DevAI Suite — White Paper de Seguridad

Versión 1.0 — Abril 2026 Clasificación: Pública · seguro para compartir con clientes potenciales, clientes y auditores.

Resumen ejecutivo

DevAI Suite es la primera plataforma Manufacturing-as-a-Service (MaaS) — un entorno unificado de gestión de programas APQP, excelencia de proveedores e inteligencia de planta para fabricantes de automoción y aeroespacial. La plataforma maneja **datos regulados de manufactura** (envíos PPAP, libros FMEA, cualificación de proveedores, telemetría de planta) para clientes que operan bajo **IATF 16949** y **AS9100 / AS9145**.

Este white paper documenta cómo se diseña y opera DevAI Suite para proteger esos datos. Está estructurado conforme al **NIST Cybersecurity Framework 2.0** (Govern, Identify, Protect, Detect, Respond, Recover) y mapea cada función a controles, tecnologías y prácticas operativas concretas. Cuando hay certificaciones pendientes lo decimos honestamente con un calendario defendible en lugar de exagerar.

La audiencia prevista son los equipos de revisión de seguridad empresarial, las funciones de gestión de riesgo de proveedores y los responsables de compras durante la due diligence previa a la firma del contrato.

Postura	Estado
Aislamiento multi-tenant	PostgreSQL Row-Level Security en cada tabla de cliente
Cifrado en tránsito	TLS 1.2+, HSTS con `includeSubDomains`, solo cifrados AEAD
Cifrado en reposo	Cifrado a nivel de volumen + envelope para credenciales sensibles
Autenticación	Contraseñas Argon2id, Google + Microsoft Entra ID SSO, MFA TOTP
Registro de auditoría	Log append-only por org de eventos relevantes para seguridad
RGPD / CCPA	Conforme; DPA disponible; flujo DSR en producción
SOC 2 Type II	En curso (observación 2.º semestre 2026, auditoría 4.º trim. 2026)
ISO/IEC 27001:2022	Planeado (auditoría de certificación 1.º semestre 2027)
Recuperación ante desastres	Runbook documentado, RTO 30 min, backups Point-in-Time de 5 días, simulacros trimestrales

1. Govern — Programa de seguridad y rendición de cuentas

1.1 Propiedad de la seguridad

Un responsable de seguridad designado reporta al fundador/CEO y posee:

- Redacción y revisión de políticas de seguridad (cadencia anual; ad-hoc tras incidentes).

- Revisiones de riesgo de proveedores y decisiones de incorporación de subencargados.
- Coordinación de respuesta a incidentes.
- Gestión del programa de auditoría y certificación.
- Respuestas a cuestionarios de seguridad de clientes.

El rol platform-admin está vinculado a una cuenta de usuario específica vía la flag `is_platform_admin` (que reemplaza la verificación obsoleta por comparación de email); el privilegio deriva de la fila de base de datos, no del email. MFA es obligatorio para el rol platform-admin.

1.2 Políticas en vigor

Política	Alcance	Cadencia de revisión
Política de seguridad de la información	Toda la organización	Anual
Política de uso aceptable	Personal	Anual
Política de control de acceso	Ingeniería + ops	Anual
Política de gestión de proveedores	Compras	Anual
Plan de respuesta a incidentes	Ingeniería + dirección	Anual + post-incidente
Plan de recuperación ante desastres	Ingeniería	Simulacro trimestral
Ciclo de vida seguro de desarrollo	Ingeniería	Anual
Política de retención y eliminación de datos	Toda la organización	Anual
Gestión de claves criptográficas	Ingeniería	Anual
Gestión del cambio	Ingeniería	Anual

Las políticas se mantienen bajo control de versiones y son revisadas por el responsable de seguridad antes de cada ciclo de release. Los extractos elegibles para clientes están disponibles bajo NDA.

1.3 Seguridad del personal

- Verificaciones de antecedentes para personal que maneja datos regulados de cliente.
- Acuerdos de confidencialidad firmados al inicio del contrato.

- La incorporación incluye formación en concienciación sobre seguridad; refresco anual.
- El aprovisionamiento de acceso sigue el principio de menor privilegio; revisión trimestral de accesos.
- La baja se automatiza cuando es posible (de-aprovisionamiento SSO, rotación de secretos, revocación de repositorios).

