# 

# Sistema Corporativo Documental con Capacidades de IA

**Código de expediente:** EXP-TD-OPER-2025-12  
**Organismo/Empresa promotora:** TeFinancia S.A. – Dirección de Medios   
**Régimen de contratación:** Proyecto llave en Mano  
**Duración estimada:** 30 meses (con posibilidad posterior de ampliación para las fases de diseño, implementación y   
**Presupuesto base de licitación (PBL):** 2.000.000  
**Tratamiento de datos personales:** Sí (conforme a normativa aplicable)



## 1. TeFinancia

### 1.1 Nuestro proyecto empresarial

TeFinancia ofrece productos de financiación al consumo, como préstamos personales, tarjetas de crédito, y financiación para coches y reformas, actuando también como socio de comercios y concesionarios para facilitar la compra a sus clientes.

Nuestra misión se centra en ser un socio estratégico que facilita la gestión del presupuesto diario de las personas a través de soluciones de financiación y seguros, promoviendo un crédito responsable y la educación financiera.  Queremos:

* **Ser un socio financiero para los particulares:**

Ayudar a las personas a gestionar su presupuesto y a realizar sus proyectos mediante productos de crédito y financiación.

* **Ser un aliado para el comercio:**

Colaborar con comercios y marcas para que puedan ofrecer financiación a sus clientes, facilitando la compra.

* **Promover un crédito responsable:**

TeFinancia está comprometida con el crédito responsable, ofreciendo productos adaptados a las necesidades de los clientes y fomentando la educación financiera.

* **Integrar la sostenibilidad:**

Generar un impacto positivo en la sociedad y el medio ambiente integrando la sostenibilidad en sus operaciones.

### 1.2 Servicios de TeFinancia

* **Préstamos y financiación:**

Ofrecemos una amplia gama de productos, incluyendo:

* + Préstamos personales.
  + Financiación para coches (nuevos y usados) y motocicletas.
  + Financiación para reformas del hogar y salud dental.
* **Tarjetas de crédito:**

Productos financieros para gestionar las finanzas diarias de los particulares.

* **Seguros:**

Disponemos de seguros para complementar sus servicios de financiación.

* **Soluciones para comercios:**

TeFinancia trabaja con grandes cadenas de distribución, tiendas físicas, concesionarios de automóviles y comercios electrónicos para ofrecer soluciones de financiación para sus clientes.

### 1.3 Contexto

Enmarcado dentro del ambicioso Plan Estratégico de Transformación Digital y Modernización llamado **FinancIA 2030**, TeFinancia busca reinventar su sistema documental corporativo e implementar un nuevo sistema documental disruptivo que simplifique y automatice las tareas documentales embebidas en los procesos con clientes, tanto personas físicas como figuras jurídicas.

Actualmente la documentación aportada por clientes, la contractual y la generada por la entidad, está distribuida en diferentes repositorios documentales, sin una visión única, y en la que no es posible tener una visión única e integrada de los activos documentales correspondientes a un cliente de una manera ágil y rápida. Esta situación no solo conlleva una pérdida de confianza por parte de los clientes sino que presenta claros riesgos de cara al cumplimiento regulatorio.

Por otro lado, los procesos actuales adolecen de fases manuales, que suponen un coste muy alto para la organización y un alto riesgo de error humano, lo que nos lleva a una baja eficiencia que impacta gravemente en la competitividad de la empresa.

Con esta RFP se busca poner fin a esta situación con un sistema moderno e innovador, buscando traer al presente como será la gestión documental en el futuro.

## 2. Objeto y alcance del contrato

### 1.1 Objeto

Contratación de una **solución integral de gestión documental** basada en capacidades cloud y de la Inteligencia artificia (IA) que permita:

