayor a 5 días.
- El responsable del entregable modificará el entregable en consecuencia y lo enviará al gestor del proyecto 2 días antes de la fecha de entrega.
- En caso de que el entregable cumpla con los objetivos requeridos, el gestor del proyecto enviará el entregable a la dirección del proyecto, y cargará simultáneamente la versión PDF en MS Teams.

Es obligatorio que todos los entregables de más de 50 páginas incluyan un resumen ejecutivo al inicio del entregable. Se debe evitar documentos que superen las 100 páginas, en cuyo caso se podrá considerar el uso de anexos.

El resumen ejecutivo es un documento breve o una sección de un documento que resume un informe o propuesta más extenso o un grupo de informes relacionados de tal manera que los lectores puedan familiarizarse rápidamente con una gran cantidad de material sin tener que leerlo todo. Por lo general, contiene una breve exposición del problema o propuesta cubierta en los documentos principales, información general, análisis conciso y conclusiones principales.

Un resumen ejecutivo se diferencia de un resumen en que un resumen suele ser más breve y normalmente pretende ser una descripción general u orientación en lugar de ser una versión condensada del documento completo. Un resumen es una breve exposición resumida leída por las partes que intentan decidir si leer o no el documento principal, mientras que "n resumen ejecutivo, a diferencia de un resumen, es un documento en miniatura que puede leerse en lugar del documento más largo.

### 3.2.2 Agenda y actas de las reuniones.

La información de las reuniones es obligatoria para garantizar que las decisiones tomadas sean conocidas y aceptadas por todas las personas que trabajan en el proyecto.

El gestor del proyecto propondrá un orden del día para cada reunión con al menos 8 días de antelación a la fecha acordada. El responsable de la reunión será responsable de levantar el acta en un plazo no mayor a 15 días. Tanto el orden del día de la reunión como el acta debe subirse al repositorio de ficheros en MS Teams en el directorio correspondiente.

### 3.2.3 Reporte de gestión del proyecto

Se va a producir dos informes de gestión del proyecto a final del primer año de ejecución y al final del proyecto. Los Informes Periódicos del Proyecto incluirán:

- Un resumen del estado del proyecto
- Una lista de los objetivos del proyecto abordados en el período del informe y los avances y logros del trabajo durante el período.

- Un informe de gestión.
- La explicación sobre el uso de recursos y estados financieros.

### 3.3 Gestión del Riesgo

Para este proyecto se ha identificado las siguientes áreas principales de posibles riesgos:

- Técnico: falta de competencia para superar dificultades inesperadas.
- Financiero: deterioro de la situación económica de un socio que le impone una paralización o una reducción inaceptable de todas sus actividades.
- Disponibilidad de recursos clave: abandono de la participación al proyecto de recursos con roles clave.

También podrían ocurrir varias combinaciones de estos tres factores negativos que podrían aumentar el impacto negativo. En la propuesta técnica se identificó un conjunto preliminar de riesgos antes del inicio del proyecto. Estos riesgos serán monitoreados continuamente, y en caso de ser necesario se activarán los planes de contingencia preestablecidos.

#### 3.3.1 Definiciones

##### **Riesgo**

El riesgo es una medida de la incapacidad de lograr los objetivos generales del proyecto dentro de las limitaciones definidas de costo, cronograma y técnicas (rendimiento y calidad) y tiene dos componentes: Probabilidad de no lograr un resultado particular y Consecuencias (impacto) de no lograr ese resultado.

El riesgo es una medida de la diferencia entre el desempeño real de un proceso y las mejores prácticas conocidas para realizar ese proceso.

El riesgo también puede ser la posibilidad de que una amenaza determinada explote las vulnerabilidades de un activo o grupo de activos para causar pérdida o daño a los activos. Normalmente se mide mediante una combinación de efecto y probabilidad de ocurrencia.

##### **Evento de riesgo**

Los eventos de riesgo son aquellos eventos, si salen mal, podrían resultar en problemas en el desarrollo de los resultados esperados, la producción y evaluación de los prototipos y la difusión de los resultados. Los eventos de riesgo deben definirse a un nivel tal que el riesgo y las causas sean comprensibles y puedan evaluarse con precisión en términos de probabilidad y consecuencias, para establecer el nivel de riesgo.

##### **Tipo de riesgo**

Un Riesgo Técnico es el riesgo asociado con la evolución de los resultados de la investigación y el desarrollo de prototipos que afecta el nivel de rendimiento necesario para cumplir con los requisitos de la propuesta técnica.

Un riesgo financiero está asociado con la capacidad del proyecto para lograr sus objetivos de costos según lo determinado en la propuesta económica. Dos áreas de riesgo que influyen en el costo son:

- Riesgo de que las estimaciones de costos y los objetivos no sean precisos y razonables y
- Riesgo de que la ejecución del proyecto no cumpla con los objetivos de costos como resultado de no mitigar los riesgos técnicos.

Los Riesgos de planificación son aquellos asociados a la adecuación del tiempo estimado y asignado para el desarrollo, producción y puesta en marcha del sistema. Dos áreas de riesgo relacionadas con el riesgo de planificación son:

- Riesgo de que las estimaciones y objetivos de la planificación no sean realistas y razonables
- Riesgo de que la ejecución del programa no alcance los objetivos de la planificación como resultado de no mitigar los riesgos técnicos.

### 3.3.2 Clasificación del riesgo

Este es el valor que se le da a un evento de riesgo con base en el análisis de la probabilidad e impacto del evento. Las calificaciones de riesgo de Bajo, Moderado o Alto se asignan según los siguientes criterios.

- Bajo riesgo: Tiene poco o ningún potencial para causar aumento en el costo, interrupción del cronograma o degradación del desempeño. Las acciones dentro del alcance del proyecto planificado y la atención normal de la gestión deberían dar como resultado el control de un riesgo aceptable.
- Riesgo moderado: Puede causar algún aumento en el costo, interrupción del cronograma o degradación del desempeño y/o la calidad. Es posible que se requieran medidas especiales y atención de gestión para controlar el riesgo aceptable.
- Alto riesgo: Es probable que cause un aumento significativo en el costo, interrupción del cronograma o degradación del rendimiento y/o la calidad. Se requerirán acciones adicionales significativas y atención de alta prioridad por parte de la gerencia para controlar el riesgo aceptable.

### 3.3.3 Organización y responsabilidades de la gestión de riesgos

El gestor del proyecto por parte de expert.ai es el administrador general de riesgos y responsable de informar a la dirección del proyecto en UPM sobre el estado de los riesgos del proyecto y realizar el seguimiento de los esfuerzos para reducir los riesgos de nivel alto a niveles aceptables.

Los participantes e involucrados en el proyecto son responsables de la evaluación de riesgos en los paquetes de trabajo en los que participen identificando y analizando los riesgos e informando de ellos al gestor del proyecto. Además, deben hacer seguimiento del riesgo y de los esfuerzos realizados para reducir el riesgo a niveles aceptables.

