



OBT INTELLIGENCE SUITE

Productos de IA para SharePoint



OBT IA CRAWLER

Clasificador IA para SharePoint

Hoja de Producto

Ficha comercial y ejecutiva

Versión 2.0 · Julio 2026

Documento confidencial — Ortega Technologies (OB Technologies)

Qué es OBT IA Crawler

OBT IA Crawler es el clasificador documental con IA de la OBT Intelligence Suite. Barre un origen de datos —un servidor de ficheros o un SharePoint— extrae el texto de cada documento, lo clasifica con un modelo de lenguaje y produce metadatos estructurados y un informe de inventario, exposición de datos sensibles y arquitectura de información propuesta. Es la capa de valor que convierte una migración a SharePoint Online en un repositorio gobernado, buscable y conforme.

Principio de diseño: en la Fase 1 el producto solo lee y propone. No modifica ni un solo documento ni metadato hasta que el cliente revisa el informe y da su aprobación (ADR-07, no destructivo).

El problema que resuelve

- **Repositorios heredados sin gobierno:** millones de ficheros en carpetas o en SharePoint on-premise, sin metadatos, con duplicados, obsoletos y datos personales dispersos.
- **Migraciones que solo mueven el desorden:** trasladar a SharePoint Online sin clasificar reproduce el caos en la nube.
- **Riesgo de cumplimiento invisible:** no se sabe dónde hay PII, contratos o documentación regulada, ni quién tiene acceso indebido.
- **Coste de clasificación manual inviable:** etiquetar cientos de miles de documentos a mano no escala.

Cómo funciona

El motor recorre el origen de forma reanudable e idempotente, reparte el trabajo por una cola y clasifica cada documento con un embudo de triaje que descarta lo trivial de forma barata y escala a modelos mayores solo lo ambiguo o sensible. Por cada documento produce metadatos de negocio y de sensibilidad:

- **Tipo de documento** (contrato, factura, informe, procedimiento...), proyecto, cliente y fecha.
- **Área o departamento** e idioma del documento.
- **Sensibilidad** (público, interno, confidencial, restringido).
- **Presencia de PII** y su tipo (email, teléfono, DNI/NIF, IBAN, tarjeta), validada con checksums (Luhn, mod-97, letra del DNI).
- **Confianza** de la clasificación; lo dudoso se marca «en revisión», no se inventa.

Fases de actuación

Fase	Qué hace	Salida
1 · Lectura e informe	Barrido en solo lectura, inventario y clasificación propuesta	Informe (PDF/Excel/JSON) + estimación de coste
2 · Validación del	El cliente revisa el informe y la arquitectura	Aprobación y taxonomía

Fase	Qué hace	Salida
cliente	de metadatos	acordada
3 · Escritura y migración	Escritura de metadatos en columnas SharePoint durante o tras la migración	Repositorio clasificado y gobernado

Sobre esta base se construyen los servicios recurrentes: *assessment de datos confidenciales*, *suscripción de monitorización (DLP continuo)* y *packs verticales por sector (farma primero)*.

Arquitectura en una línea

Motor en Node/TypeScript desplegado como Azure Functions v4 (plan Flex Consumption) en la UE, con OCR self-host en Azure Container Apps, cola en Service Bus, secretos en Key Vault y clasificación por Mistral (La Plateforme, UE) por defecto. Capa LLM y capa OCR intercambiables por configuración (Mistral, Claude, OpenAI, Azure OpenAI; RapidOCR, Azure Document Intelligence o Mistral OCR).

Diferenciadores

- **No destructivo y con aprobación previa:** la Fase 1 nunca toca el origen; la escritura se activa solo con el permiso del cliente.
- **Soberanía de datos UE:** LLM y OCR en la Unión Europea; el OCR no almacena la imagen (la procesa y la descarta). Idóneo para farma y sector regulado (RGPD).
- **Coste bajo el control del margen:** OCR propio ~25× más barato que Azure Document Intelligence y embudo de triaje que descarta ~88% de forma barata.
- **Resiliencia probada:** reintentos con backoff y respeto de límites de tasa en toda la capa HTTP, cola con detección de duplicados y alerta de mensajes muertos.
- **Bilingüe y multi-idioma:** informes en ES/EN/FR/DE.

Planes

Precio fijo por tamaño: cada plan se factura como un precio inicial más un precio por TB analizado. Ejemplo: Professional, 5 TB = 3.000 € + 5 × 1.500 € = 10.500 €.

Plan	Incluye	Precio inicial	Precio por TB
Esencial	PoC / auditoría: barrido + clasificación en solo lectura, informes. El cliente aporta su clave de IA. Sin escritura de metadatos.	1.500 €	900 €/TB
Professional	Gobierno en producción: OBT gestiona el LLM (Mistral UE), OCR opcional, escritura de metadatos tras aprobación.	3.000 €	1.500 €/TB
Enterprise	Grandes volúmenes / regulado: todos los proveedores de IA, migración asistida, residencia UE garantizada, soporte prioritario.	6.000 €	2.200 €/TB

Enterprise > 25 TB: a medida. Servicios de proyecto opcionales: despliegue de OCR self-host (desde 1.200 €), migración asistida + mapeo (desde 2.500 €), arquitectura de destino (desde 900 €) y formación (desde 600 €). Precios sin IVA, julio 2026.

Distinción de licencias y cálculo

Las tres licencias se calculan con la misma fórmula —precio inicial + precio por TB— pero cada una habilita un alcance distinto: Esencial no escribe metadatos (solo informe), Professional escribe tras aprobación con LLM gestionado por OBT, y Enterprise añade todos los proveedores de IA, migración asistida y soporte prioritario. Precio total según el tamaño analizado:

Licencia	Precio inicial	€/TB	1 TB	5 TB	15 TB
Esencial	1.500 €	900 €	2.400 €	6.000 €	15.000 €
Professional	3.000 €	1.500 €	4.500 €	10.500 €	25.500 €
Enterprise	6.000 €	2.200 €	8.200 €	17.000 €	39.000 €

Fórmula: total = precio inicial + (TB × precio por TB). La calculadora de licencias (Excel adjunto) devuelve el precio de las tres licencias para cualquier volumen.

En síntesis

El informe inicial es la puerta de entrada: precio bajo, sin compromiso de migración, y con el número real de documentos en la mano se cotizan la clasificación y la migración. Coste y precio escalan juntos: nunca se cotiza a ciegas sobre volumen desconocido.

Ortega Technologies (OB Technologies) · soporte@obtechnologies.com ·
www.obtechnologies.com