1.4 Hoja de ruta de cumplimiento y certificación

Estándar	Estado	Objetivo
RGPD (UE 2016/679)	Conforme	Continuo
CCPA / CPRA	Conforme	Continuo
SOC 2 Type II	Periodo de observación 2.º sem. 2026; auditoría 4.º trim. 2026	Informe 1.º trim. 2027
ISO/IEC 27001:2022	Definición SGSI 2.º sem. 2026	Certificación 1.º sem. 2027
ISO/IEC 27701	Agrupada con 27001	2.º sem. 2027
IEC 62443 (industrial)	Hoja de ruta	Bajo demanda Enterprise

Para los nombres de las firmas de auditoría, las narrativas de control y los paquetes de evidencia pre-auditoría, contactar con security@devaisuite.com.

2. Identify — Gestión de activos y riesgo

2.1 Inventario de activos

La superficie de activos de la plataforma está acotada e inventariada:

- **Apps de aplicación:** `devai-api` (backend FastAPI), `devai-web` (frontend Next.js), `devai-admin` (admin/facturación), `devai-landing` (sitio de marketing). Todas desplegadas en Fly.io (Frankfurt FRA).
- **Almacenes de datos:** Fly Postgres (`devai-db`) para todos los datos relacionales incluidos los logs de auditoría; buckets de Cloudflare R2 `devai-suite-prod` (plantillas globales, sin datos de cliente) y `devai-tenant-docs-prod` (documentos por org).
- **Caché / colas:** Redis (rate limit, efímero); Inngest (orquestación de jobs en background).
- **Proveedores externos:** ver §6 Gestión de proveedores.

Una lista completa de subencargados con atribución de residencia de datos vive en `/subprocessor-list.html` .

2.2 Clasificación de datos

Clase	Ejemplos	Manejo
Pública	Páginas de marketing, este white paper	Sin restricción
Interna	Código fuente, runbooks internos	Acceso controlado, no visible al cliente
Confidencial del cliente	Proyectos APQP, FMEA, PPAP, registros de proveedores, documentos cargados, prompts/respuestas IA	Aislada por RLS; cifrado en reposo; acceso registrado
PII regulada	Email, nombre, dirección (donde se recoja), credenciales de auth	Igual que arriba + manejo PII dedicado por §3.4
Secretos	API keys, claves de cifrado, secrets OAuth	Fly Secrets store; envelope encryption para credenciales almacenadas; nunca en código ni logs

2.3 Gestión de riesgos

Los riesgos materiales se llevan en un registro interno revisado trimestralmente. Principales elementos a 2.º trim. 2026:

- **Postgres en una sola región** — punto único de fallo para la capa de base de datos; mitigado por backups Point-in-Time de 5 días + restauración documentada. La migración a multi-región o Fly Managed Postgres está en la hoja de ruta cuando la escala de cliente lo justifique.
- **Dependencia de proveedor LLM** — el AI Coordinator depende de APIs de modelos de terceros (Anthropic, OpenAI, Groq); cadena de fallback configurada, sin lock-in con un solo proveedor.
- **Bloqueo de dimensión de embeddings** — el esquema de pgvector está fijado en Voyage AI `voyage-3-large` 1024

dimensiones; la rotación de proveedor requiere una re-migración coordinada documentada en el runbook DR.

2.4 Modelo de amenazas

La plataforma se modela frente a:

- **Atacantes externos** — credential stuffing, account takeover, injection, SSRF, compromiso de la cadena de suministro.
- **Tenants maliciosos o comprometidos** — intentos de acceso a datos cross-tenant vía enumeración de API, RAG poisoning, prompt injection LLM.
- **Amenaza interna** — ingenieros sobre-privilegiados, personal que ha dejado la empresa.
- **Compromiso de tercero** — brecha de subencargado que afecte a nuestros datos.

Las mitigaciones se describen en las secciones correspondientes a continuación.

3. Protect — Controles

3.1 Multi-tenancy y aislamiento de datos

Los datos de cada cliente residen en una única base de datos PostgreSQL compartida, aislados mediante **PostgreSQL Row-Level Security**. Cada tabla con datos de cliente tiene una política RLS atada al contexto del tenant de la solicitud. La capa de aplicación define el tenant en cada sesión de base de datos a través del middleware de autenticación de la solicitud:

```
ALTER TABLE deliverables ENABLE ROW LEVEL SECURITY;
CREATE POLICY tenant_isolation ON deliverables
  USING (org_id = current_setting('app.current_org_id', true)::int);
```

Una solicitud que filtre el contexto del tenant (incluso por un bug interno de aplicación) no puede leer ni escribir filas pertenecientes a otro tenant — la base de datos lo rechaza. Las consultas SELECT cross-tenant devuelven cero filas por diseño.