* **Capturar cualquier documento** que se requiera en los procesos con clientes, productos y servicios, tanto documentación aportada, como generada internamente o por terceros que deba guardarse en los sistemas de TeFiancia, bien por requerimientos comerciales, bien por requerimientos regulatorios. Se considerará como documento cualquier activo generado a partir de la relación con los clientes: textos comerciales, jurídicos, recibos, mensajes SMS o audios de dispositivos móviles, calls y audios de llamadas, transcripciones, videos, fotos, etc.
* **Entendimiento** automatizado del contenido del documento. Indexado del documento, almacenaje de datos y metadatos asociados, etiquetado y clasificación siguiendo taxonomías predefinidas.
* **Almacenamiento inteligente de los documentos**, con procesos de ciclo de vida documental automatizados, respetando el marco regulatorio y obedeciendo también a criterios comerciales.
* **Búsquedas avanzadas sobre los documentos**, no sólo a través de consultas paramétricas sino también a través del uso del lenguaje natural. Es decir, debe permitirse búsqueda léxica y semántica. Estas búsquedas, perimetradas y con los permisos adecuados, estarán habilitadas tanto a gestores y personal interno autorizado como a los clientes y proveedores, según aplique. El resultado tiene que ofrecerse en también en lenguaje natural con diferentes niveles niveles de resumen y estructurando adecuadamente referencias y trazabilidad hacia las fuentes de información

* **Chatbots y agentes inteligentes** para ingesta o consulta documental
* **Evaluar y gestionar la vigencia** y validezde los documentos lanzando tareas y alertas a gestores y clientes, según aplique.
* **Evaluación del riesgo** legal, operativo, de cumplimiento y reputacional de cada contrato, póliza, documento legal, o documento aportado por cliente.
* Relación con Scoring de clientes y operaciones en base no sólo a la información estructurado de los datos de cliente, sino también a información no estructurada obtenido de cualquier documento o fichero del cliente.
* Analítica clásica (descriptiva) y avanzada (predictiva) sobre el contenido y el uso del gestor documental.
* Para **contratos de terceros**, ejecutar **validaciones y comprobaciones automáticas** de viabilidad (p. ej., detección de proveedores con restricciones/sanciones, conflictos de interés, incompatibilidades, condiciones económicas abusivas o no conformes).
* Implantación de un gobierno Gobierno Documental y de la IA.
* Servicio inteligente de Discovery e Ideación para identificar casos de uso con valor comercial a partir de la documentación y el uso de técnicas, modelos y/o herramientas de inteligencia artificial.
* Creación de **APIs** documentales e integración con **mapa de procesos corporativo** así como con las aplicaciones que hacen uso de tareas documentales, sea cual sea el canal (terminal financiero, aplicaciones, apps, portales web, etc). Un ejemplo sería diseñar un nuevo visualizador corporativo de documentos.
* Definición del Plan para Migración Documental (desde los sistemas antiguos)
* Definición de Resilencia y Planes de Contingencia

### 1.2 Alcance funcional (visión de alto nivel)

La solución deberá cubrir, como mínimo, los siguientes dominios:

1. **Captura, Ingesta, entendimiento y clasificación documental**: OCR avanzado, detección de idioma, firma digital, control de versiones y deduplicación.
2. **Validación y compliance**: reglas declarativas, detección de gaps vs. plantillas, chequeos de terceras partes evaluación de compatibilidad con políticas internas.
3. **Evaluación de riesgo**: scoring multidimensional configurable (legal, financiero, operativo, ESG, privacidad, ciber), explicabilidad y evidencias.
4. **Búsqueda y analítica**: buscador léxico/semántico, RAG sobre repositorio corporativo, análisis comparativo de redlines y cláusulas alternativas.
5. **Asistencia inteligente**: copiloto para redacción/negociación, sugerencias de enmiendas alineadas a estándares internos, generación de reportes ejecutivos.
6. **Gobierno, seguridad y auditoría**: gestión de identidades y zonas de datos, Trazabilidad (lineage), telemetría, observabilidad, controles ENS/ISO, registro de decisiones.
7. **La gestión documental** debe permitir trabajar con los documentos en multiidioma: castellano, inglés, catalán, euskera, gallego, francés y alemán.

### 1.3 Exclusiones

1. No se solicitará la externalización de decisiones jurídicas finales (las decisiones corresponderán al área legal).
2. No se requiere la provisión de BPMS corporativo (sí la integración).
3. El proyecto no subcontratará partes a terceros proveedores. Los participantes en el proyecto deben ser empleados directos del licitante.

### 1.4 Entornos y despliegue

La solución deberá operar en **entornos segregados** (DEV, PRE, PROD) y soportar despliegueen**, cloud privado o cloud público** según política de la Entidad.

