

EJIE-2026-099

Servicio de Centro de Excelencia IA para EJIE

Pliego de condiciones técnicas

Julio de 2026

Este documento es propiedad de Eusko Jaurlaritzaren Informatika Elkarte – Sociedad Informática del Gobierno Vasco, S.A. (EJIE). Este documento no puede ser reproducido, en su totalidad o parcialmente, ni mostrado a otros, ni utilizado para otros propósitos que los que han originado su entrega, sin el previo permiso escrito de EJIE. En el caso de ser entregado en virtud de un contrato, su utilización estará limitada a lo expresamente autorizado en dicho contrato. EJIE no podrá ser considerada responsable de eventuales errores u omisiones en la edición del documento.

Versión	Fecha	Resumen de cambios	Elaborado por:	Aprobado por:
0.1	20/02/2026	Primera versión	Javier Cela	
1.0	13/07/2026	Versión Definitiva	Javier Cela Aitor Cabero	Iratxe Montoya

Índice

1	Introducción	5
2	Antecedentes	6
3	Alcance	8
3.1	Alcance funcional	8
3.2	Alcance tecnológico	10
4	Objeto del Servicio y Líneas de Trabajo	11
4.1	Línea 1: Oficina Técnica del CoE-IA y Modelo Operativo	11
4.2	Línea 2: Gestión y Cualificación de Demanda de IA	11
4.3	Línea 3: Ingeniería y evolución de la Plataforma de IA	12
4.4	Línea 4: Arquitectura, patrones y diseño de soluciones de IA	18
4.5	Línea 5: Ingeniería IA, MLOps, LLMOps y operación	19
4.6	Línea 6: Seguridad, privacidad, cumplimiento e IA responsable	20
4.7	Línea 7: Acompañamiento a casos de uso, adopción y formación	20
4.8	Línea 8: Desarrollo e Implementación de casos de uso extremo a extremo	21
4.9	Línea 9: Soporte especializado en producción y operación de la plataforma IA	22
4.10	Línea 10: FinOps, valor y mejora continua	23
5	Modelo Operativo de IA que se pretende alcanzar	25
5.1	Principios del modelo operativo objetivo	25
5.2	Gestión de Excepciones	26
6	Composición del Equipo de Trabajo	27
6.1	Equipo propuesto	27
6.2	Roles Clave y Responsabilidades Detalladas	29
6.2.1	Responsable del Servicio	29
6.2.2	Product Owner Plataforma de IA	30
6.2.3	Arquitecto de IA	30
6.2.4	Consultor Infraestructura IA/Cloud	30
6.2.5	Ingeniero de Plataforma IA	30
6.2.6	Ingeniero LLMOps / MLOps	31
6.2.7	Ingeniero de Implantación de IA	31
6.2.8	Desarrollador de Soluciones IA	31
6.2.9	Ingeniero de Datos	32
6.2.10	Consultor Gobernanza IA, Riesgo y Cumplimiento	33
6.2.11	Ingeniero de Seguridad IA	33
6.2.12	Consultor de Observabilidad IA	33
6.2.13	Ingeniero de Soporte y Operación de la Plataforma IA	33
6.2.14	Consultor de FinOps IA	34

6.2.15	Consultor de Negocio IA	34
6.2.16	Consultor de Adopción y Formación IA	35
6.3	Flexibilidad y Perfiles Complementarios	35
7	Metodología de Trabajo y Entorno Tecnológico	36
7.1	Principios Metodológicos	36
7.2	Ciclo de Trabajo para evolución y ciclo de vida de la plataforma	36
7.3	Ciclo de Trabajo para casos de uso	36
7.4	Entorno Tecnológico	37
7.5	Indicadores Mínimos del Servicio	38
7.6	Reuniones y Mecanismos de Seguimiento	39

1 Introducción

EJIE, Eusko Jaurlaritzaren Informatika Elkartea – Sociedad Informática del Gobierno Vasco, es la Empresa pública de servicios de las tecnologías de la información y las comunicaciones (TIC), cuya razón de existir es contribuir a la consecución de un Sector Público Vasco, moderno y eficiente, en el Marco Legal establecido por el Gobierno, con la seguridad y calidad necesarias y con el debido respeto al medio ambiente.

EJIE tiene como meta final la consecución de la satisfacción de sus clientes, siendo el instrumento común de prestación de servicios TIC en el Sector Público Vasco, y comprometiéndose en:

Construir y mantener con eficiencia y calidad la infraestructura de los Sistemas de Información, posibilitando su continuidad y seguridad.

Garantizar la interoperabilidad entre las distintas administraciones.

Servir de apoyo a las necesidades de planificación y realización de la función informática de los Departamentos y Organismos Autónomos del Gobierno, asegurando la cobertura de sus demandas con el compromiso y profesionalidad adecuados a las relaciones contractuales que se establezcan.

Por tanto, EJIE debe ser, un instrumento común de referencia para la prestación de servicios TIC en el Sector Público Vasco:

- Aportando valor añadido.
- Proporcionando soluciones competitivas.
- Transmitiendo confianza a sus clientes.
- Contando con personas cualificadas y comprometidas.

Se puede obtener información más detallada y extensa en nuestra dirección de Internet <http://www.ejie.eus>

Con fecha 16 de agosto de 2016 se firmó el Acuerdo entre el Ente Público de Derecho Privado Osakidetza y Eusko Jaurlaritzaren Informatika Elkartea—Sociedad Informática del Gobierno Vasco, S.A. (EJIE) para la integración del servicio de infraestructuras y operaciones (I.O.) de OSAKIDETZA con el análogo servicio (I.O.) en EJIE, en el marco del acuerdo adoptado por el Consejo de Gobierno sobre el proceso de convergencia en materia de tecnología de la información y comunicación – TIC/ICPS el 27 de julio de 2015 y del Acuerdo de 21 de junio de 2016 relativo a dicho proceso de convergencia por el que se aprueba el documento ejecutivo “plan general de actuación en materia de convergencia TIC/ICPS”.

2 Antecedentes

En los últimos años, la Inteligencia Artificial se ha consolidado como una capacidad transversal con potencial para mejorar la productividad, automatizar procesos, facilitar el acceso al conocimiento, optimizar la prestación de servicios públicos y habilitar nuevos modelos de asistencia, análisis y toma de decisiones. No obstante, su adopción requiere un modelo operativo robusto que combine gobierno, seguridad, arquitectura, plataforma tecnológica, datos, operación, adopción y medición de valor.

EJIE se acoge a la estrategia de IA establecida por la DED, que desarrolla un modelo de gobierno y gestión de casos de uso al que el servicio objeto de este contrato deberá adherirse. Esta estrategia en general define y establece un modelo integral de gobernanza para soluciones de IA, especialmente IA Generativa, con principios, controles y mecanismos técnicos y organizativos que garanticen un uso seguro, controlado, transparente y conforme al marco normativo vigente, incluyendo RGPD y AI Act.

Asimismo, existe una primera versión básica de una plataforma corporativa de IA, que deberá ser asumida, mantenida, evolucionada y ampliada en el marco del presente contrato. El servicio a contratar deberá reforzar dicha plataforma como el mecanismo principal para escalar la IA de forma segura, reutilizable y gobernada, evitando la proliferación de soluciones aisladas, arquitecturas específicas por proyecto, herramientas no aprobadas o iniciativas de *shadow IT*.

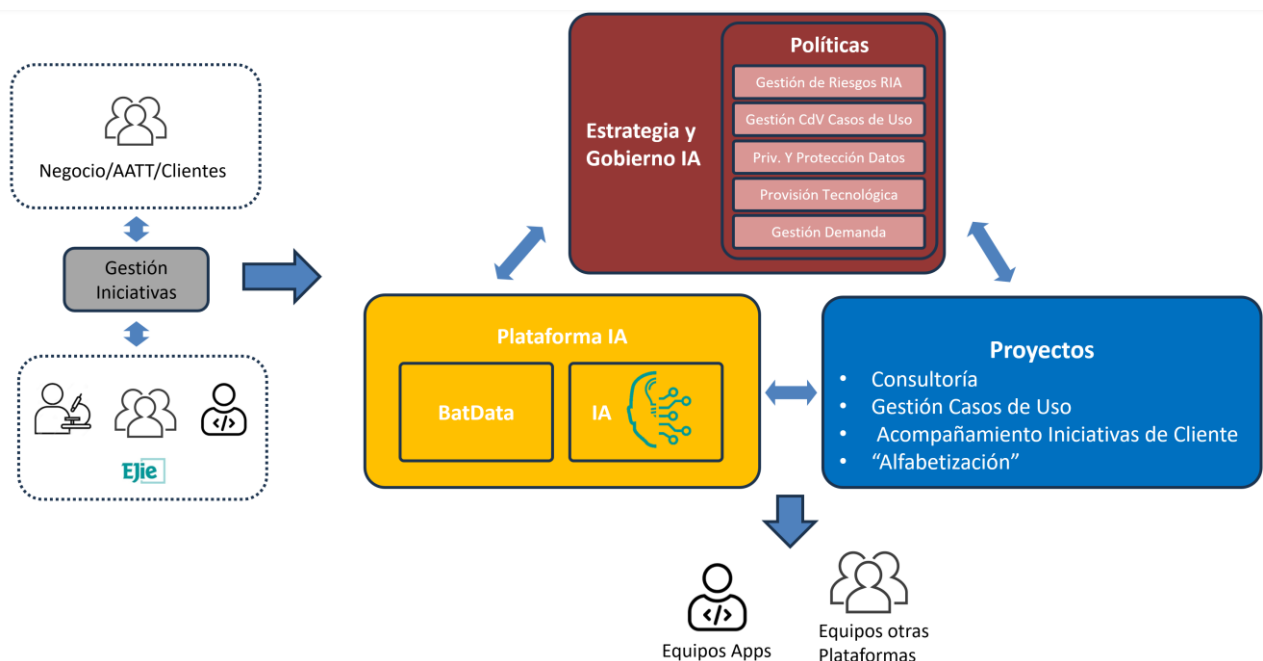


Ilustración 1 - Modelo Funcionamiento Actual

En el contexto de Batera la infraestructura de IA se concibe como un conjunto de capacidades de plataforma escalables y reutilizables, y no únicamente como acceso a GPUs o capacidad de cómputo. EJIE pretende ser un habilitador y acelerador para que sus clientes o él mismo puedan implantar con éxito casos de uso en todo su ciclo de vida, participando desde la cualificación inicial de la demanda hasta la puesta en producción y mantenimiento posteriores.