### 3.3.4 Proceso de gestión del riesgo

La evaluación de riesgos incluye la identificación de eventos (o procesos) de riesgo críticos, que podrían tener un impacto adverso en el proyecto, y el análisis de estos eventos para determinar la probabilidad de ocurrencia y variación del proceso y sus consecuencias.

La evaluación de riesgos es un proceso iterativo. Cada evaluación de riesgos es una combinación de los riesgos identificados y analizados en la fase anterior y la identificación y análisis de riesgos en los hitos y entregables actuales de acuerdo con la planificación del proyecto.

#### 3.3.4.1 Identificación de riesgos

La identificación de riesgos es el primer paso en el proceso de evaluación. El proceso básico implica buscar en todo el plan del proyecto para determinar aquellos eventos críticos que impedirían que se alcance los objetivos.

Todos los riesgos identificados se documentarán en la Tabla de Riesgos, con una declaración del riesgo y una descripción de las condiciones o situaciones que causan preocupación y el contexto del riesgo. El procedimiento básico para la identificación de riesgos consta de los siguientes pasos:

- Comprender los requisitos y los objetivos generales de calidad y rendimiento del proyecto. Examinar las condiciones operativas (funcionales y ambientales) bajo las cuales se deben alcanzar los objetivos con relación a la planificación del proyecto.
- Identificar los procesos y actividades que se necesitan para alcanzar los resultados.
- Evaluar cada proceso y actividad frente a áreas de riesgo.

### Indicadores de riesgo

Los siguientes indicadores son útiles para identificar riesgos:

- Recursos insuficientes o inadecuados: las personas, los fondos, el cronograma y las herramientas son inadecuados.
- Las tendencias o pronósticos negativos son motivo de preocupación y pueden requerir acciones específicas para revertirlas.
- No utilizar las mejores prácticas garantiza que el proyecto experimentará algún riesgo.
- La comunicación es un factor crítico de éxito. No proporcionar activamente la información disponible ni exigir activamente la información requerida introducirá un riesgo considerable.
- Falta de estabilidad, claridad o comprensión de los requisitos: los requisitos impulsan la investigación y el diseño de los prototipos. Los requisitos cambiantes o mal establecidos garantizan la introducción de un deterioro del rendimiento, un aumento de costos y problemas de planificación.

#### 3.3.4.2 *Análisis de riesgos*

El análisis de riesgos es una evaluación de los eventos de riesgo identificados para determinar los posibles resultados, y la probabilidad de que ocurran esos eventos y el impacto de esos eventos en los resultados del proyecto. Una vez que se ha determinado esta información, el evento de riesgo se puede calificar según los criterios del proyecto y se puede asignar una evaluación general de bajo, moderado o alto.

El procedimiento básico para el análisis de riesgo comprende los siguientes pasos:

- Recopilar todos los riesgos identificados.
- Asignación de probabilidad e impacto a cada evento de riesgo para establecer una calificación de riesgo.
- Análisis cuantitativo.

Para cada riesgo identificado durante el proceso de identificación de riesgos se realizará una evaluación con base en la probabilidad de que el riesgo ocurra y en el impacto del riesgo en el proyecto:

- **Probabilidad:** Para cada área de riesgo identificada, se debe determinar la probabilidad del riesgo. Se definen cuatro niveles de probabilidad en el proceso de evaluación de riesgos: **remoto**, **improbable**, **probable** y **altamente probable**. Si la probabilidad de que ocurra un evento es cero, no hay riesgo según nuestra definición.

- **Impacto:** Para cada área de riesgo identificada, se debe responder la siguiente pregunta: Dado que ocurre el evento, ¿cuál es la magnitud del impacto? Hay cuatro niveles de consecuencias para este proyecto, desde 1 menor impacto o no impacto a 4 alto impacto. Además, hay **cuatro áreas** que se evalúan al determinar el impacto: **desempeño técnico, planificación, costo** e impacto en **otros paquetes de trabajo**. Es necesario que se aplique al menos una de las cuatro áreas de consecuencias para que exista un riesgo; si no hay consecuencias adversas en ninguna de las áreas, no hay riesgo alguno.
  - Desempeño Técnico: Esta categoría se refiere a los objetivos técnicos.
  - Planificación: Esta categoría se refiere a los impactos en el marco temporal general del proyecto.
  - Costo: Costos directos asociados al proyecto.
  - Impacto en otros paquetes de trabajo: Tanto la consecuencia de un riesgo como las acciones de mitigación asociadas con la reducción del riesgo pueden afectar a otro paquete de trabajo. Esto puede implicar coordinación adicional o atención de gestión y, por lo tanto, puede aumentar el nivel de riesgo.

Además, para cada área se define los siguientes criterios para asignar el impacto:

- Desempeño Técnico:
  - Impacto 1: Impacto mínimo o inexistente
  - Impacto 2: Aceptable con una reducción importante del margen
  - Impacto 3: Aceptable con ningún margen restante.
  - Impacto 4: No aceptable
- Planificación:
  - Impacto 1: Impacto mínimo o inexistente
  - Impacto 2: Desviación menor en hitos clave. Incapacidad de cumplir las fechas.
  - Impacto 3: Desviación mayor en hitos clave.
  - Impacto 4: Imposibilidad de alcanzar hitos del proyecto.
- Costo:
  - Impacto 1: Incremento en el costo total menor a 5%
  - Impacto 2: Incremento en el costo total entre 5% y 7%
  - Impacto 3: Incremento en el costo total entre 7% y 10%
  - Impacto 4: Incremento en el costo total mayor al 10%
- Oros paquetes de trabajo
  - Impacto 1: Ninguno
  - Impacto 2: Moderado
  - Impacto 2: Importante
  - Impacto 4: No aceptable

La clasificación final del riesgo se calcula teniendo en cuenta la siguiente matriz donde para una determinada probabilidad y el impacto esperado en el proyecto se define la correspondiente evaluación entre riesgo bajo, medio y alto.

Probabilidad\Impacto	1	2	3	4
<b>Remoto</b>	Bajo	Bajo	Bajo	Medio
<b>Improbable</b>	Bajo	Bajo	Medio	Alto
<b>Probable</b>	Bajo	Medio	Alto	Alto
<b>Altamente Probable</b>	Medio	Alto	Alto	Alto

Las categorías de clasificación de los riesgos se describen arriba en la sección Clasificación del Riesgo.

#### 3.3.4.3 Evaluación de Riesgos

Durante el Análisis de Riesgos, es posible que los escenarios identificados de eventos de riesgo que ocurran afecten varias áreas de impacto. En este caso, está presente una combinación de impactos y posiblemente con diferentes probabilidades lo que influye en las acciones requeridas. Todas las áreas de impacto identificadas para un evento de riesgo deben registrarse y el área de impacto con mayor probabilidad que causa la máxima evaluación final debe identificarse claramente.

Una vez finalizado el análisis de riesgos, se realiza el análisis cuantitativo y se asigna una calificación a cada riesgo: bajo, medio, alto. Esto finalmente produce una visión general del estado de riesgo durante todo el transcurso del proyecto y forma parte de la tabla de riesgos.