TeFinancia tiene sus sistemas, en un 80% en infraestructura On-Premises y un 20% de las aplicaciones sobre infraestructura Cloud en AWS y en Microsoft Azure. La solución debe contemplar esta situación híbrida e integrarse con todos los sistemas de la casa que tengan tareas documentales.

## 2. Requisitos funcionales

### 2.1 Ingesta y preparación

* **Fuentes**: DMS/ECM (p. ej., SharePoint/Alfresco), correo, SFTP, API REST, carpetas vigiladas.
* **Formatos**: DOCX, PDF (nativo/escaneado), ODT, TXT, imágenes (PNG/JPG/TIFF), firma XAdES/PAdES. También formato de impresión AFP, audios (mp3) y videos (mp4) y sus correspondientes transcripciones. Debe tenerse un conversor de formatos que permita desde otros formatos no soportados convertir a los formatos permitidos.
* **OCR**: multilenguaje (ES/EN/FR/PT al menos), detección de tablas y sellos; precisión OCR ≥ 98% en documentos estándar.
* **Control de versiones** y **deduplicación**
* **Estados de documentos** (pendiente de firma, completado, etc).

### 2.2 Etiquetado, extracción y taxonomías

* **Datos de contenido, metadatos asociados e indicadores de compliance en base a taxonomías**.
* **En el caso de documentos contractuales, detección de cláusulas**: clasificación por plantillas estándar (corporativas) y variantes; identify de cláusulas ausentes/atípicas.
* **Metadatos obligatorios**, metadatos de contenido y metadatos auxiliares.
* **Taxonomías**: adopción/alineación a **ontologías corporativa**.

### 2.3 Validaciones automáticas (contratos propios y de terceros)

* **Reglas de conformidad** parametrizables (motor de reglas) frente a **estándares internos**.
* **Checks de proveedores/terceros**: integración con listas públicas (p. ej., sanciones, inhabilitaciones), registros mercantiles, scoring de riesgo de terceros, cumplimiento ESG, histórico de incidencias; soporte de **listas de bloqueo/observación** internas.
* **Compatibilidad económica**: validación de términos de pago, revisión de cláusulas de precio/variación, penalizaciones y descuentos conforme a umbrales corporativos.
* **Alertas** de incompatibilidades y **sugerencias de redacción** (IA generativa) conforme a guías internas.

### 2.4 Evaluación y explicación del riesgo

* **Modelo de scoring** configurable por dimensiones y pesos.
* **Explicabilidad**: evidencias trazables (párrafos, cláusulas, registros externos consultados) y **racional explícito** de la puntuación.
* **Histórico**: evolución de riesgo por proveedor/contrato, tendencias y comparativas inter‑contratos.

### 2.5 Búsqueda, recuperación y analítica

* **Búsqueda híbrida**: léxica (BM25) + **semántica/vectorial**; soporte de **filtros facetados** por metadatos.
* Respuestas con **citas a evidencias** (páginas/secciones) y **controles anti‑alucinación**.
* **Comparativa**: diff jurídico entre versiones; alineación contra cláusulas de referencia.
* **Cuadros de mando**: KPIs de pipeline, riesgos, tiempos de ciclo, productividad (HITL), calidad de extracción, SLA y compliance.
* **Técnicas de Few-shot** para el aprendizaje continuo e incremento del accuracy en las búsquedas.

## 3. Requisitos no funcionales y técnicos

### 3.1 Rendimiento y disponibilidad

* **Disponibilidad** objetivo PROD ≥ 99,9%; **MTTR** ≤ 4 h; **RPO** ≤ 1 h; **RTO** ≤ 4 h.
* **Latencia** consultas de búsqueda ≤ 2 s p95 (corpus ≤ 5 M de páginas indexadas).
* **Throughput**: ingestión mínima 10.000 páginas/hora por nodo (escalable horizontalmente).