Por ello, el contrato deberá contribuir a consolidar un Centro de Excelencia de IA y una plataforma común de IA que permitan:

- Adherirse al modelo de gobierno de IA existente.
- Gestionar de forma estructurada la demanda de casos de uso.
- Evolucionar la plataforma de IA ya existente.
- Definir servicios reutilizables y patrones estándar.
- Facilitar el paso de pilotos a producción.
- Incorporar controles de seguridad, privacidad, trazabilidad, observabilidad y coste.
- Acompañar a los equipos promotores de casos de uso.
- Medir el valor generado y la adopción efectiva.
- Asegurar un modelo de operación sostenible en un entorno híbrido, cloud (principalmente sobre AWS) y coexistencia con infraestructura on-premises.

En paralelo a esto y también por la apuesta estratégica de la DED, EJIE y sus clientes han ido consumiendo gradualmente cada vez más suscripciones de cloud público. Incluso se prevé que ese uso aumente ya que por la evolución del mercado es progresivamente más habitual que las soluciones que se buscan para dar respuesta a las necesidades de los clientes tengan su mejor encaje posible en soluciones de este tipo (siendo además un albergue especialmente adecuado para implementaciones de IA). Todo esto motivó que en 2023 se licitase el contrato EJIE-082-2023 "Suministro de suscripción multicloud a cloud hiperescalares y solución de analítica de datos sobre cloud", así como otro que sucedió a éste y actualmente en vigor EJIE-2026-014 "Compra conjunta de suministro de suscripciones a cloud hiperescalares y servicios asociados". En la actualidad en EJIE, de los hiperescalares en el mercado, AWS es el principal ya que se realiza sobre él aproximadamente 3/4 partes del consumo anual.

3 Alcance

El alcance del contrato comprende la prestación de un servicio especializado para la evolución, operación, soporte, gobierno técnico y acompañamiento del Centro de Excelencia de Inteligencia Artificial y de la plataforma corporativa de IA de EJIE.

El contrato no se limitará a la ejecución de proyectos puntuales de IA, sino que deberá proporcionar capacidades transversales y reutilizables que permitan a EJIE escalar la adopción de IA de forma ordenada, segura y alineada con su estrategia y modelo de gobierno.

3.1 Alcance funcional

El servicio incluirá, al menos, las siguientes capacidades:

1. Apoyo al Centro de Excelencia de IA

- Soporte a la consolidación del modelo operativo del CoE-IA
- Coordinación con los órganos de gobierno de IA existentes.
- Apoyo a la gestión de demanda, priorización y seguimiento de casos de uso.
- Definición de roles, responsabilidades, procesos y mecanismos de coordinación.
- Elaboración de guías, procedimientos, plantillas y documentación de referencia.

2. Evolución y mantenimiento de la plataforma corporativa de IA

- Asunción de la versión inicial existente de la plataforma.
- Evolución funcional y técnica de la plataforma.
- Mantenimiento correctivo, adaptativo, preventivo y evolutivo.
- Definición de catálogo de servicios aprobados de IA.
- Definición de patrones y *golden-paths*.
- Integración de controles de seguridad, identidad, datos, logging, observabilidad, costes y operación.

3. Gobierno técnico y arquitectura de IA

- Definición y actualización de arquitecturas de referencia.
- Definición de patrones reutilizables para RAG, inferencia, agentes, orquestación, evaluación y observabilidad.
- Evaluación de tecnologías, modelos, componentes y alternativas de despliegue.
- Alineamiento con otras arquitecturas corporativas, seguridad, datos, cloud y operación.
- Alineamiento con otros equipos de EJIE como BatTech (Arquitectura de Aplicaciones e Ingeniería Software), CoE de Cloud, BatSDLC (CI/CD), Gestión de Iniciativas, etc.

- Alineamiento con BatData (Plataforma de Analítica de EJIE) y otras plataformas de EJIE y su ingeniería, así como otros servicios de catálogo de Servicios Convergentes de Batera.

4. Gestión de casos de uso de IA

- Cualificación técnica temprana de casos de uso.
- Clasificación de la demanda según patrones estándar.
- Identificación de reutilización de capacidades existentes.
- Acompañamiento a equipos promotores de aplicaciones/casos de uso.
- Apoyo al diseño, prototipado, industrialización y paso a producción.
- Seguimiento de resultados, riesgos y valor generado.

5. Operación, observabilidad y LLMOps/MLOps

- Definición y automatización de procesos de despliegue.
- Gestión específica de versiones de modelos, prompts, embeddings, pipelines y configuraciones.
- Monitorización de rendimiento, calidad, coste, trazabilidad, disponibilidad y errores.
- Gestión de incidencias, problemas, cambios y mejoras.
- Integración con procesos corporativos de operación y gestión del servicio.

6. Seguridad, privacidad, cumplimiento e IA responsable

- Incorporación de controles de seguridad y privacidad desde el diseño.
- Evaluación de riesgos específicos de IA.
- Soporte a cumplimiento normativo, incluyendo RGPD, AI Act, ENS y otra normativa que pueda ser de aplicación.
- Definición de criterios de uso aceptable, supervisión humana, trazabilidad y responsabilidad.
- Coordinación con seguridad corporativa, seguridad, gobierno del dato y responsables de servicio.

7. Adopción, formación y comunidad

- Apoyo a la adopción de capacidades de IA.
- Definición de guías y buenas prácticas.
- Formación y acompañamiento a perfiles técnicos, funcionales y de gobierno.
- Dinamización de “comunidad de práctica”.
- Soporte a usuarios clave (“*AI Champions*”) o referentes distribuidos.

8. Medición de valor, FinOps y mejora continua

- Definición de indicadores de valor y eficiencia.
- Medición de costes de plataforma, modelos, inferencia, almacenamiento y consumo.
- Seguimiento de adopción, reutilización y excepciones.
- Identificación de oportunidades de mejora.
- Elaboración de informes periódicos de actividad, avance, riesgos y resultados.

3.2 Alcance tecnológico

El entorno tecnológico a cubrir será híbrido, con presencia de infraestructura on-premises y cloud pública, con mayor protagonismo de AWS como nube principal. El servicio deberá contemplar, al menos:

- Infraestructura cloud en AWS como proveedor principal, complementado por GCP y con posibilidad de incluir Azure en fases posteriores.
- Integración con entornos on-premises.
- Servicios de cómputo, almacenamiento, red, seguridad e identidad;
- Componentes de plataforma IA.
- Capacidades de GenAI.
- Modelos fundacionales, modelos open source y/o servicios gestionados.
- Pipelines de datos y conocimiento. Tratamiento inteligente de documentos.
- Vectorización, recuperación semántica y RAG.
- Observabilidad, logging, trazabilidad y auditoría.
- Mecanismos de FinOps específicos e integración con los generales de EJIE.
- Integración con herramientas corporativas existentes.
- Soporte a despliegues seguros y gobernados en diferentes entornos como laboratorio, pruebas y producción.

Es decir, la plataforma de IA debe incluir servicios, patrones y controles compartidos para acceso a modelos, cómputo, pipelines de datos, seguridad, guardrails, observabilidad y gestión, de modo que múltiples casos de uso a su vez de diferentes tenants/clientes puedan consumirlos de forma consistente.

4 Objeto del Servicio y Líneas de Trabajo

El objeto del contrato es la prestación de un servicio especializado para consolidar, evolucionar, operar y dar soporte al Centro de Excelencia de IA y a la plataforma corporativa de IA de EJIE, asegurando su alineamiento con la estrategia de IA, el modelo de gobierno existente, la arquitectura corporativa, los requisitos de seguridad y cumplimiento, y las necesidades de los casos de uso impulsados por los clientes de EJIE u otras áreas del propio EJIE.

El servicio deberá estructurarse en las siguientes líneas de trabajo.

4.1 Línea 1: Oficina Técnica del CoE-IA y Modelo Operativo

Esta línea tendrá como finalidad apoyar la consolidación del Centro de Excelencia de IA como función transversal de coordinación, gobierno técnico, arquitectura, plataforma, acompañamiento y mejora continua.

Actividades mínimas:

- Definir y mantener el modelo operativo objetivo del AI CoE.
- Mantener el mapa de roles, responsabilidades y relaciones con otras áreas.
- Coordinar la relación entre CoE-IA, Gobierno del Dato, Seguridad, Arquitectura, Otros Servicios Convergente, CoE-Cloud, Aplicaciones, PMO y áreas demandantes.
- Elaborar procedimientos, plantillas, checklists y guías de trabajo.
- Apoyar la preparación de comités, informes y documentación de seguimiento.
- Asegurar la adhesión al modelo de gobierno y gestión de casos de uso ya definido.
- Proponer mejoras al modelo operativo en función de la evolución de la demanda y madurez.

Entregables:

- Modelo operativo actualizado del CoE-IA
- Mapa de roles y responsabilidades.
- Procedimiento de coordinación con gobierno de IA y gestión de casos de uso.
- Informe periódico de actividad.
- Plan de Gestión del Cambio, para adaptar las dinámicas organizativas al marco de este contrato.
- Registro de decisiones, riesgos, dependencias y acciones.

4.2 Línea 2: Gestión y Cualificación de Demanda de IA

Esta línea tendrá como finalidad asegurar que los casos de uso de IA se analizan, clasifican y enrutan correctamente antes de su financiación, diseño técnico detallado o ejecución. Es decir, se cualifica la demanda de IA de forma temprana mediante un tratamiento estructurado, antes de que se convierta en proyecto o despliegue como tal, para clasificar solicitudes, identificar oportunidades de reutilización y determinar qué necesidades deben ir a la plataforma compartida. Actualmente EJIE ya

dispone de circuitos establecidos para gestionar iniciativas y su gobierno que el adjudicatario deberá asumir, mantener y evolucionar.