#### 3.3.5 Plan de Contingencia

Una vez identificados y evaluados los riesgos del proyecto, se debe desarrollar el enfoque para manejar cada riesgo significativo, es decir con un resultado de evaluación alto. Existen cuatro técnicas u opciones para manejar los riesgos:

- Evitar: Aplicar de tareas para evitar el evento de riesgo.
- Controlar: vigilar las condiciones ambientales para detectar influencias en un riesgo ya evaluado.
- Transferir: Aplicar tareas para fijar un riesgo a un nivel inferior.
- Asumir: basar una decisión para manejar planes en la suposición de que ocurre el evento de riesgo.

Para todos los riesgos identificados, las diversas técnicas de manejo deben evaluarse en términos de viabilidad, efectividad esperada, implicaciones de costo y cronograma, el efecto sobre la calidad y rendimiento técnico del sistema y la técnica más adecuada seleccionada.

Los resultados de la evaluación y selección se incluirán y documentarán en la tabla de riesgos.

#### 3.3.6 Procedimiento de monitorización del riesgo

Cada responsable debe hacer seguimiento e informar la efectividad de las acciones de manejo de los riesgos asignados. Los riesgos clasificados como Altos se informarán a la dirección del

proyecto, quien los manejará y rastreará hasta que el riesgo se considere Medio o Bajo y se recomiende su "Cierre".

Los riesgos clasificados como Medio se informarán al gestor de proyecto de expert.ai, quien los rastreará hasta que el riesgo se considere Bajo y se recomiende su "Cierre". Los riesgos clasificados como Bajos se rastrean dentro del paquete de trabajo y se realiza seguimiento continuamente para garantizar que se mantengan bajos.

El proceso de gestión de riesgos es continuo. La información obtenida del proceso de seguimiento se retroalimenta para reevaluaciones y evaluaciones de las acciones de manejo para mejorar el proceso mismo.

### 3.3.7 Tabla de riesgo

La tabla de riesgos es la principal herramienta para realizar un seguimiento de los diferentes riesgos. Contiene todos los campos para evaluar, monitorear y mitigar correctamente un riesgo. La tabla de riesgos Implementa la matriz de evaluación de riesgos descrita anteriormente y un indicador de riesgo global que considera la evaluación de las cuatro áreas de consecuencias en su conjunto.

De esta manera, el miembro del proyecto que identifica un riesgo sólo tiene que indicar la probabilidad del riesgo (MP=Muy Probable=4; P=Probable=3; I=Improbable=2; R=Remoto=1) y el impacto de cada uno de ellos en las áreas de impacto (1 Mínimo, 4 Máximo). La tabla es capaz de traducir la evaluación en las tres categorías (riesgo alto, riesgo moderado, riesgo bajo) y calcular el indicador global como promedio de las diferentes áreas (0 Mínimo, 4 Máximo).

La tabla de riesgos se implementa como una hoja de cálculo que puede ser editada colaborativamente por todos los miembros del proyecto y que va a estar disponible en el repositorio de contenidos en MS Teams. Una parte de la tabla de riesgos actual se presenta en la siguiente figura:

Id.	Autor	Descripción del riesgo	Fecha	Probabilidad	Impacto por área				Evaluación	Responsable	Mitigación del riesgo
					Desempeño Técnico	Planificación	Costo	Otros paquetes de trabajo			
Riesgo 1	JMGP	Bajo rendimiento del licitador	04/12/2023	Remoto	4	3	4	4	Medio	JMGP	La experiencia del licitador, conocimientos, posición en el campo y internacional avala la trayectoria de expert.ai en temas de investigación en Sin embargo, el desempeño del equipo con respecto a los objet monitoreado y documentado en los informes de PT7 y reuniones periód; dirección de RESData. Los posibles casos de bajo rendimiento se i; tempranamente, antes de que puedan causar perturbaciones, y se <del>evocaron conclusiones para evitar repeticiones negativas en el proyecto.</del>
Riesgo 2	JMGP	Selección irrelevante de dominios y casos de uso	04/12/2023	Remoto	4	4	3	3	Medio	JMGP	La experiencia del licitador en el desarrollo de modelos híbridos que combi; grafos de conocimiento permite partir de una base sólida para la sel; dominios, parámetros y casos de uso principales, así como los asociados a ellos. Por otro lado, los requisitos mínimos están definidos a técnico de la licitación y han sido documentados en esta propuesta (Tabl; la selección de dominios relevantes contamos con la colaboración con la <del>del proyecto RESData.</del>
Riesgo 3	JMGP	Falta de detalle en los casos de uso asociados	04/12/2023	Improbable	3	3	3	3	Medio	JMGP	La metodología agil (sección 2.6.1) permite refinar la definición de los cas; en las diferentes iteraciones de desarrollo. Además, la experiencia de si; este tipo de proyectos garantiza un nivel de detalle adecuado en la di; casos de uso y requisitos.
Riesgo 4	JMGP	Baja calidad de los entregables producidos	04/12/2023	Remoto	3	4	3	3	Medio	JMGP	Todos los entregables pasarán por una revisión por pares antes de su e; dirección del proyecto.
Riesgo 5	JMGP	Pérdida de competencia críticas o personal clave en el proyecto	04/12/2023	Probable	4	4	1	2	Alto	JMGP	El grupo expert.ai reemplazará el personal que pueda dejar el proyecto co; <del>habilidad en caso de que sea necesario.</del>
Riesgo 6	JMGP	Conflictos de propiedad	04/12/2023	Remoto	3	3	1	3	Bajo	JMGP	La dirección del proyecto asegura que se asignen licencias abiertas o software, modelos y datasets, generados en el proyecto, de acuerdo <del>condiciones de la licitación.</del>
Riesgo 7	JMGP	Desacuerdo entre el licitador y la dirección del proyecto INESData en el alcance y propósito de los principales entregables e hitos	04/12/2023	Remoto	3	3	3	3	Bajo	JMGP	Esta propuesta se ha redactado teniendo en cuenta los requisitos y tareas en los pliegos técnicos de la licitación y la experiencia del licitador en investigación similares y proyectos relacionados, con el objeto de <del>desacuerdos en el alcance y propósito de los entregables. Además, el E7.1</del> producirá en M1 un plan de trabajo inicial y plan de calidad <del>revisado con la dirección del proyecto INESData.</del>
Riesgo 8	JMGP	Deterioro de la situación económica	04/12/2023	Improbable	3	2	2	2	Medio	JMGP	El grupo expert.ai movilizará recursos humanos, técnicos y financieros di; <del>antes en caso de ser necesario.</del>
Riesgo 9	JMGP	Otras iniciativas desarrollan una tecnología similar al inicio del proyecto o producen un cambio drástico en el estado de la cuestión al final	04/12/2023	Altamente probable	3	3	2	1	Alto	JMGP	Hasta el inicio del proyecto y durante su ejecución monitorearemos con; métodos novedosos, resultados de investigación y avances en el c; aseguramos de que aprovechamos las tecnologías y métodos más a; efectivos. Esta visión actualizada del estado de la cuestión se reflej; estrategia (E7.1). Los resultados relevantes que estén disponibles con lic; código abierto serán cuidadosamente analizados y alimentados en nue; enfoque para garantizar desarrollos a la vanguardia de la tecnología.