### 3.2 Calidad y métricas de IA

* **Extracción/Clasificación**: F1 ≥ 0,85 sobre golden set acordado (por entidad/cláusula prioritaria).
* **Riesgo**: correlación con dictámenes expertos ≥ 0,7 en pilotos; drift monitoring y recalibración.
* Grounded answer rate ≥ 95% (con evidencia citada).
* **Trazabilidad**: versión de modelo/datos, features, prompts, parámetros y resultados.
* **Explicabilidad e Responsible AI**

### 3.3 Seguridad y cumplimiento

* **Identidades y acceso**: SSO (OIDC/SAML), MFA, RBAC/ABAC, segregación por dominio/área legal.
* **Cifrado**: en tránsito (TLS1.2+) y en reposo (AES‑256 o equivalente).
* **Protección de datos**: minimización, retención, **pseudonimización/anonimización**, **DPIA** cuando aplique, gestión de derechos ARSOPL.
* **Auditoría**: logs inmutables, retención ≥ 2 años, exportación a SIEM, **detección de prompt‑injection** y salida sensible.
* **Marcos**: alineamiento con **RGPD/Ley Orgánica**, **ENS**, **ISO 27001/27701**, **ISO 42001 (IA)** o equivalente. Alineamiento con RDA.

### 3.4 Arquitectura y tecnología mínimas

* **Indexación y búsqueda**: motor léxico (p. ej., OpenSearch/Solr) + **BD vectorial** (p. ej., pgvector/OpenSearch k‑NN/FAISS) con **consistencia de metadatos**.
* **Repositorio de conocimiento**: almacenamiento documental (objeto/ficheros) + metadatos en DB relacional/noSQL; **catalogación** y **data lineage**.
* **Capa de IA**:
  + **IA generativa**: LLM gobernado (zona segura), **RAG** con control de fuentes; opciones on‑prem, cloud soberano o proveedor cualificado; **guardrails**.
  + **ML clásico**: pipelines de extracción/clasificación (NER, segmentadores, clasificadores de cláusulas), **MLOps** (registro de modelos, métricas, despliegue canario, rollback).
* **Integración**: APIs REST/GraphQL, colas/eventos (Kafka/RabbitMQ), ETL/ELT; conectores OOTB a DMS/ERP/CRM.
* **Observabilidad**: métricas (Prometheus), trazas (OpenTelemetry), logs; tableros operativos.
* **IaC y CI/CD**: Terraform/Ansible; GitOps; pruebas automatizadas (incl. **pruebas de LLM**).

### 3.5 Portabilidad y reversibilidad

* **Exportación** íntegra de datos, embeddings, modelos, artefactos y metadatos en formatos abiertos.
* **Derechos de uso**: código/artefactos entregados según régimen acordado; documentación técnica y funcional completa.

## 4. Requisitos de arquitectura (detallados)

### 4.1 Principios

* **Segregación por zonas**: datos brutos, procesados, features, entregables; **datos sensibles en zona confinada**.
* **Zero‑trust** y **least privilege**.
* **Defensa en profundidad**: WAF, EDR, DLP, escaneo SCA/SAST/DAST, revisión de prompts y plantillas.

### 4.2 Componentes mínimos

* **Capa de ingesta** con colas; **servicio de OCR** escalable; **normalizador** de formatos.
* **Servicio de parsing**
* **Capa de extracción y clasificación** (NER/cláusulas) con modelos entrenables.
* **Almacén documental** con **control de versiones** y **retención**.
* **Índices léxicos y vectoriales** sincronizados; **orquestador RAG** con **citación de fuentes** obligatoria.
* **Motor de reglas** (consola de configuración por Legal/Compliance).
* **Módulo de scoring de riesgo** con **explicabilidad** y **versionado**.
* **Capa de APIs** y **portal** para usuarios (buscador, revisión, dashboard, copiloto).
* **MLOps/LLMOps**: registro de datasets, modelos, prompts, experimentos; tests de seguridad de IA.

### 4.3 Entornos

* **DEV/PRE/PROD** separadas; **paso a producción** con criterios de aceptación definidos (calidad, seguridad, rendimiento) y **pruebas automatizadas**.

## 5. Seguridad, gobierno de la información y cumplimiento

### 5.1 Seguridad

* **ENS (cuando aplique)**: niveles y medidas organizativas/técnicas; **anexo de cumplimiento**.
* **ISO 27001/27701**: controles relevantes; **gestión de vulnerabilidades** y parches; **copias de seguridad**.
* **Seguridad de IA**: evaluación de amenazas (prompt injection, data exfiltration, model poisoning), **guardrails**, listas de bloqueo, sanitización de prompts/outputs.