Actividades mínimas:

- Participar en el proceso de recepción de casos de uso.
- Analizar viabilidad técnica, datos requeridos, sensibilidad, escala, criticidad y expectativas de servicio.
- Atender a las implicaciones según los niveles de riesgo conforme a la regulación y las políticas de gobierno.
- Clasificar cada solicitud según patrón/arquetipo
- Identificar capacidades de plataforma reutilizables.
- Detectar necesidades no cubiertas por la plataforma.
- Proponer excepciones temporales cuando proceda.
- Apoyar la priorización desde criterios de valor, riesgo, complejidad, reutilización y alineamiento estratégico.
- Coordinarse con el modelo de gobierno y gestión de casos de uso ya existente.

Conjunto mínimo de patrones/arquetipos:

Arquetipo	Descripción	Implicaciones plataforma
Experimentación	Pruebas controladas de modelos, prompts, RAG o agentes	Entornos sandbox, capacidad compartida, límites de consumo y datos gobernados.
Fine-tuning / adaptación	Adaptación de modelos a dominios o tareas específicas	Pipelines, datasets gobernados, versionado, capacidad específica y controles reforzados.
Inferencia batch	Procesos diferidos de clasificación, extracción, resumen o análisis documental	Priorización de coste, ventanas de ejecución, trazabilidad y completitud.
Inferencia en tiempo real	APIs, asistentes, copilotos o servicios interactivos	Requisitos de latencia, disponibilidad, logging, cuotas y soporte operativo.
Workflows agénticos	Flujos con agentes que planifican, usan herramientas o coordinan acciones	Orquestación, memoria, estado, supervisión humana, auditoría y control de herramientas.

Entregables:

- Ficha de cualificación de caso de uso.
- Matriz de clasificación de demanda.
- Informe de encaje con plataforma.
- Informe de riesgos, dependencias y requisitos.
- Recomendación de patrón de carga y servicio de plataforma aplicable.
- Registro de excepciones.

4.3 Línea 3: Ingeniería y evolución de la Plataforma de IA

Esta línea constituye uno de los pilares principales del contrato. La plataforma de IA será tratada como un producto interno corporativo, orientado a proporcionar capacidades comunes, reutilizables, seguras y gobernadas para múltiples casos de uso.

Siguiendo el paradigma de *Infrastructure Platform Engineering*, se busca no diseñar de forma específica para cada proyecto, caso de uso o aplicación, sino que se identifican patrones recurrentes y alinean servicios aprobados a dichos patrones para que los equipos puedan avanzar con mayor consistencia y menor diseño a medida.

Actividades mínimas:

- Asumir la plataforma de IA existente.
- Analizar su estado actual, arquitectura, componentes, automatizaciones, controles y limitaciones.
- Definir un roadmap de evolución.
- Vigilancia tecnológica para estar al día de la evolución del contexto la IA en el mercado, y poder identificar “*best practices*”, tendencias, evoluciones que incorporar.
- Diseñar y mantener un catálogo interno de servicios aprobados.
- Implementar, mantener y evolucionar los servicios, plantillas, automatizaciones y componentes técnicos que materialicen los golden paths definidos en la línea de arquitectura y patrones.
- Evolucionar servicios de acceso a modelos, inferencia, RAG, agentes, observabilidad, seguridad, integración y FinOps.
- Automatizar despliegues y provisión de capacidades bajo un paradigma de IaC/GitOps.
 - Existe un marco de desarrollo (en adelante MD-IaC) que estandariza el uso de la Infraestructura como Código (IaC) en entornos cloud de EJIE. Este marco integra las plantillas y módulos de IaC (principalmente con Terraform), sobre los pipelines CI/CD corporativos (GitLab), siguiendo un modelo GitOps.
 - Estas implementaciones técnicas se podrán consumir como building blocks desde el portal interno del desarrollador de EJIE BatArin (Basado en Backstage)
 - Las plantillas deberán alinearse con mejores prácticas (p.ej., mantener la separación plan/apply, usar OIDC, asegurar el estado de Terraform). El adjudicatario deberá implantar, aplicar y mantener este marco en la organización, abordando: convenciones de naming y tagging, gestión de secretos, bibliotecas de módulos reutilizables, y CI/CD pipeline con validaciones automáticas de seguridad y calidad en cada push (e.g., linting, análisis SAST/DAST, integration tests, etc.). Esta línea de trabajo incluye la gobernanza del marco: asegurar su adopción por todos los proyectos cloud y actualizarlo continuamente conforme a nuevas mejores prácticas (por ejemplo, integración de recomendaciones del Well-Architected Framework, best practices de AWS y GCP, etc.)
- Incorporar controles de identidad, seguridad, datos, logging, trazabilidad y costes.
- Definir niveles de servicio diferenciados según patrón/arquetipo.
- La plataforma tendrá que cubrir distintos requisitos no funcionales que guiarán las soluciones técnicas para adecuarlas a las características de distintos tipos de los casos de uso a cubrir:
 - Rendimiento, disponibilidad, escalabilidad, capacidad de recuperación ante desastres
- Monitorizar factores clave como adopción, solicitudes fallidas, bypass, excepciones y salud del servicio.
- Revisión tecnológica (al menos semestral) que permita reorientar arquitecturas, modelos, frameworks y buenas prácticas que eran referencia en su momento pero hayan podido quedar obsoletas en meses, dado el ritmo excepcional al que evoluciona el mercado de la IA.
- Asegurar la documentación técnica y funcional de la plataforma.

La plataforma deberá cubrir como mínimo las diferentes funcionalidades:

La plataforma deberá evolucionar para cubrir, como mínimo, los siguientes elementos:

Capa	Capacidad / elemento	Descripción	Ejemplos prácticos
Acceso y gobierno	Identidad y permisos	Gestión de identidades, roles, permisos y control de acceso a la plataforma, modelos, datos, herramientas y casos de uso.	Integración con Entra ID/LDAP; roles diferenciados para desarrolladores, equipos funcionales, administradores, auditores y aplicaciones consumidoras.
	Políticas de uso	Definición y aplicación de reglas de uso corporativo, límites, autorizaciones, condiciones de acceso y criterios de cumplimiento.	Política que impide enviar datos sensibles a modelos externos; límites de tokens por usuario; restricción de modelos según criticidad del caso de uso.
	Registro de casos de uso	Inventario gobernado de casos de uso de IA, con responsables, finalidad, criticidad, datos utilizados, modelos empleados y estado de aprobación.	Ficha de caso de uso para “asistente de consulta normativa”, indicando propietario, datos tratados, modelo autorizado, riesgos y estado de validación.
	Catálogo de modelos aprobados	Catálogo de modelos autorizados para uso corporativo, con reglas de uso, trazabilidad, criterios de aprobación y condiciones de consumo.	Catálogo con Claude, GPT, Llama, Mistral o modelos internos, indicando para qué usos están permitidos y en qué condiciones.
	Auditoría y trazabilidad	Capacidades para registrar accesos, decisiones, ejecuciones, consumos, cambios de configuración y evidencias de cumplimiento.	Registro de qué aplicación invocó qué modelo, cuándo, con qué coste, con qué usuario o servicio técnico y bajo qué política aplicada.
Hub de modelos e inferencia	Gateway de acceso a modelos aprobados	Servicio de acceso controlado a modelos internos, externos, propietarios u open source, actuando como punto único de entrada para aplicaciones y servicios consumidores.	API corporativa tipo LiteLLM, Portkey, Kong AI Gateway o similar para enrutar peticiones a Bedrock, Azure OpenAI, modelos propios o proveedores externos.
	Acceso gobernado a patrones / arquetipos	Servicio de acceso controlado a plataformas, patrones o servicios específicos de IA de mayor valor añadido que el simple acceso a modelos.	Plantillas corporativas para RAG, resumen documental, clasificación automática, extracción de entidades,

			generación de borradores o asistentes conversacionales.
	Guardrails	Aplicación de controles de seguridad, filtrado, moderación, protección de datos, restricciones de salida y cumplimiento de políticas durante el uso de modelos.	Bloqueo de prompts con datos personales; detección de fuga de información; filtrado de respuestas no permitidas; validación de formato de salida.
	Acceso a modelos externos	Integración gobernada con proveedores externos de modelos fundacionales o servicios de IA.	Consumo de modelos en AWS Bedrock, Azure AI Foundry, Google Vertex AI o APIs SaaS, con control contractual, trazabilidad y FinOps.
	Modelos open source	Soporte para el consumo, evaluación, despliegue o integración de modelos open source autorizados.	Despliegue interno de Llama, Mistral, Qwen o DeepSeek sobre GPU corporativa usando vLLM, TGI, NIM u otro runtime aprobado.
	Inferencia batch	Ejecución diferida de procesos de IA, con logging, control de coste, trazabilidad, planificación y gestión de cargas masivas.	Procesamiento nocturno de miles de expedientes para clasificación, resumen, extracción de metadatos o anonimización previa.
	Inferencia en tiempo real	Exposición de endpoints productivos con requisitos de disponibilidad, latencia, cuotas, escalabilidad, seguridad y soporte operativo.	Endpoint para un chatbot corporativo, asistente de atención ciudadana o copiloto interno con SLA, autoscaling y cuotas por aplicación.
RAG y conocimiento	Conectores a fuentes	Integración con repositorios documentales, bases de datos, APIs, sistemas corporativos y otras fuentes de conocimiento.	Conectores a SharePoint, , BBDDs Oracle/PostgreSQL, APIs internas, gestor documental, data lake o repositorios normativos.
	IDP / Tratamiento de Documentos	Capacidades de ingesta, limpieza, extracción, transformación, particionado y enriquecimiento de documentos.	OCR de PDFs escaneados, extracción de tablas, chunking por secciones, limpieza de cabeceras/pies y enriquecimiento con metadatos.
	Embeddings	Generación y gestión de representaciones vectoriales de textos, documentos u otros contenidos.	Generar embeddings de normativa, procedimientos internos, tickets históricos o documentación técnica para búsqueda semántica.