Ilustración 5. Parte de la tabla de riesgos preparada para el proyecto.

## Apéndice A: Descripción detallada del trabajo

A continuación, se presenta una descripción detallada de cada uno de los paquetes de trabajo.

PT:	1	Mes inicial y Final	M2-M3
Título PT:	Inyección de conocimiento en LLM - Estado del arte, guías y métricas		
Esfuerzo en PM: <sup>2</sup>	2		
Requisitos:	R1, R2, R3, R4		

**Objetivo:** En este paquete de trabajo se definirán las guías y principios básicos que regirán el desarrollo del proyecto desde un punto de vista técnico y que permitirán centrar el desarrollo de métodos para el enriquecimiento de LLM mediante el uso de conocimiento estructurado y razonamiento simbólico en el ámbito de INESData. Para ello, PT1 abarcará tres áreas de trabajo principales. En primer lugar, tratándose de un campo en el que las innovaciones tecnológicas se suceden a gran velocidad, se revisará el estado del arte de métodos de integración de LLM y grafos de conocimiento y uso de herramientas externas por parte de LLM con el fin de recoger una visión actualizada de los métodos y recursos de interés para el proyecto. De especial interés será la aparición de nuevos y posiblemente más eficaces LLM fundacionales que sean particularmente atractivos para fines de integración con grafos de conocimiento, bien por sus capacidades técnicas, por sus esquemas de licenciamiento u otras características. Además, catalogaremos y haremos aprovisionamiento de aquellos grafos de conocimiento, herramientas externas y APIs relevantes en los que centraremos nuestro estudio.

Tareas a realizar:

**T1.1: Análisis actualizado del estado del arte (M2-M3).** Esta tarea definirá una visión actualizada del estado de la cuestión en métodos para el enriquecimiento de LLM mediante grafos de conocimiento y herramientas externas de razonamiento simbólico sobre la que partirá el desarrollo de los métodos en los que se centrará el proyecto. La tarea T1.1 cubrirá todas las áreas técnicas del proyecto propuesto, que incluyen de manera fundamental la inyección de conocimiento procedente de grafos de conocimiento (PT2) y el uso de herramientas externas (PT3).

**T1.2: Definición de guías con los enfoques propuestos para la inyección de conocimiento (M2-M3).** La tarea T1.2 tiene como objetivo definir una hoja de ruta que guiará el desarrollo y la experimentación durante el proyecto. Para ello, esta tarea incluirá el desarrollo de un plan de experimentación y desarrollo que considerará factores como la optimización del tiempo y los recursos necesarios para el entrenamiento de los modelos y la experimentación con los métodos desarrollados en el proyecto, así como la selección de LLM fundacionales para su uso y experimentación en el proyecto, teniendo en cuenta no sólo sus características técnicas sino también sus licencias, con preferencia por LLMs de código abierto cuya licencia también permita su modificación y explotación comercial. Basándose en estos criterios, T1.2 definirá los métodos fundamentales más prometedores de entre los ofrecidos por el estado del arte, sobre los que basaremos la experimentación y el desarrollo de extensiones y nuevos métodos para el enriquecimiento de LLM mediante conocimiento

<sup>2</sup> Person month PM indica el tiempo aproximado que el personal de la empresa dedica a una tarea o proyecto. Se ha calculado un total de 37 PM en el proyecto teniendo en cuenta los costes salariales y un valor medio aproximado de 4.500€ por PM.

estructurado y herramientas externas, así como su utilización en dominios y aplicaciones concretas.

**T1.3: Diseño de una metodología de evaluación (M2-M3).** Esta tarea definirá las métricas para la evaluación de inyección de conocimiento y de medición de la adaptación del modelo de lenguaje. T1.3 revisará las métricas y marcos de evaluación (benchmarks) relevantes para los métodos de enriquecimiento de LLM que se desarrollarán durante el proyecto. La tarea T1.3 propondrá una metodología de evaluación de factualidad y sesgo en LLM que reutilizará y extenderá dichos marcos de evaluación, teniendo en cuenta las lenguas y dominios de interés para INESData. T1.3 partirá de modelos y datasets existentes para medir la factualidad en LLM, como FEVER (Thorne et al., 2018), FactCC (Kryscinski et al., 2020), QAGS (Wang et al., 2020), WICE (Kamoi et al., 2023), RARR (Gao et al., 2022) o FacTool (Chern et al., 2023). Datasets como StereoSet (Nadeem et al., 2021) y CrowS-Pairs (Nangia et al., 2020) se considerarán como punto de partida para evaluar sesgo en LLM. T1.3 también evaluará el volumen de datos de entrenamiento, la cantidad de recursos hardware y consumo energético requerido para entrenar los modelos. Para la evaluación de tareas como la extracción de grafos de conocimiento a partir de texto mediante LLM (T4.4), T1.3 se basará en datasets como LAMA (Petroni et al., 2019), TAC,<sup>3</sup> KnowledgeNet (Mesquita et al., 2019) y LM\_KBC22:<sup>4</sup>

**T1.4: Selección de grafos de conocimiento y herramientas (M2-M3).** Esta tarea hará una selección de fuentes de conocimiento estructurado, herramientas externas a LLM y APIs que se utilizarán en el enriquecimiento y ajuste de LLM. T1.4 tendrá en cuenta los criterios de selección derivados de los dominios de aplicación y casos de uso identificados en T4.1. Los grafos de conocimiento y APIs de interés para T1.4 pueden incluir grafos de conocimiento específicos de dominio, recursos como Wikidata, Wikipedia, Dbpedia, grafos de conocimiento lexico-semánticos como WordNet, de sentido común como ATOMIC o ConceptNet, servicios de razonamiento lógico, simbólico o aritmético como Wolfram Alpha, herramientas externas, como motores de búsqueda, calculadoras o calendarios y repositorios de herramientas como Toolforming Hub. Entre otras fuentes, T1.4 prestará especial atención al contenido del catálogo del espacio de datos lingüístico de INESData y otras iniciativas europeas.

Entregables:

**E1.1: Documento del estado de la cuestión, guías para la inyección de conocimiento en LLM y métricas (M3).** Documento que recoge los resultados de las tareas T1.1, T1.2 y T1.3 y T1.4. Incluye un análisis actualizado del estado del arte, las guías con los enfoques propuestos para el desarrollo de métodos de inyección de conocimiento en LLM y el diseño de métricas de evaluación. También incluye una selección de grafos de conocimiento y herramientas externas a LLM, así como sus APIs de interés para los dominios del proyecto.