### 5.2 Protección de datos

* **RGPD y LOPDGDD**: bases legitimadoras, contratos de encargo, **DPIA** en su caso, transferencias internacionales, minimización y retención.
* **Clasificación de datos** y **políticas de acceso**; redacción automática assistida de **cláusulas de privacidad** con revisión humana.

### 5.3 Gobierno del dato

* **Catálogo de datos**, **calidad**, **linaje**, **diccionario de metadatos**, **comité de gobierno** y **roles** (data owner, steward, custodio, MLOps lead).
* **Política de datasets de entrenamiento**: procedencia, sesgo, licencias, derechos, registros de consentimiento.

## 6. Metodología de trabajo requerida

El adjudicatario deberá describir una **metodología híbrida** que incluya:

* Descubrimiento y priorización de casos
* Diseño centrado en usuario (UX de legal/compliance)
* Definición de **golden sets** y KPIs de calidad
* Pilotos controlados (sandbox) y criterios Go/No‑Go
* Gestión de cambios
* Aseguramiento de calidad (QA)
* Pruebas automatizadas (incl. **pruebas de LLM**)
* Gestión de riesgos y plan de mitigación;
* Plan de formación y **adopción** (playbooks, guías de uso responsable de IA).

## 7. Plan de implantación y perfiles

### 8.1 Plan de implantación (a presentar por el licitador)

El licitador deberá aportar un **plan detallado** con fases, hitos, entregables, recursos, dependencias, plan de datos y seguridad, estrategia de migración, y plan de reversibilidad.

### 8.2 Perfiles mínimos (a detallar por el licitador)

Se deberá proponer **equipo multidisciplinar** indicando dedicación estimada por fase y **actividades por rol**, incluyendo al menos: Dirección de Proyecto, Arquitecto/a de IA, Especialista LegalTech, Científico/a de Datos, Ingeniero/a MLOps/LLMOps, Ingeniero/a de Datos, Ingeniero/a Backend/Integraciones, Especialista en Búsqueda/Información, Especialista en Seguridad/ENS, UX, y Analista de Compliance.

## 8. Hitos y entregables (mínimos)

* **D1**: Especificación funcional y de taxonomías (incl. mapeo a estándares internos).
* **D2**: Diseño de arquitectura lógico/físico con **matriz de seguridad** y **ENS/ISO**.
* **D3**: Plan de datos, gobierno y DPIA (si procede).
* **D4**: Prototipo funcional (MVP) con corpus acotado y KPIs de calidad iniciales.
* **D5**: Plataforma en PRE con pipelines, buscador y RAG; evidencias de pruebas.
* **D6**: Paso a PROD con criterios de aceptación cumplidos; manuales y formación.
* **D7**: Plan de operación, MLOps/LLMOps y **cuadro de mando** de servicio.
* **D8**: Plan de reversibilidad/portabilidad y entrega de artefactos.

## 9. Mantenimiento y niveles de servicio (SLA)

* **Soporte**: horario laboral ampliable a 24x7 para incidencias críticas; **tiempos de respuesta/recuperación** por severidad (S1–S4).
* **Mantenimiento correctivo/evolutivo**: ventanas planificadas; gestión de versiones de modelos.
* **SLA sugeridos** (a concretar en oferta): disponibilidad, tiempos de procesamiento, niveles de precisión mínimos, tiempos de resolución de incidencias, tasa de grounded answers, tasa de falsos positivos/negativos en checks críticos.

## 10. Criterios de evaluación de ofertas

### 10.1 Ponderación

* **Criterios evaluables mediante juicio de valor** (máx. **60 puntos**):
  + **Enfoque técnico y calidad de la solución** (30): cobertura funcional, arquitectura, seguridad, RAG/LLM guardrails, MLOps/LLMOps, portabilidad.
  + **Metodología y plan de implantación** (15): gestión de riesgos, calidad, datos, adopción, formación.
  + **Experiencia y equipo** (15): casos similares en legal/compliance, perfiles propuestos y organización.
* **Criterios evaluables automáticamente** (máx. **40 puntos**):
  + **Precio** (25): fórmula de adjudicación indicada en Pliegos Administrativos.
  + **SLA y garantías** (10): disponibilidad, tiempos de respuesta, métricas de IA comprometidas.
  + **Mejoras/innovación verificables** (5): p. ej., explicabilidad avanzada, evaluadores automáticos de LLM, anonimización diferencial, etc.