	Vectorización	Procesos de indexación vectorial, almacenamiento en bases vectoriales y actualización de índices.	Indexación en OpenSearch, pgvector, Milvus, Weaviate, Pinecone o Redis Vector de documentación departamental.
	Recuperación semántica	Capacidades de búsqueda semántica, grounding, selección de contexto relevante y recuperación aumentada de información.	Un asistente responde preguntas citando únicamente fragmentos recuperados de normativa vigente o documentación validada.
	Citas y procedencia	Identificación de fuentes, referencias, evidencias, trazabilidad documental y control de procedencia.	Cada respuesta incluye enlaces al documento fuente, página, sección, fecha de actualización y grado de confianza.
Agentes y orquestación	Workflows	Orquestación de pasos, herramientas, usuarios, aprobaciones y procesos de IA.	Flujo que analiza una solicitud, consulta normativa, genera propuesta, solicita validación humana y registra la decisión.
	Herramientas	Integración de herramientas, APIs, servicios corporativos y acciones externas invocables por modelos o agentes.	Agente que consulta un expediente, abre un ticket ITSM, consulta una API de inventario o ejecuta una búsqueda en un sistema documental.
	Memoria y contexto	Gestión de memoria, estado, contexto conversacional, preferencias, histórico y datos necesarios para agentes persistentes.	Asistente que recuerda el expediente o trámite sobre el que se está trabajando durante una sesión, sin exponer información fuera del ámbito autorizado.
	Human-in-the-loop	Capacidades para incorporar validación, supervisión, aprobación o intervención humana.	Una propuesta generada por IA requiere aprobación de un técnico antes de enviarse al ciudadano o incorporarse al expediente.
	Control de acciones	Gobernanza, autorización, trazabilidad y supervisión de las acciones ejecutadas por agentes.	El agente puede consultar información, pero necesita autorización explícita para modificar datos, enviar comunicaciones o ejecutar operaciones en sistemas.
LLMOps / MLOps	Investigación y desarrollo de modelos	Marco para desarrollo de modelos y técnicas como fine tuning, cuantización, evaluación experimental o adaptación de modelos.	Laboratorio para probar fine tuning de un modelo open source con documentación interna o evaluar

			cuantización para reducir consumo GPU.
	Entorno de experimentación IA	Sandbox para pruebas controladas, con límites de consumo, datos permitidos, aislamiento y mecanismos de promoción.	Espacio para que un equipo pruebe un caso de uso RAG con datos anonimizados, presupuesto limitado y modelos no productivos.
	Versionado	Gestión de versiones de modelos, prompts, datasets, configuraciones, pipelines, evaluaciones y artefactos.	Versionar el prompt de un asistente, el dataset de evaluación, la configuración de retrieval y el modelo usado en cada release.
	Evaluación de modelos y prompts	Validación de calidad, seguridad, groundedness, robustez, sesgos, regresión, rendimiento y adecuación funcional.	Batería de pruebas para comparar Claude, GPT y Llama en precisión, coste, latencia, alucinaciones y cumplimiento de instrucciones.
	Despliegue	Mecanismos para promocionar, publicar y operar modelos, prompts, agentes o pipelines desde prueba hacia producción.	Pipeline CI/CD que despliega una nueva versión de un asistente tras superar pruebas funcionales, seguridad y evaluación de calidad.
	Monitorización	Seguimiento técnico y funcional de modelos, prompts, endpoints, pipelines y agentes.	Panel con latencia, errores, consumo de tokens, evolución de calidad, tasa de fallback y satisfacción de usuarios.
	Rollback	Capacidad para revertir versiones de modelos, prompts, configuraciones o despliegues.	Volver automáticamente a la versión anterior de un prompt si aumenta la tasa de respuestas incorrectas o se degrada la calidad.
Observabilidad y FinOps	Logs	Registro centralizado de invocaciones, errores, entradas, salidas, eventos relevantes y evidencias operativas.	Logs anonimizados de prompts y respuestas, eventos de guardrails, errores de proveedor, trazas de ejecución y auditoría de acceso.
	Métricas	Captura de indicadores técnicos, funcionales y económicos.	Métricas de latencia p95/p99, tokens por petición, coste por aplicación, tasa de error, throughput y disponibilidad.

	Calidad	Monitorización de calidad de respuestas, groundedness, exactitud, regresión, drift y comportamiento funcional.	Evaluar periódicamente si un asistente sigue respondiendo correctamente tras cambios en documentos, modelos o prompts.
	Coste	Visibilidad del consumo y coste por caso de uso, modelo, patrón, unidad funcional, aplicación o proveedor.	Dashboard de coste mensual por departamento, modelo, aplicación, entorno y tipo de operación: chat, RAG, embeddings o batch.
	Uso	Análisis de adopción, frecuencia de uso, usuarios, aplicaciones consumidoras y evolución del consumo.	Seguimiento de usuarios activos, casos de uso más utilizados, aplicaciones con mayor consumo y patrones infrautilizados.
	Salud del servicio	Monitorización de disponibilidad, errores, capacidad, saturación, degradaciones, incidencias y cumplimiento de SLA.	Alertas cuando aumenta la latencia del gateway, falla un proveedor externo, se agota una cuota o se degrada un endpoint productivo.

Entregables:

- Catálogo de servicios aprobados.
- Golden paths como implementaciones técnicas reutilizables.
- Automatización de despliegues.
- Componentes de plataforma.
- Servicios de acceso a modelos.
- Servicios RAG comunes.
- Observabilidad técnica de plataforma.
- Integración con IaC/GitOps.
- Evolución técnica y roadmap.
- Documentación técnica de plataforma.

4.4 Línea 4: Arquitectura, patrones y diseño de soluciones de IA

Esta línea tendrá como finalidad definir, mantener y aplicar arquitecturas de referencia y patrones reutilizables (Golden-Paths) para los distintos tipos de casos de uso de IA.

Esta línea no tendrá como finalidad la construcción directa de componentes de plataforma, sino la definición, revisión y gobierno técnico de los patrones de arquitectura que deben seguir tanto la plataforma como los casos de uso consumidores.

Actividades mínimas:

- Definir arquitecturas de referencia para:
 - RAG.
 - Asistentes y copilotos.
 - Clasificación y extracción documental.
 - Inferencia batch.

- Inferencia en tiempo real.
- Workflows agenticos
- Agentes con supervisión humana.
- Experimentación segura con modelos
- Despliegue para procesamiento HPC Fine-Tuning
- Integración con sistemas corporativos.
- Evaluación y observabilidad.
- Revisar diseños propuestos por proyectos.
- Asegurar alineamiento con la plataforma corporativa.
- Definir criterios de selección de modelos y componentes.
- Evaluar alternativas cloud AWS (u otros hiperescalares), on-premises o híbridas.
- Documentar decisiones de arquitectura.
- Coordinarse con el CoE de Cloud, arquitectura, seguridad, datos e integración.

Entregables:

- Arquitecturas de referencia.
- Patrones de solución.
- Guías de diseño.
- Revisión de arquitectura de casos de uso.
- Matrices/árboles de decisión tecnológica.
- Criterios de selección de modelos/componentes.
- Registro de decisiones de arquitectura.
- Alineamiento con arquitectura empresarial, seguridad, datos y cloud.

4.5 Línea 5: Ingeniería IA, MLOps, LLMOps y operación

Esta línea tendrá como finalidad facilitar el paso de prototipos y pilotos desarrollados por los clientes, proveedores u otras áreas de EJIE a servicios productivos operables, monitorizados y mantenibles.

Actividades mínimas:

- Definir pipelines de despliegue y operación integrados con los mecanismos de SDLC de EJIE.
- Gestionar versiones de modelos, prompts, embeddings, configuraciones y datasets.
- Automatizar pruebas y despliegues.
- Implantar evaluación continua.
- Monitorizar calidad, coste, latencia, disponibilidad, errores y seguridad.
- Definir procedimientos de rollback.
- Integrar la plataforma con los procesos de gestión de incidencias, problemas y cambios alineados con ITIL que EJIE tiene implantados.
- Apoyar el soporte de segundo y tercer nivel especializado que pueda ser de aplicación sobre componentes técnicos o productos sobre los que se apoya la plataforma.

Entregables:

- Framework LLMOps/MLOps.
- Pipelines de despliegue.
- Guías de operación.
- Cuadros de mando técnicos.
- Informes de incidencias y problemas.

- Registro de cambios.
- Procedimientos de recuperación y rollback.

4.6 Línea 6: Seguridad, privacidad, cumplimiento e IA responsable

Esta línea tendrá como finalidad asegurar que las soluciones de IA se diseñan, despliegan y operan con controles adecuados de seguridad, privacidad, cumplimiento y responsabilidad, alineadas con el marco de gobernanza para la IA de EJIE y del Gobierno Vasco. De esta forma se asegura que se cubre la necesidad de mitigar riesgos de seguridad, éticos y legales, se asegura el gobierno del dato y de los modelos, se facilita observabilidad, auditoría y rendición de cuentas, y se permite una experimentación controlada y escalable.

Actividades mínimas:

- Definir y mantener controles técnicos de IA segura.
- Evaluar riesgos específicos de IA y GenAI.
- Revisar riesgos como la fuga de datos, prompt injection, jailbreak, uso indebido de herramientas y exposición de información.
- Coordinarse con otras áreas o funciones como Seguridad Corporativa, Seguridad y Gobierno del dato.
- Apoyar evaluaciones de impacto cuando proceda.
- Definir criterios de supervisión humana.
- Asegurar trazabilidad, auditoría y rendición de cuentas.
- Revisar proveedores, modelos y servicios externos desde el punto de vista de seguridad y cumplimiento.
- Asegurar el alineamiento con RGPD, AI Act, ENS y normativa aplicable.

Entregables:

- Checklists de seguridad IA.
- Matriz de riesgos IA.
- Informe de cumplimiento técnico.
- Guías de uso responsable.
- Informe de revisión de proveedores/modelos.
- Registro de controles y evidencias.
- Recomendaciones de mitigación.

4.7 Línea 7: Acompañamiento a casos de uso, adopción y formación

Esta línea tendrá como finalidad asegurar que la plataforma y el modelo operativo generan adopción real y valor para los clientes u otras áreas de EJIE que la consuman.