El helper `app.db.tenant.set_tenant_context` es el único punto por el que se aplica el contexto de tenant; se ejecuta en cada solicitud que resuelva un usuario, antes de que se ejecute cualquier lógica de negocio.

3.2 Almacenamiento de documentos por tenant

Los documentos para RAG viven en Cloudflare R2 con claves de objeto prefijadas por `org_id`. Las lecturas están limitadas por el mismo contexto de tenant, con una verificación en servidor antes de la emisión de cualquier presigned-URL.

Los embeddings (Voyage AI `voyage-3-large`, 1024 dim) se almacenan en tablas pgvector que llevan una columna `org_id` bajo la misma política RLS que los documentos fuente.

3.3 Cifrado

En tránsito. TLS 1.2+ en todos los endpoints accesibles al cliente. Solo cipher suites fuertes (AES-GCM, ChaCha20-Poly1305). HSTS aplicado con `max-age=31536000; includeSubDomains`. La directiva `preload` se omite intencionadamente hasta que cada subdominio cumpla los criterios de elegibilidad. X-Frame-Options está en `DENY` y X-Content-Type-Options en `nosniff` en cada respuesta.

En reposo. Los datos de Postgres se cifran en reposo mediante la capa de volumen de Fly. Cloudflare R2 cifra en reposo por defecto (AES-256). Las credenciales sensibles de integración almacenadas en `integration_connectors.credentials_json` se cifran mediante envelope con una clave Fernet por despliegue (`DEVAI_CREDENTIALS_ENCRYPTION_KEY`); la clave se rota en una cadencia documentada y nunca sale del store de secretos.

Gestión de claves. Las claves de aplicación (firma JWT, Fernet, secrets OAuth, API keys de terceros) se almacenan en Fly Secrets, se exponen al

proceso en arranque y nunca se persisten en disco. Los procedimientos de rotación están documentados por proveedor en `docs/runbooks/disaster-recovery.md` §3.

3.4 Autenticación y control de acceso

Opciones de identidad:

- Email/contraseña — contraseñas hasheadas con Argon2id (memory cost 64MB, time cost 2, parallelism 1; ajustable). No se aplica historial de contraseñas; longitud mínima 12 caracteres con requisitos de complejidad (configurable por org).
- Google SSO vía OAuth 2.0.
- Microsoft Entra ID SSO vía OAuth 2.0 / OIDC, por defecto al tenant `organizations` (cuentas de trabajo/escuela) — las cuentas MSA personales se bloquean mediante el gate `is_business_domain`.
- Los usuarios solo-SSO tienen contraseña `NULL` y la ruta de login con contraseña los rechaza en lugar de ofrecer un fallback inseguro.

Autenticación multifactor. Basada en TOTP, opcional según política de org, obligatoria para el rol platform-admin. La verificación MFA se ejecuta en la dependencia de auth en cada solicitud protegida; los usuarios con MFA habilitado pero no verificado solo pueden alcanzar los endpoints `/auth/mfa/verify` y `/auth/mfa/status` hasta que verifiquen.

Autorización basada en roles. Cada ruta protegida llama `ensure_permission(db, user, resource_type, action, org_id)` que verifica el grant rol-permiso del usuario contra la acción solicitada. Las decisiones se registran en una tabla de auditoría `policy_decisions` append-only.

Tokens de sesión. JWTs de corta duración (12 horas por defecto) con rotación en acciones sensibles (cambio de contraseña, alta MFA). Almacenados en cookies `HttpOnly; Secure; SameSite=Lax` Y reflejados en `localStorage` para el SPA — ambos deben coincidir en cada solicitud. La reutilización de tokens tras invalidación se detecta verificando contra una blocklist actualizada en logout / revocación.

CAPTCHA. Cloudflare Turnstile protege registro, login, restablecimiento de contraseña, alta MFA y la ruta de provisioning del sandbox público de demo. El helper `verify_turnstile_token` registra el campo `error-codes` de Cloudflare en caso de fallo para diagnóstico.

Rate limiting. Un middleware con backend Redis aplica límites por endpoint y por IP — endpoints de auth limitados a 10 solicitudes/minuto, AI Coordinator feedback a 30/minuto, defecto en 100/minuto. Los límites fallan abiertos cuando Redis no está disponible (disponibilidad por encima de límites rígidos) pero registran el fallo para revisión.