PT:	2	Mes inicial y Final	M4-M13
Título PT:	Inyección de conocimiento procedente de grafos de conocimiento		
Esfuerzo en PM:	8		
Requisitos:	R5, R6, R7, R8, R9		

<sup>3</sup> <https://tac.nist.gov/2010/KBP>

<sup>4</sup> <https://github.com/lm-kbc/dataset2022>

**Objetivo:** En este paquete de trabajo se desarrollarán los métodos necesarios para la inyección de conocimiento estructurado en LLM. Para ello, PT2 se centrará en los escenarios fundamentales en los que la inyección de conocimiento puede plantearse, determinados por las distintas etapas del ciclo de vida del LLM. El primer escenario es el de la adaptación de LLM que ya fueron entrenados previamente, con el fin de dotarlos de capacidades que les permitan adaptarse a un dominio específico (T2.2), así como resolver problemas de manera eficaz (T2.3). El segundo escenario cubre la incorporación de conocimiento estructurado durante el preentrenamiento del LLM (T2.4). Estos escenarios responden a distintas necesidades en un paradigma PLN como el actual, basado en la reutilización de LLM para su adaptación a tareas y dominios específicos, así como para enriquecer el preentrenamiento del LLM, optimizando el dataset de entrenamiento. De especial utilidad para la adaptación de LLM a dominios específicos, en la tarea T2.1 PT2 desarrollará métodos que permitan determinar el alcance del conocimiento contenido en LLM, perfilándolos de manera que la información así extraída se pueda usar como guía para seleccionar grafos de conocimiento y sus subgrafos con el conocimiento necesario para compensar de manera más efectiva los posibles defectos de cobertura de dominio o inconsistencias en el LLM a enriquecer. El software y recursos resultantes asociados a los métodos para la inyección de conocimiento en LLM desarrollados en PT2 serán desplegados en la plataforma de INESData.

Tareas a realizar:

**T2.1: Desarrollo de métodos para la extracción de subgrafos de conocimiento (M4-M13).**

Esta tarea tendrá en cuenta la inyección de fragmentos de grafos de conocimiento en el LLM. El objetivo de T2.1 será el desarrollo de métodos que permitan identificar y extraer subgrafos de conocimiento, considerando el dominio y el LLM en el que ese conocimiento será inyectado, con el fin de mejorar de la forma más precisa posible la cobertura del LLM en esos dominios, así como su factualidad y ausencia de sesgos. Para ello, T2.1 desarrollará un método que permita identificar y extraer del grafo de conocimiento el conocimiento factual válido para los dominios verticales del proyecto identificados en T5.1.

**T2.2: Desarrollo de métodos para la adaptación de LLM mediante conocimiento estructurado (M4-M13).**

Basada en los resultados de T2.1, el propósito de esta tarea es el desarrollo de métodos para facilitar la adaptación de LLM pre-entrenados a un dominio específico. T2.2 propondrá un marco flexible y simple que admita la inyección de múltiples tipos de conocimiento en grandes modelos previamente entrenados, sin necesidad de reentrenar todo el LLM. Este marco se basará en mecanismos como Adapters (Houlsby et al., 2021) y propuestas de métodos de inyección de conocimiento en LLM, como K-Adapter (Wang et al., 2021) y Kformer (Yao et al., 2022). Los métodos propuestos en T2.2 adaptarán los LLM ya pre-entrenados, inyectando conocimiento procedente de uno o varios (sub)grafos de conocimiento con sus propias entidades y relaciones. De esta manera, los métodos propuestos en T2.2 no necesitarán reentrenar todo el LLM sino sólo las capas de adaptación añadidas al modelo, lo que contribuirá a su frugalidad en términos de volumen de datos e infraestructura de entrenamiento. Mediante estos métodos, T2.2 también pretende prevenir el fenómeno conocido como olvido catastrófico, en el que parte del conocimiento preexistente en el LLM se pierde durante el ajuste del modelo sobre nuevos datasets. La tarea T2.2 se centrará en los LLM y grafos de conocimiento seleccionados en el PT1.

**T2.3: Inyección de estrategias de resolución de problemas (M4-M13).**

Esta tarea llevará a cabo la inyección de conocimiento estructurado sobre métodos de resolución de problemas (PSM) con el fin de dotar al LLM de estrategias de razonamiento para la resolución efectiva de problemas complejos. T2.3 persigue entrenar al LLM a descomponer una tarea PLN en subtareas más sencillas para conseguir un determinado objetivo. Esta tarea se enfocará en PSM generales, como por ejemplo clasificación, validación, diseño, predicción o planificación, que se pueden aplicar en multitud de dominios y tareas. T2.3 obtendrá conocimiento

estructurado sobre PSM procedente de recursos generados por trabajos clásicos en áreas de inteligencia artificial relacionadas con la adquisición de conocimiento (Gómez-Pérez, 2010), McDermott, 1988; Benjamins, 1993; Fensel, 2000; Schreiber et al., 2000).

**T2.4: Desarrollo de métodos para la inyección de conocimiento durante preentrenamiento (M4-M13).** De manera complementaria a T2.2, el objetivo de esta tarea es facilitar la inyección de conocimiento factual no sólo a posteriori, sino también durante el propio preentrenamiento del LLM, mediante el desarrollo de métodos para utilizar conocimiento procedente de grafos de conocimiento general como Wikidata o Dbpedia, recursos léxico-semánticos como WordNet y grafos de conocimiento de sentido común como ATOMIC (Sap et al., 2019). T2.4 persigue así que los LLM resultantes exhiban una mejora en sus habilidades generales de razonamiento y comprensión de lenguaje, manifestada a través de su rendimiento en tareas como desambiguación con respecto al LLM base.<sup>5</sup> Al igual que T2.2, T2.4 seguirá un enfoque frugal en cuanto a volumen de datos e infraestructura de entrenamiento y experimentará con arquitecturas de LLM fundacionales de tamaño contenido, de manera que el entrenamiento de los LLM enriquecidos con grafos de conocimiento sea abordable mediante la infraestructura al alcance de los medios de la empresa licitadora y de INESData. Aun así, los resultados de T2.4 serán generalizables para LLM de mayores dimensiones. Para ello, T2.4 se basará en los modelos de lenguaje y datasets liberados por el proyecto MarIA,<sup>6</sup> promovido por el Plan Nacional de Tecnologías del Lenguaje y basados en RoBERTa y GPT-2.

Entregables:

**E2.1v1: Métodos de inyección de conocimiento en LLM (M7).** Este entregable consiste en la primera versión del software resultante de la implementación de los métodos desarrollados para la extracción de subgrafos de conocimiento (T2.1), la adaptación de LLM preexistentes (T2.2 y T2.3) y el enriquecimiento del entrenamiento de LLM mediante el uso de conocimiento estructurado (T2.4). El software vendrá acompañado de un documento donde se presentará el desarrollo de los métodos y su implementación.

**E2.1v2: Métodos de inyección de conocimiento en LLM (M13).** Este entregable consiste en la versión final del software resultante de la implementación de los métodos desarrollados en las tareas T2.1, T2.2, T2.3 y T2.4. El software vendrá acompañado de un documento donde se presentará el desarrollo de los métodos y su implementación, centrado en los cambios sobre la primera versión.