Actividades mínimas:

- Acompañar a promotores de casos de uso.
- Facilitar talleres de ideación, cualificación y diseño.
- Apoyar pruebas de concepto y pilotos.
- Definir criterios de paso a producción.

- Elaborar guías de uso para perfiles técnicos y funcionales.
- Diseñar acciones de formación.
- Dinamizar comunidades de práctica.
- Apoyar a usuarios clave (*"AI Champions"*) o referentes distribuidos.
- Medir satisfacción, adopción y controlar barreras limitantes o bloqueantes.

Entregables:

- Plan de adopción.
- Materiales formativos.
- Guías de uso.
- Registro de sesiones.
- Informe de adopción y satisfacción.
- Recomendaciones de mejora.
- Recomendaciones de mitigación.

4.8 Línea 8: Desarrollo e Implementación de casos de uso extremo a extremo

Esta línea complementa a la línea 7, que se dará en casos puntuales para iniciativas que no sean de mucha envergadura o en las que el cliente no disponga de recursos o capacidades para desarrollar la aplicación o sistema que permita cubrir el caso de uso que se desea implantar entregándolo en formato proyecto.

Actividades mínimas:

- Analizar y concretar los requisitos funcionales y técnicos del caso de uso.
- Diseñar la solución extremo a extremo apoyándose en los patrones, arquitecturas y servicios homologados de la plataforma corporativa de IA.
- Desarrollar, configurar e integrar los componentes necesarios para cubrir el caso de uso.
- Implementar asistentes, copilotos, soluciones RAG, clasificación documental, extracción de información, agentes, workflows o patrones equivalentes según las necesidades del caso.
- Desarrollar conectores e integraciones con aplicaciones, APIs, repositorios documentales, sistemas corporativos, plataformas de datos u otros servicios necesarios.
- Preparar, transformar, enriquecer e integrar las fuentes de información requeridas para la solución.
- Configurar modelos, prompts, guardrails, mecanismos de observabilidad, controles de acceso y capacidades de trazabilidad.
- Realizar pruebas funcionales, técnicas, de rendimiento, seguridad y aceptación junto con el área promotora.
- Gestionar el paso entre entornos de laboratorio, pruebas y producción utilizando los pipelines y mecanismos LLMOps/MLOps corporativos.
- Elaborar la documentación técnica y funcional necesaria para la operación, soporte y mantenimiento de la solución.
- Transferir conocimiento a los equipos promotores, responsables funcionales y equipos de operación.
- Proporcionar soporte o acompañamiento posterior a la puesta en producción durante el período que se determine para cada implantación.
- Medir los resultados obtenidos y proponer mejoras evolutivas para incrementar el valor generado.

Entregables:

- Documento de análisis y diseño de la solución.
- Arquitectura técnica específica del caso de uso.
- Backlog y planificación de implantación.
- Componentes desarrollados, configuraciones, código fuente y artefactos generados.
- Conectores, integraciones y pipelines implementados.
- Configuración de modelos, prompts, agentes, workflows y guardrails.
- Casos de prueba, evidencias de validación y resultados de aceptación.
- Documentación técnica, funcional y operativa.
- Manuales de usuario y materiales de transferencia de conocimiento.
- Informe de puesta en producción.
- Informe de estabilización post-arranque e incidencias detectadas.
- Informe de resultados, adopción, valor generado y recomendaciones de evolución.

La ejecución de esta línea de trabajo tendrá carácter subsidiario respecto al modelo general de acompañamiento y habilitación promovido por el CoE-IA. Su utilización se reserva para iniciativas de alcance limitado, aceleradores corporativos, prototipos avanzados, demostradores, necesidades internas de EJIE o situaciones en las que el promotor como se dice más arriba carezca de capacidades suficientes para ejecutar directamente la solución, evitando que esta línea se convierta en un mecanismo alternativo para la ejecución de proyectos de desarrollo de gran tamaño.

4.9 Línea 9: Soporte especializado en producción y operación de la plataforma IA

Esta línea tendrá como finalidad asegurar el soporte especializado, operación técnica, gestión de incidencias, problemas, cambios y estabilización de la plataforma corporativa de IA y de los servicios IA desplegados sobre ella, una vez se encuentren en entornos productivos o preproductivos sujetos a operación.

El adjudicatario deberá prestar soporte especializado sobre las capas funcionales y técnicas propias de la plataforma IA, incluyendo componentes de acceso a modelos, orquestación, RAG, agentes, pipelines de IA, prompts, embeddings, guardrails, observabilidad específica, evaluación, trazabilidad, configuración, servicios reutilizables y artefactos desplegados sobre la plataforma.

Quedan fuera de la responsabilidad directa del adjudicatario las capas inferiores de infraestructura gestionadas por otros de los que dispone EJIE, tales como infraestructura física, virtualización/kubernetes, almacenamiento, networking, landing zones cloud, servicios cloud gestionados de base, conectividad, backup corporativo, seguridad perimetral o capacidades IaaS/PaaS subyacentes.

No obstante, el adjudicatario deberá coordinarse con dichos equipos, aportar diagnóstico especializado, evidencias, logs, métricas, trazas, análisis técnico y recomendaciones cuando una incidencia pueda estar relacionada con dependencias de infraestructura, cloud, red, identidad, seguridad, observabilidad corporativa o servicios gestionados externos.

Actividades mínimas:

- Prestar soporte especializado de segundo y tercer nivel sobre la plataforma corporativa de IA y los servicios IA desplegados sobre ella.
- Atender, analizar, diagnosticar y resolver incidencias relacionadas con componentes propios de la plataforma IA.
- Gestionar problemas recurrentes, errores conocidos, degradaciones, deuda operativa y acciones correctivas.
- Coordinar cambios operativos sobre componentes de IA, configuraciones, modelos, prompts, pipelines, guardrails, conectores, servicios de inferencia y servicios reutilizables de plataforma.
- Ejecutar procedimientos de recuperación, rollback, mitigación, escalado y estabilización ante incidencias relevantes.
- Coordinar la resolución de incidencias con los equipos responsables de infraestructura on-premises, cloud pública, hiperescalares, red, identidad, seguridad, observabilidad corporativa, service desk y otros servicios transversales de EJIE.
- Aportar evidencias técnicas, logs, métricas, trazas, análisis de causa raíz y recomendaciones cuando una incidencia afecte o pueda afectar a capas gestionadas por otros equipos.
- Mantener runbooks, playbooks y procedimientos de operación para los servicios IA en producción.
- Supervisar la salud operativa de la plataforma IA, incluyendo disponibilidad, latencia, errores, consumo, saturación, calidad de respuesta, eventos de guardrails y estado de integraciones.
- Participar en la gestión de cambios, releases y despliegues productivos de componentes IA en coordinación con los procesos corporativos de EJIE.
- Realizar seguimiento de incidencias relevantes, problemas abiertos, acciones de mejora y riesgos operativos.
- Proponer mejoras con mentalidad SRE de resiliencia, automatización, observabilidad, seguridad y operabilidad derivadas de la experiencia en producción. Incluso evolucionar a técnicas de AIOps y automeremediación.
- Asegurar la correcta transferencia de conocimiento operativo a equipos de soporte, operación y responsables funcionales cuando proceda.
- Mantener una matriz de responsabilidades operativas que delimite claramente qué componentes son responsabilidad del adjudicatario y cuáles corresponden a otros equipos de soporte.

Entregables:

- Procedimiento de soporte y escalado de la plataforma IA.
- Matriz de responsabilidades operativas por capas, servicios y componentes.
- Catálogo de servicios IA en producción y responsables operativos asociados.
- Runbooks y playbooks de operación, diagnóstico, recuperación y rollback.
- Registro de incidencias, problemas, cambios y acciones correctivas.
- Informes de causa raíz para incidencias relevantes.
- Informes periódicos de salud operativa de la plataforma IA.
- Registro de problemas conocidos, riesgos operativos y deuda técnica.
- Registro de dependencias con equipos de infraestructura, cloud, seguridad, identidad, observabilidad y proveedores externos.
- Recomendaciones SRE de mejora operativa, resiliencia, automatización y observabilidad.
- Evidencias de soporte, trazabilidad, escalados y actuaciones realizadas.

4.10 Línea 10: FinOps, valor y mejora continua

Esta línea tendrá como finalidad medir y optimizar el coste, uso, valor y madurez del servicio y de la plataforma. Tiene por tanto una parte de “showback” y otra de “chargeback” que permita la imputación

de costes y facturación a los clientes integrada con los mecanismos más generales cuando la plataforma se consume bajo el contexto del catálogo de servicios.

Actividades mínimas:

- Alineamiento con las prácticas FinOps más generales de la Plataforma Cloud de EJIE.
- Definir indicadores de uso, coste, valor y eficiencia.
- Medir costes por modelo, servicio, patrón, caso de uso y unidad funcional.
- Identificar oportunidades de optimización.
- Medir reutilización de capacidades de plataforma.
- Medir excepciones y bypass.
- Medir tiempos desde recepción de una necesidad hasta disponer de entorno usable.
- Elaborar informes ejecutivos y técnicos.
- Proponer acciones de mejora continua.

Entregables:

- Cuadro de mando FinOps IA.
- Informe de valor generado.
- Informe de uso y adopción de plataforma.
- Informe de excepciones y bypass.
- Recomendaciones de optimización.
- Plan de mejora continua.

5 Modelo Operativo de IA que se pretende alcanzar

El modelo operativo objetivo se basará en un enfoque “Hub & Spoke”, en el que el CoE-IA actúa como núcleo central de coordinación, gobierno técnico, arquitectura, plataforma y habilitación, mientras que otras áreas funcionales, técnicas y de negocio participan mediante referentes distribuidos, equipos promotores y responsables de casos de uso. Este CoE-IA se coordinará a su vez con el CoE-Cloud en las partes que por su ámbito sean de aplicación. El CoE-IA no se compone únicamente de recursos de esta contratación, y tendrá un liderazgo y participación activa de personal de EJIE.

Este modelo deberá adherirse a la estrategia de IA, la estrategia Cloud y al modelo de gobierno y gestión de casos de uso ya establecidos.