La aplicación del **umbral del 50%** sobre los criterios de juicio de valor podrá ser exigida para continuar la evaluación.

### 10.2 Método de valoración (resumen)

* Para **precio**: fórmula de proporcionalidad inversa.
* Para **SLA**: escalas por tramos; penalizaciones por incumplimiento.
* Para **calidad de IA**: compromisos mínimos y bandas de mejora con evidencias en PRE.

## 11. Condiciones de ejecución

* **Adscripción de medios**: equipo clave mínimo; régimen de sustituciones; dedicación.
* **Propiedad intelectual**: condiciones de uso, licencias de terceros, derechos de explotación.
* **Confidencialidad** y **ética de IA**: uso responsable; prohibición de entrenar modelos externos con datos del cliente sin autorización.
* **Penalizaciones**: por incumplimiento de SLA o entregables críticos.
* **Facturación**: por hitos (a detallar en el PCAP); aceptación formal por la Dirección del Proyecto.

## 13. Documentación que el **proveedor** deberá presentar en su oferta

**Importante:** El licitador **no debe resolver** los puntos de diseño por la Entidad, sino **proponer** su enfoque cumpliendo los requisitos.

1. **Detalle de Arquitectura**
   * Arquitectura lógica y física, patrones y componentes; despliegue (on‑prem/nube); zonas y flujos de datos; **diagrama de seguridad** (ENS/ISO).  
     Modelo de indexación y búsqueda (léxica + vectorial), orquestación RAG, motor de reglas, scoring de riesgo y guardrails.
   * Estrategia de escalado, alta disponibilidad, continuidad (RPO/RTO), observabilidad y **hardening**.
2. **Detalle de la Solución Técnica**
   * Pipelines de ingesta, OCR, parsing, NER, clasificación de cláusulas, validadores y checks externos.
   * Esquemas de metadatos taxonómicos; formatos de intercambio; APIs; conectores; UI/UX.
   * MLOps/LLMOps: repositorios, datasets, versionado, evaluaciones, despliegues, rollback.
3. **Seguridad, Gobierno de la Información y Cumplimiento**
   * Controles de seguridad, cifrado, identidades y accesos; gestión de secretos; segregación y minimización.
   * Cumplimiento **RGPD/LOPDGDD**, **ENS**, **ISO 27001/27701/42001** o equivalentes; **DPIA** (si aplica); tratamiento de datos en IA generativa; registro y auditoría.
   * Política de datasets de entrenamiento y uso de modelos fundacionales; medidas anti‑sesgo; explicabilidad.
4. **Metodología de Trabajo**
   * Plan de proyecto, governance, gestión de riesgos, QA y pruebas (incl. LLM); gestión de cambios; formación y adopción.
5. **Plan de Implantación Detallado**
   * Fases, hitos, entregables, criterios de aceptación; dependencias; gestión de cut‑over/migración; plan de reversibilidad.
6. **Perfiles y Organización del Equipo**
   * Roles, funciones y actividades de cada perfil; dedicaciones y plan de cobertura; cuadro RACI.
7. **Compromisos de Servicio (SLA) y Mantenimiento**
   * Cobertura, tiempos de respuesta/recuperación, ventanas de mantenimiento; evolución de modelos y revisiones periódicas.
8. **Propuesta Económica** (según PCAP)
   * Desglose por hitos/mantenimientos; licencias/infraestructuras; hipótesis de dimensionamiento.

## 14. Aceptación, pruebas y paso a producción

* **Criterios de aceptación** por entregable (funcionalidad, rendimiento, seguridad, calidad IA).
* **Pruebas**: unitarias, integración, rendimiento, seguridad (SAST/DAST), **pruebas de LLM** (robustez, groundedness, red‑teaming).
* **Plan de UAT** con usuarios legales/compliance; **manuales** y **formación** previas al Go‑Live.