PT:	3	Mes inicial y Final	M4-M13
Título PT:	Uso de herramientas externas		
Esfuerzo en PM:	8		
Requisitos:	R9, R10, R11		

**Objetivo:** En este paquete de trabajo se desarrollarán los métodos necesarios para entrenar LLM en el uso de herramientas externas, mediante la invocación a sus APIs, como motores

<sup>5</sup> La descripción de estas capacidades también formará parte de la metodología y marco de evaluación definidos en T1.3.

<sup>6</sup> <https://huggingface.co/PlanTL-GOB-ES>

de búsqueda, calculadoras o calendarios. De esta manera, PT3 busca atajar la incidencia de alucinaciones o sesgos que podrían ser evitados mediante el acceso a este tipo de servicios. Este paquete de trabajo persigue salvar las limitaciones de enfoques anteriores, que necesitaban grandes cantidades de anotaciones o limitaban el uso de las herramientas a unas pocas tareas PLN, lo que dificulta una adopción más generalizada del uso de herramientas por parte de LLM. Para ello PT3 se basará en el concepto de Toolformer (Schick et al., 2023), que utiliza un enfoque auto supervisado y por tanto no necesita datos anotados para su entrenamiento, sin pérdida de generalidad y decidiendo por sí mismo cuándo y cómo usar qué herramientas, independientemente de la tarea PLN a resolver. Para experimentación y desarrollo de métodos que permitan el uso de herramientas externas por parte de LLM, PT3 se centrará en los LLM, herramientas y APIs seleccionados en PT1. El software y recursos resultantes asociados a los métodos desarrollados en PT3 serán desplegados en la plataforma de INESData.

Tareas a realizar:

**T3.1: Desarrollo de métodos para habilitar el uso de herramientas externas (M4-M13).**

Esta tarea desarrollará métodos que permitan el acceso por parte de LLM a servicios externos de razonamiento lógico, simbólico o aritmético, como Wolfram Alpha, herramientas externas, como motores de búsqueda, calculadoras o calendarios, y otras herramientas contenidas en repositorios como Toolforming Hub. T3.1 partirá del concepto de Toolformer para dotar a los LLM de la habilidad de determinar cuándo necesitan recuperar un hecho factual y decidir en tiempo de inferencia si es más conveniente recuperarlo de su memoria interna, con el riesgo de hacer predicciones alucinatorias, o acceder a un recurso externo. Al inicio, T3.1 se centrará en explorar las limitaciones del método de Toolformer, de reciente aparición, para a continuación extenderlo resolviendo las limitaciones que afecten especialmente a los dominios de interés definidos en T4.1. T3.1 aumentará los LLM seleccionados para tal fin en T1.2 con las herramientas seleccionadas en T1.4.

**T3.2: Desarrollo de métodos de recuperación de información factual (M4-M13).** Esta tarea desarrollará métodos para la recuperación de información factual mediante el acceso a servicios de recuperación de información y bases de conocimiento. T3.2 tendrá en cuenta tanto información factual en forma de texto no estructurado, recuperado de colecciones de documentos locales o de la Web, así como conocimiento factual estructurado procedente de grafos de conocimiento. T3.2 desarrollará la noción de Toolformer en el ámbito de la recuperación de información factual con el fin de habilitar al LLM para invocar herramientas externas de consulta, búsqueda y recuperación de información. Dada una petición al LLM, estos métodos permitirán recuperar información relacionada e incluir esa información en la petición, enriqueciéndola. La consulta así lanzada por el modelo para recuperar información factual externa reflejará el conocimiento contenido en las representaciones internas del LLM, incluido el conocimiento adicional procedente de grafos de conocimiento que pueda haber sido incorporado en el LLM mediante los métodos desarrollados en PT2.

**T3.3: Desarrollo de métodos basados en propiedades no funcionales (M4-M13).** Entre las principales limitaciones del método Toolformer documentadas hasta ahora, se encuentra el no considerar las propiedades no funcionales de los servicios para cuyo uso se entrena al LLM. Esto hace que el LLM no tenga en cuenta cuestiones como el coste de acceder a un determinado servicio o el tiempo de computación necesario durante la llamada a su API. La tarea T3.3 desarrollará métodos para resolver esta limitación. Por otro lado, otra limitación de Toolformer es que entrena al LLM en el uso de determinadas herramientas, pero no tiene en cuenta qué ocurre si por ejemplo esa herramienta concreta no está disponible en tiempo de inferencia. T3.3 también explorará métodos que permitan al LLM identificar herramientas basadas en una colección de requisitos funcionales y no funcionales relacionados con la operación a realizar, de forma que pueda recurrir a servicios alternativos cuando sea necesario.

Entregables:

**E3.1v1: Métodos de uso de herramientas externas por LLM (M7).** Este entregable consiste en la primera versión del software resultante de la implementación de los métodos desarrollados para el uso de herramientas externas por parte de LLM (T3.1 y T3.2) y su extensión de forma que también tengan en cuenta las propiedades no funcionales de las herramientas (T3.3). El software vendrá acompañado de un documento donde se presentará el desarrollo de los métodos y su implementación.

**E3.1v2: Métodos de uso de herramientas externas por LLM (M13).** Este entregable consiste en la versión final del software resultante de la implementación de los métodos desarrollados en las tareas T3.1, T3.2 y T3.3. El software vendrá acompañado de un documento donde se presentará el desarrollo de los métodos y su implementación, centrado en los cambios sobre la primera versión.

PT:	4	Mes inicial y Final	M1-M13
Título PT:	Aplicación en dominios verticales		
Esfuerzo en PM:	9		
Requisitos:	R12, R13, R14, R15, R18, R19		

**Objetivo:** En este paquete de trabajo, los resultados de PT2 y PT3 se aplicarán en los LLM seleccionados para casos de uso en dominios verticales, dotándolos de representaciones internas factuales, fiables y libres de sesgo, así como de la capacidad de acceder a fuentes de conocimiento factual y herramientas de razonamiento simbólico externas relevantes para esos dominios cuando sea necesario. Los LLM resultantes se ofrecerán a través de la infraestructura cloud de INESData. PT4 aportará los recursos necesarios para permitir la reproducibilidad de los experimentos llevados a cabo, tales como la metodología seguida para la extracción del conocimiento, el conocimiento en sí, la arquitectura desarrollada, el LLM usado y la descripción de la infraestructura utilizada para su entrenamiento. PT4 seleccionará una serie de dominios de interés y casos de uso (tarea T4.1) sobre los que se validará la eficacia de los métodos desarrollados en el proyecto, teniendo en cuenta su interés y posibilidad de impacto tanto general como en el marco de INESData. PT4 adaptará una selección de los LLM propuestos en T1.2 a dichos dominios verticales y casos de uso (T4.3) aplicando los métodos desarrollados en PT2 y PT3 y evaluará los resultados obtenidos mediante la metodología de evaluación definida en PT1. Finalmente, PT4 definirá una serie de recomendaciones para optimizar las peticiones y mejorar los resultados de inferencia de los LLM en los dominios seleccionados (T4.2).