5.1 Principios del modelo operativo objetivo

El servicio deberá contribuir a implantar y consolidar los siguientes principios:

1. **Adhesión al gobierno de IA existente** El contrato no deberá crear un modelo paralelo, sino integrarse con la estrategia, órganos, políticas, procesos y mecanismos de gestión de casos de uso definidos por EJIE.
2. **Plataforma común como eje de escalado:** La plataforma corporativa de IA será el mecanismo preferente para experimentar, desplegar y operar soluciones de IA.
3. **Cualificación temprana de la demanda:** Todo caso de uso deberá cualificarse antes de consolidar financiación, diseño técnico o provisión de infraestructura, identificando patrón de carga, datos, sensibilidad, dimensión, criticidad, requisitos de servicio y reutilización posible.
4. **Estandarización pragmática:** Se definirán arquetipos, patrones y servicios estándar sólo cuando exista demanda repetida o requisitos materialmente diferenciados (entorno de ejecución, control o nivel de servicio).
5. **Camino homologado más fácil que la excepción:** Los golden paths y servicios homologados deberán integrar controles de identidad, modelos, datos, logging, cuotas, costes y soporte, evitando que los equipos de desarrollo de los clientes tengan que recurrir a vías paralelas.
6. **Gobierno embebido en la plataforma:** Los controles de seguridad, privacidad, datos, costes y trazabilidad deberán estar integrados en los flujos de solicitud, despliegue y operación, y no depender exclusivamente de aprobaciones manuales externas.
7. **Modelo híbrido con protagonismo cloud:** El modelo deberá contemplar despliegues híbridos, con mayor peso de AWS (que en una siguiente fase se podrá extender a GCP y Azure), pero garantizando integración con entornos on-premises y compatibilidad con requisitos de soberanía, seguridad, integración y operación.
8. **Reutilización y reducción de duplicidades:** La plataforma deberá favorecer la reutilización de servicios, componentes, patrones y conocimientos, evitando soluciones específicas por proyecto.
9. **Operabilidad desde el diseño:** Todo caso de uso deberá considerar desde fases tempranas observabilidad, soporte, gestión de incidencias, niveles de servicio, seguridad, costes y ciclo de vida.
10. **Medición de valor y mejora continua:** El modelo deberá incorporar métricas de valor, coste, adopción, calidad, riesgo, reutilización, excepciones y salud de plataforma.

5.2 Gestión de Excepciones

En el contexto del servicio se deberá definir y operar un procedimiento de excepciones gobernadas inspirado en este planteamiento:

Situación	Tratamiento
El caso encaja en patrón existente	Uso obligatorio del servicio aprobado de plataforma.
El caso no encaja pero está justificado	Excepción temporal revisada por plataforma, seguridad, datos y arquitectura.
La misma excepción se repite	Evaluación para incorporarla al catálogo o retirarla.
La excepción genera riesgo no aceptable	Rechazo, rediseño o elevación al órgano de gobierno correspondiente.

Se utilizarán excepciones temporales cuando los arquetipos y patrones estándar no soporten una necesidad justificada, y se revisarán excepciones repetidas para decidir si la plataforma necesita un nuevo servicio, un cambio de política o un nivel de servicio diferente

6 Composición del Equipo de Trabajo

La empresa licitadora deberá proponer un equipo multidisciplinar con capacidad para cubrir gobierno técnico, arquitectura, plataforma, ingeniería, operación, seguridad, datos, adopción y gestión de valor.

Las dedicaciones indicadas a continuación son un **dimensionamiento anual**, expresado en horas/año. Podrán ser ajustadas por EJIE en función de necesidad, demanda real, modelo definitivo de prestación. La oferta deberá justificar la adecuación de la dedicación propuesta para cubrir los trabajos exigidos.

6.1 Equipo propuesto

Rol	Descripción	Dedicación Fija (horas/año)	Dedicación Variable (horas/año)
Responsable del Servicio	Responsable global del contrato, coordinación con EJIE, seguimiento de compromisos, planificación, riesgos, calidad y relación con órganos de gobierno.	900	
Product Owner Plataforma de IA	Responsable funcional y técnico del roadmap de la plataforma de IA como producto interno; prioriza catálogo, golden paths, adopción, mejoras y coordinación con stakeholders.	1.000	
Arquitecto de IA	Define arquitecturas de referencia, patrones, decisiones tecnológicas, criterios de integración y alineamiento con arquitectura corporativa, cloud, seguridad y datos.	1.000	
Consultor Infraestructura IA y Cloud	Lidera la evolución de la infraestructura IA en AWS y on-premises, definiendo capacidad, conectividad, resiliencia, automatización y operación.		900
Ingeniero Plataforma IA	Implementa y evoluciona componentes de plataforma, automatizaciones, servicios reutilizables, despliegues, integración y plantillas técnicas.	1.200	
Ingeniero LLMOps/MLOps	Define e implementa pipelines de despliegue, versionado, evaluación, monitorización, rollback y operación de modelos, prompts, RAG y agentes. Participación en incidencias complejas y/o problemas.		900
Ingeniero Implantación de IA	Desarrolla e integra soluciones de IA, RAG, agentes, APIs, workflows y componentes de inferencia sobre la plataforma corporativa.		1.200

Desarrollador de Soluciones IA	Perfil técnico senior encargado del desarrollo, implementación e integración extremo a extremo de casos de uso de IA en el marco de la Línea 8. Desarrolla componentes RAG, agentes, asistentes, APIs, conectores, workflows, automatizaciones, integraciones con sistemas corporativos y configuraciones sobre la plataforma corporativa de IA. Además, coordina técnicamente la implantación cuando sea necesario, asegurando alineamiento con arquitectura, seguridad, datos, LLMOps/MLOps, operación y adopción		1.200
Ingeniero de Datos	Diseña pipelines de datos, integración de fuentes, preparación de conocimiento, vectorización, metadatos, calidad y soporte a RAG.		700
Consultor Gobernanza IA, Riesgo y Cumplimiento	Asegura adhesión al modelo de gobierno, evaluación de riesgos, cumplimiento, IA Responsable, trazabilidad y coordinación con Seguridad Corporativa, Delegado de Protección, etc.	1.700	
Ingeniero de Seguridad IA	Revisa seguridad específica de IA, amenazas GenAI, controles de acceso, logging, fuga de datos, prompt injection, uso de herramientas y despliegues seguros.		700
Consultor de Observabilidad IA	Define y opera observabilidad, salud del servicio, cuadros de mando, gestión de incidencias, problemas, cambios y niveles de servicio.	400	400
Ingeniero de Soporte y Operación de la Plataforma IA	Perfil técnico responsable del soporte especializado en producción de la plataforma corporativa de IA y de los servicios IA desplegados sobre ella. Atiende incidencias, problemas, cambios operativos, análisis de causa raíz, escalados, recuperación, rollback, estabilización, coordinación con equipos de infraestructura/cloud y mantenimiento de runbooks/playbooks operativos.	800	400
Consultor de FinOps IA	Mide coste, consumo, reutilización, eficiencia, valor generado, excepciones y optimización económica de la plataforma y casos de uso.	400	
Consultor de Negocio IA	Apoya recepción de iniciativas, cualificación y traducción de necesidades funcionales a casos de uso viables, gobernables y medibles.		600
Consultor de Adopción y Formación IA	Define materiales, formación, guías, comunidad de práctica, acompañamiento a usuarios y AI Champions.		1200
Estrategia y consultoría	Especialización de nivel muy alto y experiencia en el ámbito de las tecnologías cloud definidas en el entorno tecnológico, prestando labores de asistencia técnica y complementando al resto de roles.		2.000

	Estos perfiles intervendrán bajo demanda, previa valoración por jornadas aceptada expresamente por EJIE y se facturarán al precio ofertado en el mes siguiente en el que se hayan prestado. El nivel de dedicación de la tabla se debe interpretar por tanto para este caso como un máximo.		
--	---	--	--

Adicionalmente las horas de cada perfil, se aportarán otras **2.000 horas anuales para un posible servicio de consultoría adicional** que se prestará mediante los recursos especializados que proporcione el adjudicatario. Para la ejecución de estas horas, se seguirá la siguiente metodología:

- Mensualmente EJIE establecerá tareas de consultoría que se ejecutarán bajo petición para resolver una necesidad concreta que surja en el ámbito objeto del contrato.
- Una vez recibida la petición, la empresa adjudicataria presentará un plan detallado de recursos y horas para ejecutar dichas tareas, que tendrá que ser aprobado por EJIE previamente a su ejecución.
- La empresa adjudicataria asumirá los retrasos que se deriven de la ejecución de las tareas de consultoría acordadas
- La dedicación se facturará al precio ofertado y por servicio efectivamente ejecutado con el visto bueno de EJIE en el mes siguiente en el que se hayan prestado. El nivel de dedicación de la tabla se debe interpretar, por tanto, para este caso, como un máximo.

Estos volúmenes son los que se han utilizado para calcular el presupuesto máximo.

Nota: Los roles anteriores pueden ser cubiertos por uno o varios profesionales, siempre que se garantice una cobertura completa de las competencias y dedicaciones mínimas señaladas. Por ejemplo, algunos roles pueden ser asumidos por una misma persona con perfil mixto, siempre que esta posea la experiencia acreditada en ambas áreas y la dedicación total sea la suma indicada. En cualquier caso, los licitadores deberán proponer una organización del equipo que incluya al menos estos roles y justifique las horas anuales asignadas a cada uno, de modo de asegurar la disponibilidad del conocimiento y la capacidad de respuesta necesarias para cumplir con las obligaciones del pliego.

6.2 Roles Clave y Responsabilidades Detalladas

6.2.1 Responsable del Servicio

Responsable de la correcta prestación del servicio, interlocución principal con EJIE, coordinación del equipo, cumplimiento de compromisos, gestión de riesgos y calidad de entregables.

Responsabilidades:

- Coordinar el servicio.
- Planificar actividades y recursos.
- Asegurar cumplimiento de SLAs y entregables.
- Gestionar riesgos y dependencias.
- Preparar informes de seguimiento.
- Coordinar comités y reuniones.
- Asegurar la alineación con la estrategia y gobierno de IA de EJIE.

- Apoyo documental, actas, informes, repositorio documental, seguimiento de entregables, calidad y control administrativo.

6.2.2 Product Owner Plataforma de IA

Rol crítico en el marco de platform engineering para tratar la plataforma de IA como producto interno corporativo.

Responsabilidades:

- Mantener el roadmap de plataforma.
- Definir y priorizar el catálogo de servicios.
- Definir golden paths por patrón/arquetipo.
- Coordinar demanda, arquitectura, seguridad, datos y operación.
- Medir adopción, uso, excepciones, bypass y satisfacción.
- Gestionar backlog de evolución.
- Asegurar que el camino homologado sea más sencillo que las alternativas ad hoc.

6.2.3 Arquitecto de IA

Responsable de la coherencia arquitectónica de las soluciones de IA

Responsabilidades:

- Definir arquitecturas de referencia.
- Revisar diseños.
- Definir patrones RAG, agentes, inferencia, integración y observabilidad.
- Evaluar tecnologías.
- Asegurar alineamiento con cloud, datos, gobierno, seguridad y aplicaciones.
- Documentar decisiones de arquitectura.

6.2.4 Consultor Infraestructura IA/Cloud

Responsable de coordinar las capacidades de infraestructura IA en el entorno híbrido.

Responsabilidades:

- Diseñar la infraestructura IA sobre AWS y on-premises.
- Definir modelos de capacidad compartida y reservada.
- Coordinar redes, seguridad, identidad, conectividad y resiliencia.
- Integrar la plataforma IA con servicios cloud, la plataforma cloud y otras plataformas o sistemas corporativos.
- Participar desde la cualificación temprana de demanda.

6.2.5 Ingeniero de Plataforma IA

Responsable de construir y mantener servicios técnicos reutilizables de plataforma.

Responsabilidades:

- Implementar componentes de plataforma.
- Automatizar despliegues

- Mantener plantillas, APIs, catálogos y servicios
- Coordinación con otros equipos de referencia en ámbitos de plataformas, integración, ..
- Integrar logging, seguridad, identidad, datos y costes.
- Resolver incidencias técnicas.
- Apoyar a proyectos consumidores.

6.2.6 Ingeniero LLMOps / MLOps

Responsable de industrializar el ciclo de vida de soluciones IA.

Responsabilidades:

- Definir pipelines de despliegue.
- Gestionar versionado de modelos, prompts y configuraciones.
- Automatizar evaluación.
- Monitorizar calidad, coste, latencia y errores.
- Gestionar rollback.
- Integrar operación y mejora continua.

6.2.7 Ingeniero de Implantación de IA

Responsable de implementar soluciones de IA sobre la plataforma. En general servirá de acompañamiento a los equipos de proyecto de los promotores/desarrolladores de los casos de uso que se quieran implantar por parte de los clientes. En algunos casos podrán implementar casos de uso específicamente cuando el promotor carezca de recursos suficientes o sea un caso de uso interno de EJIE.

Responsabilidades:

- Construir componentes RAG.
- Integrar modelos con aplicaciones.
- Implementar agentes y workflows.
- Diseñar APIs y conectores y coordinarse con otros equipos de integración en este ámbito
- Crear prototipos y pilotos.
- Apoyar industrialización.

6.2.8 Desarrollador de Soluciones IA

Perfil responsable de desarrollar e implantar soluciones de IA extremo a extremo dentro del alcance de la Línea 8, especialmente en aquellos casos de uso de alcance limitado, prototipos avanzados, aceleradores corporativos, necesidades internas de EJIE o iniciativas en las que el promotor no disponga de recursos suficientes para construir directamente la solución.

A diferencia de otros perfiles orientados principalmente a arquitectura, plataforma, acompañamiento o gobierno, este rol tendrá una orientación clara a la ejecución técnica y al desarrollo efectivo de soluciones, participando directamente en la construcción de componentes software, integraciones y artefactos necesarios para poner en marcha el caso de uso.

Responsabilidades:

- Desarrollar componentes software para casos de uso de IA extremo a extremo.

- Implementar soluciones RAG, asistentes, copilotos, agentes, workflows, procesos batch, APIs de inferencia y servicios de integración.
- Desarrollar conectores con repositorios documentales, APIs corporativas, bases de datos, sistemas transaccionales, plataformas de datos u otras fuentes necesarias.
- Construir pipelines de ingesta, tratamiento documental, extracción, enriquecimiento, chunking, generación de embeddings y vectorización, en coordinación con el perfil de datos.
- Implementar prompts, plantillas, cadenas de razonamiento, configuraciones de modelos, herramientas invocables por agentes y mecanismos de control de contexto.
- Integrar guardrails, controles de acceso, trazabilidad, logging, auditoría, cuotas, límites de consumo y mecanismos de observabilidad definidos por la plataforma.
- Desarrollar automatizaciones, scripts, plantillas, infraestructura como código o configuraciones necesarias para desplegar la solución en los entornos definidos.
- Preparar y versionar código fuente, configuraciones, artefactos, prompts, datasets de prueba y componentes reutilizables.
- Ejecutar pruebas unitarias, funcionales, técnicas, de integración, regresión, rendimiento y validación de calidad de respuestas.
- Corregir defectos, ajustar prompts, mejorar pipelines, optimizar integraciones y resolver incidencias técnicas durante el desarrollo, despliegue y estabilización.
- Colaborar con arquitectura para asegurar que la solución respeta los patrones y golden paths corporativos.
- Colaborar con seguridad, datos, LLMOps/MLOps y observabilidad para asegurar que la solución es segura, trazable, operable y mantenible.
- Participar en el despliegue entre entornos de laboratorio, pruebas y producción siguiendo los mecanismos corporativos de CI/CD, LLMOps/MLOps y gestión de cambios.
- Documentar la solución desarrollada, incluyendo arquitectura técnica, componentes, integraciones, configuración, operación, dependencias y procedimientos de mantenimiento.
- Realizar transferencia técnica a equipos de soporte, equipos promotores y responsables de mantenimiento.
- Participar en la estabilización post-arranque, análisis de incidencias y propuestas de mejora evolutiva.

Nota: Este rol tendrá una orientación eminentemente práctica y de desarrollo, participando directamente en la construcción del software, integraciones, configuraciones y artefactos técnicos necesarios para materializar los casos de uso asignados dentro de la Línea 8. Su activación no supondrá convertir el contrato en un mecanismo general de ejecución de grandes proyectos de desarrollo, sino dotar al servicio de capacidad técnica suficiente para abordar casos acotados, aceleradores, prototipos avanzados o necesidades internas que requieran una implementación directa por parte del adjudicatario.

6.2.9 Ingeniero de Datos

Responsable de datos y conocimiento utilizables por IA.

Responsabilidades:

- Integrar fuentes.
- Diseñar pipelines.
- Preparar bases de conocimiento.
- Gestionar metadatos, calidad y linaje.
- Apoyar vectorización y recuperación semántica.
- Coordinarse con Gobierno del Dato y otros equipos de referencia en ese ámbito.

6.2.10 Consultor Gobernanza IA, Riesgo y Cumplimiento

Responsable de asegurar que las iniciativas de IA cumplen el modelo de gobierno y normativa aplicable.

Responsabilidades:

- Aplicar criterios del modelo de gobierno.
- Revisar riesgos y controles.
- Coordinar cumplimiento RGPD, AI Act, ENS y políticas internas.
- Mantener evidencias.
- Apoyar revisiones de Responsable AI.
- Participar en aprobaciones y paso a producción.

6.2.11 Ingeniero de Seguridad IA

Responsable de seguridad específica de soluciones de IA.

Responsabilidades:

- Evaluar riesgos GenAI.
- Revisiones prompt injection, fuga de datos, jailbreak y control de abuso de herramientas.
- Definir controles de acceso.
- Validar logging y trazabilidad.
- Coordinarse con Seguridad EJIE y Seguridad Corporativa EJIE.
- Revisar integraciones y proveedores.

6.2.12 Consultor de Observabilidad IA

Responsable del marco de control de la salud operativa de la plataforma y servicios IA.

Responsabilidades:

- Definir métricas de observabilidad.
- Integrar monitorización.
- Gestionar incidencias, problemas y cambios.
- Definir niveles de soporte.
- Medir disponibilidad, rendimiento y calidad.
- Elaborar informes de salud.
- Coordinación con equipos de observabilidad ya existentes

6.2.13 Ingeniero de Soporte y Operación de la Plataforma IA

Perfil responsable de prestar soporte especializado de segundo y tercer nivel sobre la plataforma corporativa de IA, actuando sobre los componentes propios del ámbito IA y coordinándose con los equipos de otras plataformas de ingeniería o responsables del soporte de las capas inferiores de infraestructura, cloud, networking, seguridad, identidad, observabilidad corporativa, operación, etc.

Responsabilidades:

- Atender, analizar y diagnosticar incidencias productivas relacionadas con la plataforma IA.
- Resolver incidencias sobre el IA Gateway, modelos, servicios de inferencia, RAG, agentes, pipelines, prompts, embeddings, guardrails, trazabilidad, observabilidad específica y configuraciones de plataforma.
- Gestionar problemas recurrentes, errores conocidos, degradaciones, deuda operativa y acciones correctivas.
- Ejecutar procedimientos de recuperación, mitigación, rollback y estabilización.
- Coordinar escalados con equipos de plataforma de infraestructura on-premises, plataforma cloud pública/hiperescalares, networking, kubernetes, SDLC, identidad, seguridad, observabilidad corporativa etc.
- Aportar logs, métricas, trazas, evidencias y análisis técnico para facilitar la resolución de incidencias cuya causa pueda estar en capas subyacentes.
- Mantener runbooks, playbooks, procedimientos de diagnóstico, recuperación y escalado.
- Participar en cambios productivos, releases, ventanas de mantenimiento y despliegues controlados.
- Realizar análisis de causa raíz en incidencias relevantes.
- Proponer mejoras operativas para reducir recurrencia, mejorar resiliencia, automatizar diagnósticos y aumentar la observabilidad.
- Mantener actualizado el catálogo de servicios IA en producción y sus responsables operativos.
- Colaborar con el Consultor de Observabilidad IA en cuadros de mando, alertas, métricas y seguimiento de salud del servicio.
- Colaborar con el Ingeniero LLMOps/MLOps en versionado, despliegue, rollback y estabilización de modelos, prompts, pipelines y configuraciones.
- Colaborar con el Ingeniero de Seguridad IA cuando la incidencia tenga implicaciones de fuga de datos, prompt injection, abuso de herramientas, accesos indebidos o eventos de seguridad.