Tareas a realizar:

**T4.1: Definición y análisis de dominios de aplicación y casos de uso (M1).** Esta tarea identificará los dominios verticales y casos de uso teniendo en cuenta su interés e impacto potencial tanto general como particular para INESData. Se priorizarán dominios y casos de uso con recursos (corpora, grafos de conocimiento, APIs) existentes, con el fin de facilitar y acelerar la experimentación en PT2, PT3 y PT4. Los dominios y casos de uso así definidos guiarán el trabajo a realizar en las tareas de PT1 relacionadas con la selección y priorización de LLMs para experimentación durante el proyecto, la selección de recursos como grafos de conocimiento y APIs de servicios externos y el diseño de métricas de evaluación.

**T4.2: Diseño de secuencias de pensamiento orientadas a resolución de problemas (M4-M7).** Esta tarea se enfoca en guiar la generación de respuestas como resultado de una petición al LLM, optimizando la relevancia y coherencia de tales respuestas en el marco del dominio de aplicación. Para ello, T4.2 extenderá el estado de la cuestión en secuencias de pensamiento (Chain of Thought, en inglés) propuesto originalmente por Wei et al. (2022), donde se probó que al mostrarle al LLM ejemplos en los que se explica el proceso de razonamiento a seguir, los resultados del LLM mejoran y aumentan su factualidad. T4.2 se centrará en esta área y estudiará la formulación de secuencias de pensamiento basadas en la noción de métodos de resolución de problemas (problem-solving methods o PSM, en inglés) (McDermott, 1988). Estas secuencias de pensamiento orientadas a la resolución de problemas darán instrucciones al LLM con especial atención a cómo descomponer tareas complejas en pasos más sencillos en función del objetivo a conseguir, la tarea que se pretenda resolver y las subtareas que esa tarea conlleve.

**T4.3: Adaptación a dominio y despliegue de LLM (M8-M13).** Esta tarea llevará a cabo la adaptación de LLM y su adaptación a dominio, aplicando los métodos para el enriquecimiento de LLM mediante conocimiento estructurado y acceso a herramientas externas desarrollados en PT2 y PT3. Para ello, T4.3 seleccionará, de entre los propuestos en T1.2, los LLM más apropiados para las distintas combinaciones de dominios y casos de uso, así como los grafos de conocimiento y herramientas externas propuestos en T1.4. Para formulación de peticiones al LLM durante inferencia y evaluación, T4.3 seguirá las guías propuestas en T4.2 y las métricas y marco de evaluación definidos en T1.3, respectivamente. T4.3 acabará con el despliegue de los LLM adaptados en la plataforma de INESData.

**T4.4: Extracción de grafos de conocimiento a partir de texto usando LLM (M8-M13).** Esta tarea se centra en la aplicación de LLM enriquecidos y adaptados a dominio para la extracción de grafos de conocimiento procedentes de corpora documentales. Para ello, T4.4 desarrollará varios métodos que utilizarán tanto LLM genéricos como adaptados a dominio para extraer entidades y relaciones con las que crear un grafo de conocimiento a partir de un texto o corpus dado. Los métodos desarrollados en T4.4 se aplicarán en el marco de los dominios seleccionados por el proyecto y los grafos de conocimiento generados serán publicados en el espacio de datos lingüístico en INESData. T4.4 estará informada por los casos de uso y dominios analizados en T4.1, así como por los corpora documentales asociados a ellos.

Entregables:

**E4.1: Análisis y definición de dominios de aplicación y casos de uso (M1).** Este entregable resultante de la tarea T4.1 analiza y define los posibles dominios de aplicación y casos de uso a ejecutar en el proyecto, teniendo en cuenta sus requisitos, disponibilidad de recursos y posible impacto.

**E4.2: LLM adaptados a dominio (M13).** Este entregable resulta de las tareas T4.2 y T4.3. Contiene los modelos adaptados y su aplicación a dominio, documentación adicional e información sobre el despliegue de los modelos adaptados en INESData.

**E4.3: Métodos de extracción de grafos de conocimiento a partir de texto mediante LLM (M13).** Este entregable consiste en los métodos desarrollados en la tarea T4.4, así como los grafos de conocimiento extraídos (y desplegados en INESData) y documentación adicional.

PT:	5	Mes inicial y Final	M4-M13
Título PT:	Demostrador e implementación del marco de evaluación		
Esfuerzo en PM:	6		

Requisitos:	R4, R12, R16, R17
-------------	-------------------

**Objetivo:** Este paquete de trabajo cubrirá dos ámbitos fundamentales relacionados entre sí. Por un lado, PT5 desarrollará el marco de evaluación definido en la tarea T1.3 de PT1, generando y adaptando los recursos, software y datasets necesarios para la evaluación de los métodos desarrollados en el proyecto, con especial interés por su impacto en los dominios y casos de uso identificados en la tarea T4.1 de PT4. Por otro lado, PT5 diseñará, desarrollará y desplegará un demostrador en la plataforma de INESData que ilustre el impacto de los métodos para enriquecimiento de LLM mediante conocimiento desarrollados en el proyecto.

Tareas a realizar:

**T5.1: Desarrollo del marco de evaluación de factualidad y sesgo (M4-M7).** Esta tarea llevará a cabo los desarrollos necesarios para implementar las métricas y marco de evaluación definido en la tarea T1.3. Por un lado, T5.1 reutilizará los datasets de evaluación identificados en T1.3 y los adaptará a las necesidades del proyecto y de INESData, incluidos requisitos relacionados con la lengua y los dominios y casos de uso de interés. Por otro lado, T5.1 desarrollará el software necesario para automatizar la evaluación de LLM sobre esos datasets facilitando la carga automática de datos y modelos en el marco de evaluación y su ejecución. En este sentido, T5.1 prestará especial atención a la comparación entre los modelos base y los resultantes de su adaptación mediante la aplicación de los métodos desarrollados en PT2 y PT3 llevada a cabo en PT4. T5.1 aportará, mediante su despliegue en el espacio de datos lingüístico de INESData, los recursos generados como resultado de la ejecución de la tarea.

**T5.2: Diseño, desarrollo y despliegue del demostrador (M8-M19).** En esta tarea se desarrollará un demostrador que permita visualizar los resultados de la experimentación con los métodos de inyección de conocimiento en LLM desarrollados en el proyecto, tanto de manera general como particularmente en el ámbito de su aplicación a los casos de uso y dominios verticales llevada a cabo en PT4. T5.2 aportará los marcos de evaluación y resultados alcanzados por estos modelos frente a otros. El demostrador permitirá comparar los resultados de cada LLM sobre el marco de evaluación con los resultados de otros modelos, estableciendo así una tabla de clasificación (del inglés, leaderboard). Finalmente, la tarea T5.2 adoptará un enfoque dinámico y reproducible basado en Jupyter notebooks y explorará el uso de extensiones como Jupyter AI<sup>7</sup>, que facilita la integración con LLM.

Entregables:

**E5.1: Marco de evaluación y recursos asociados (M7).** Este entregable consiste en los datasets y el software perteneciente al marco de evaluación y documentación descriptiva de dichos datasets resultantes de la tarea T5.1.

**E5.2: Diseño del demostrador (M10).** Este documento de diseño recoge el análisis de requisitos y casos de uso del demostrador llevado a cabo en la tarea T5.2.