6.2.14 Consultor de FinOps IA

Responsable de coste, eficiencia y valor.

Responsabilidades:

- Medición de consumo por servicio, modelo y caso de uso.
- Optimización de costes.
- Medición reutilización.
- Cálculo KPIs e indicadores de valor en general.
- Detectar ineficiencias.
- Elaborar informes FinOps IA.

6.2.15 Consultor de Negocio IA

Responsable de conectar necesidades funcionales con capacidades técnicas y de plataforma.

Responsabilidades:

- Apoyar recepción de necesidades/iniciativas.
- Traducir necesidades en casos de uso.
- Definir hipótesis de valor.
- Coordinar con áreas promotoras.
- Apoyar priorización.

- Identificar dependencias y riesgos.

6.2.16 Consultor de Adopción y Formación IA

Responsabilidades:

- Diseñar materiales formativos.
- Crear guías de uso.
- Apoyar usuarios de referencia y “AI Champions”.
- Dinamizar comunidad de práctica.
- Medir adopción y satisfacción.
- Identificar barreras que limiten la adopción y éxito.

6.3 Flexibilidad y Perfiles Complementarios

La empresa adjudicataria podrá proponer perfiles complementarios, siempre que justifique su aportación al servicio, tales como:

- Especialista en UX conversacional.
- Especialista en evaluación de modelos.
- Especialista en NLP.
- Especialista en agentes.
- Especialista en integración.
- Especialista en gobierno del dato.
- Especialista en contratación tecnológica o proveedores IA.
- Especialista en formación avanzada.

7 Metodología de Trabajo y Entorno Tecnológico

7.1 Principios Metodológicos

La metodología de trabajo deberá combinar gestión de servicio, gestión ágil de evolución de plataforma, gobierno de casos de uso, prácticas DevOps/MLOps/LLMOps y mejora continua.

El servicio deberá trabajar bajo los siguientes principios:

1. **Orientación a producto** La plataforma de IA deberá gestionarse como un producto interno, con roadmap, backlog, catálogo, usuarios consumidores, métricas y mejora continua.
2. **Orientación a servicio** La operación de la plataforma deberá integrarse con procesos de soporte, incidencias, cambios, problemas, niveles de servicio y reporting.
3. **Agilidad controlada** La evolución se organizará mediante backlog priorizado, iteraciones, revisiones periódicas y entregables incrementales.
4. **Gobierno desde el diseño** Seguridad, privacidad, datos, observabilidad y coste deberán incorporarse desde la fase de cualificación y diseño.
5. **Reutilización prioritaria** Todo nuevo caso de uso deberá evaluarse primero contra capacidades existentes de plataforma.
6. **Automatización y trazabilidad** El objetivo es la automatización de los despliegues, con versionado, auditoría, logging y control de cambios.
7. **Mejora continua basada en métricas** La evolución de la plataforma deberá basarse en indicadores de uso, adopción, excepciones, salud, coste y valor.

7.2 Ciclo de Trabajo para evolución y ciclo de vida de la plataforma

El adjudicatario deberá proponer y ejecutar una vez validado por EJIE un ciclo de trabajo basado en:

- Recepción y priorización de necesidades.
- Análisis de impacto.
- Diseño funcional y técnico.
- Validación con EJIE.
- Implementación.
- Pruebas técnicas, seguridad y operación.
- Documentación.
- Despliegue.
- Monitorización.
- Mejora continua.

7.3 Ciclo de Trabajo para casos de uso

El adjudicatario deberá proponer y ejecutar una vez validado por EJIE un ciclo de trabajo que incorpore estas etapas:

- Identificación y registro.
- Cualificación inicial.

- Clasificación por arquetipo/patrón.
- Evaluación de datos, riesgo, valor y viabilidad.
- Encaje con plataforma.
- Diseño de solución.
- Implicación en tareas específicas o prerequisites que precisen desarrollo y el el promotor no pueda abordar
- Prototipo o piloto.
- Evaluación.
- Revisión de seguridad, privacidad y cumplimiento.
- Decisión de paso a producción.
- Despliegue mediante golden path.
- Operación, medición y mejora.

7.4 Entorno Tecnológico

El alcance tecnológico del CoE de IA es híbrido y multicloud. Se resume a continuación:

- **Nubes públicas:** el CoE de IA se coordinará con el CoE específico de Cloud de EJIE, que disponibilizará una plataforma de cloud/landing zones sobre la que se apoyará la plataforma de IA. En general el cloud de referencia será AWS (hiperescalar principal en EJIE) para el despliegue de la mayoría de soluciones. Adicionalmente, se contemplará Google Cloud Platform (GCP), (p.ej., para casos de uso específicos) y según la estrategia de sourcing se podría también incorporar OCI, u alguno de los otros hiperescalares relevantes en el mercado. Actualmente como herramienta de FinOps se utiliza cloudability.
- **Infraestructura On-premise:** El modelo es híbrido, por lo que algunos casos de uso se cubrirán 100% en cloud, de manera híbrida o 100% on-premises que deberán estar integrados. La plataforma de IA se apoyará para eso en la plataforma de infraestructura on-premises que disponibiliza la infraestructura tradicional del CPD de EJIE y el cloud privado del servicio IaaS de Batera construido sobre vCF).
- **Herramientas de DevOps e IaC:** La pila de herramientas corporativa incluye GitLab (como sistema de control de versiones y orquestador de pipelines CI/CD) y Terraform para la definición de Infrastructure as Code que está estandarizado en un marco de desarrollo para IaC. Se seguirá por tanto un modelo GitOps (gestión de infra/config mediante repositorio). Además, se usarán tecnologías de contenedores (Kubernetes) y su ecosistema para aplicaciones en la nube.
- **Plataforma de desarrollo interno (IDP):** se aprovechará el portal interno del desarrollador de EJIE (IDP) basado de Backstage llamada BatArin, para publicar y exponer los patrones/building blocks que disponibiliza la plataforma de IA.
- **Monitorización y gestión:** Todos los entornos cloud y recursos serán integrados en la plataforma de monitorización y gestión de EJIE, permitiendo a EJIE visualizar alertas, métricas y controlar la calidad del servicio. Por ejemplo, los logs y métricas de las aplicaciones en AWS se remitirán a las herramientas corporativas (p.ej., Grafana, ELK stack), y las cuentas/subscripciones cloud se integrarán con las herramientas de gestión de inventario y service management de EJIE (ServiceNow).
- **Seguridad y compliance:** se usarán los servicios nativos de seguridad de AWS/GCP (ej. AWS Config, CloudTrail, Security Hub, GCP Cloud Security Command Center), u otros hiperescalares implicados según el caso, integrados con la estrategia de ciberseguridad corporativa (p.ej., volcando logs al SIEM, cumpliendo las directrices de GureSek y del

Delegado de Protección de Datos). Además, se considerarán los lineamientos de ENS, la NIS2 y los marcos de compliance relevantes a la nube pública.

En lo específico al ámbito de la Inteligencia Artificial, la empresa adjudicataria en el momento del arranque del servicio deberá asumir la versión de plataforma IA existente y proponer un plan de transición y evolución que minimice riesgos, asegure continuidad y preserve el conocimiento.

Esta plataforma existente se basa en:

- Servicios cloud principalmente de AWS.
- Infraestructura on-premises para inferencia (Servicio IaaS sobre vCF y Kubernetes)
- Infraestructura on-premises para HPC organizada con Slurm y desarrollo de modelos.
- Hub de Modelos/Model as a Service basado en vLLM y LiteLLM
- Mecanismos de monitorización de uso y consumo, chargeback y showback para la infraestructura HPC y de inferencia on-premises.
- Repositorios de código y configuración.
- Repositorio plantillas para el despliegue automatizado de patrones/*blueprints* para soluciones IA en Cloud.
- Repositorio imágenes on-premises para el despliegue automatizado de patrones/*blueprints* para soluciones IA.
- Integración con una futura solución de Gobierno de IA que EJIE implantará en los próximos meses, (que se contratará en otro expediente fuera del ámbito de este contrato) y en la que se espera que el adjudicatario se implique.

7.5 Indicadores Mínimos del Servicio

El adjudicatario deberá proponer, mantener y reportar indicadores como mínimo en las siguientes dimensiones:

Dimensión	Indicadores mínimos
Demanda	Nº de casos registrados, cualificados, priorizados y en ejecución.
Plataforma	Nº de servicios disponibles, uso por servicio, adopción y disponibilidad.
Reutilización	Nº de casos por patrón, servicios reutilizados y componentes comunes.
Velocidad	Tiempo desde recepción hasta entorno usable o primer resultado técnico.
Excepciones	Nº de excepciones, recurrencia, resolución e incorporación al catálogo.
Operación	Incidencias, problemas, cambios, disponibilidad, latencia y errores.
Seguridad	Riesgos identificados, controles aplicados, evidencias y excepciones.
Datos	Fuentes integradas, datasets validados, calidad, permisos y linaje.

FinOps	Coste por servicio, modelo, patrón, caso de uso y tendencia de consumo.
Valor	Ahorro, productividad, satisfacción, reducción de tiempos y beneficios medidos.
Adopción	Usuarios formados, iniciativas acompañadas, satisfacción y uso real.

7.6 Reuniones y Mecanismos de Seguimiento

La empresa adjudicataria deberá contemplar, al menos:

- Reunión operativa semanal.
- Comité de seguimiento mensual.
- Comité técnico de arquitectura/plataforma.
- Comité de demanda y portfolio.
- Revisión periódica de seguridad, riesgos y cumplimiento.
- Revisión periódica de FinOps y valor.
- Informe mensual de actividad.
- Informe ejecutivo trimestral.
- Repositorio actualizado de documentación y evidencias.



Eusko Jaurlaritzaren Informatika Elkartea
Sociedad Informática del Gobierno Vasco