**E5.3: Demostrador y resultados del marco de evaluación (M19).** Este entregable consiste en el software del demostrador desarrollado y desplegado en la tarea T5.2, así como documentación asociada.

---

<sup>7</sup> <https://github.com/jupyterlab/jupyter-ai>

PT:	6	Mes inicial y Final	M2-M19
Título PT:	Diseminación y Entrenamiento		
Esfuerzo en PM:	2		
Requisitos:	-		

**Objetivo:**

Este paquete de trabajo agrupa las actividades de diseminación del proyecto, colaboración con otras iniciativas, y entrenamiento de usuarios de los métodos de inyección de conocimiento estructurado en LLM y su utilización. En cuanto a diseminación general se creará una web del proyecto y para diseminación científica se plantea presentar artículos en conferencias y workshops en áreas de IA, PLN y KGs. Además, se organizarán sesiones de entrenamiento con la finalidad de facilitar a la comunidad de investigadores y desarrolladores de LLM la adopción de los métodos desarrollados durante el proyecto. Finalmente se establecerán lazos con el resto de las iniciativas INESData y otras iniciativas europeas, en particular relacionadas con la creación de espacios de datos lingüísticos en el contexto europeo.

Tareas a realizar:

**T6.1 Web del proyecto, enlazada con la Plataforma INESData (M2-M19).** Se desarrollará una web informativa del proyecto donde se presente la motivación del proyecto y los métodos desarrollados en él, los objetivos y principales requisitos, y su relación con INESData. Además, la web actuará como puerta de entrada al catálogo del espacio de datos lingüísticos ya que incluirá un enlace a la interfaz web del catálogo a través del que se podrá acceder a los recursos, software y LLM enriquecidos generados durante el proyecto, así como el demostrador desarrollado en este proyecto.

**T6.2 Publicaciones científicas (M8-M19).** De acuerdo con la disponibilidad de conferencias o revistas que acepten artículos relacionados con la inyección y uso de conocimiento estructurado en LLM se publicará al menos un artículo que presente los resultados del proyecto. Conferencias PLN como SEPLN, a nivel nacional, o EMNLP, ACL, NAACL o COLING son especialmente atractivas para este trabajo. También conferencias en el ámbito de los grafos de conocimiento, como ESWC e ISWC, y en la intersección con PLN, como AKBC.

**T6.3 Organización de eventos (*hackathons*, tutoriales, seminarios) (M11-M19).** Junto a la dirección del proyecto INESData se planificarán actividades de entrenamiento de investigadores y desarrolladores de la comunidad PLN y KG interesados en los métodos desarrollados durante el proyecto. Eventos apropiados para diseminar y entrenar usuarios incluyen los *hackathons* donde se propone un caso de uso y se ofrece la documentación, datasets, LLM, grafos de conocimiento y APIs para desarrollar la solución al caso de uso. Otros formatos más tradicionales son los tutoriales donde se instruye a los asistentes en el uso de las tecnologías presentadas o los seminarios donde se presentan las funcionalidades de la implementación software de los métodos acompañada de demostraciones.

**T6.4 Colaboración con otras iniciativas europeas y en el marco de INESData (M11-M19).** Se establecerán lazos con el resto de las iniciativas INESData y otras iniciativas europeas, en particular relacionadas con la creación de espacios de datos lingüísticos en el contexto

Europeo, con el objetivo de darle mayor difusión a los resultados del proyecto y aumentar la posible base de usuarios.

Entregables:

**E6.1 Web del proyecto (M3).** Entregable de software que consiste en la web del proyecto desarrollada en la tarea T6.1. La web se actualizará con los resultados generados en cada hito (H2, H3, H4, and H5).

**E6.2 Informe de publicaciones, actividades de diseminación y colaboración (M19).** Entregable que documenta las actividades de diseminación de los resultados del proyecto y de entrenamiento de la comunidad de desarrolladores e investigadores en PLN y KG interesados en el enriquecimiento de LLM con conocimiento estructurado. Este entregable también cubre las actividades de colaboración desarrolladas en el marco del proyecto.

PT:	7	Mes inicial y Final	M1-M19
Título PT:	Gestión de proyecto		
Esfuerzo en PM:	2		
Requisitos:	R20, R21, R22		

Objetivo:

Este paquete de trabajo tiene por objetivo llevar a cabo la gestión del proyecto garantizando las entregas a tiempo según el cronograma de trabajo y el presupuesto, y asignando los recursos necesarios para el correcto desempeño de las tareas. Además, se encarga de asegurar que la propiedad intelectual de los resultados del proyecto este acorde a los requisitos definidos en los pliegos técnicos de la licitación, y de la producción de reportes periódicos de la gestión del proyecto.

Tareas a realizar:

**T7.1 Progreso del proyecto, gestión del riesgo y aseguramiento de la calidad (M1-M19).**

Esta tarea se encarga de la dirección técnica general y seguimiento de la ejecución del proyecto. Las actividades incluidas en esta tarea comprenden un seguimiento del progreso del proyecto, proporcionar una evaluación continua de riesgos y en caso de problemas, iniciará las acciones correctivas requeridas, y además coordinar las tareas de aseguramiento de la calidad.

El plan de aseguramiento de la calidad describirá los roles de los diferentes actores en la gestión del proyecto, calendarios de reuniones, y dará pautas para realizar las acciones de gestión del proyecto incluyendo instrucciones y plantillas para informes técnicos y administrativos, y plantillas y convenciones de nomenclatura/numeración para archivos y documentos técnicos y administrativos. Como parte del plan de aseguramiento de la calidad, el proyecto aplicará un procedimiento de revisión interna, basado en la revisión por pares, para garantizar la calidad de sus resultados.

En esta tarea se garantiza que la propiedad intelectual de los resultados del proyecto siga las directrices definidas por la coordinación del proyecto INESData. Para ello, se identificará claramente la propiedad intelectual del software preexistente del licitador y los nuevos desarrollos realizados durante el proyecto.

**T7.2 Gestión administrativa y financiera (M1-M19).** Esta tarea cubre el establecimiento y mantenimiento de registros financieros, y la planificación y seguimiento de gastos. Además, se incluye la preparación de reportes administrativos y financieros de acuerdo con las necesidades de gestión del proyecto INESData.

Entregables:

**E7.1 Plan de trabajo inicial y plan de calidad (M1).** Este entregable incluye la versión definitiva del plan de trabajo elaborado a partir de la planificación descrita en esta propuesta. Además, incluye el plan de calidad que define los procedimientos para asegurar la calidad de los entregables, los procedimientos de comunicación y las plantillas para presentar los reportes.

**E7.2v1 Informe periódico de la gestión del proyecto (M12).** Informe de la gestión del proyecto para ser entregado al final del primer año de ejecución. En este reporte se incluye los objetivos del proyecto alcanzados en el primer año, el avance del proyecto y los principales logros, y un reporte del uso de los recursos y estados financieros.

**E7.2v2 Informe periódico de la gestión del proyecto (M19).** Informe final de la gestión del proyecto.