3.5 Seguridad de aplicación

- Alineación con OWASP ASVS Level 2 como base.
- Validación de entrada vía esquemas Pydantic en cada límite de API; estrictez de tipos en tiempo de compilación con mypy.
- Acceso a datos solo vía ORM (SQLAlchemy); ningún SQL concatenado por strings.
- Las respuestas HTTP llevan headers de seguridad estrictos (HSTS, X-Frame-Options=DENY, X-Content-Type-Options=nosniff, Referrer-Policy=strict-origin-when-cross-origin). CSP está en la hoja de ruta y supeditada a una pasada de pruebas exhaustiva para no romper los embeds del AI Coordinator.
- Defensa contra SSRF: las salidas HTTP de la aplicación están limitadas a una allowlist conocida de hostnames de proveedores; no se permiten salidas arbitrarias desde rutas de código atadas a solicitudes.
- Codificación de salida: React (web app) y Jinja2 (admin) auto-escapan; el uso explícito de `dangerouslySetInnerHTML` se revisa en CI con grep.

3.6 Protecciones específicas de LLM

La plataforma se integra con Anthropic, OpenAI, Groq, Voyage AI, RunPod y (por org y opt-in) Together. Riesgos específicos:

- **Inyección de prompt desde documentos cargados.** Todo contenido recuperado por RAG se entrega al LLM en un bloque "context" claramente delimitado; las instrucciones en documentos recuperados se acotan frente al system prompt y se filtran post-hoc para patrones conocidos de jailbreak. Las propuestas de tool-calling se exponen para revisión del usuario antes de ejecutarse (ciclo de vida AIProposal).
- **Exfiltración de datos vía LLM.** Ningún dato de cliente sale del contexto de tenant de la plataforma salvo como parte de una llamada de inferencia al LLM; las cláusulas de manejo de datos del proveedor IA se revisan en la incorporación, y se requiere opt-in por org antes de enviar datos a proveedores que los retengan para entrenamiento.
- **Agotamiento de coste.** Cuotas por org y por usuario (configurables; defecto conservador). El sandbox público de demo aplica una cuota de 10 mensajes del AI Coordinator por sesión.
- **Fine-tuning privado por tenant.** La fase 3 del track de IA introduce adaptadores LoRA por tenant entrenados solo con datos de la propia org. Los adaptadores nunca cruzan tenants; el router de inferencia selecciona el adaptador correcto por `org_id` antes de llamar al servidor de modelo.

3.7 Gestión de secretos

Superficie	Almacén	Rotación
Secretos de aplicación (JWT, Fernet, API keys)	Fly Secrets	Por proveedor — ver `docs/runbooks/disaster-recovery.md` §3
Credenciales de integración del cliente (Salesforce, SAP, etc.)	Cifradas con envelope en `integration_connectors.credentials_json`; la clave en Fly Secrets	Iniciada por el cliente
Secrets de cliente OAuth	Fly Secrets	Anual o post-incidente
API keys de terceros (Anthropic, OpenAI, Voyage, etc.)	Fly Secrets	Por proveedor; documentado en runbook DR

Ningún secreto se commitea jamás al control de versiones; los hooks pre-commit y CI escanean por commits accidentales.

4. Detect — Monitorización y auditoría

4.1 Logging de aplicación

Se emiten logs estructurados en JSON por cada solicitud y evento relevante para seguridad. El saneado de PII se aplica en el filtro de logging (`app.core.logging_filters.PIIScrubFilter`) antes de que los logs salgan de la aplicación; emails, direcciones IP y otros identificadores se redactan en entornos no-debug.

4.2 Trazabilidad de auditoría

Una tabla dedicada `org_security_audits` append-only registra eventos relevantes para seguridad acotados por org:

- Autenticación: login, logout, alta MFA, resultado de challenge MFA, restablecimiento de contraseña, vinculación/desvinculación SSO.
- Autorización: grants de rol, revocaciones de rol, denegaciones de permiso.
- Acceso a datos: envíos PPAP, ediciones FMEA, aceptar/rechazar propuestas IA, lectura/escritura de credenciales de integración, toggle del sandbox.
- Ciclo de vida de la org: creación, solicitud de eliminación, cancelación de eliminación, conversión desde demo.

Los administradores cliente pueden solicitar exportación del histórico de auditoría de su org vía el flujo DSR. Retención: 7 años para artefactos regulados, 1 año para eventos generales, configurable por cliente en el DPA.

4.3 Seguimiento de errores

Sentry captura excepciones de aplicación con la integración FastAPI (transacciones etiquetadas por endpoint), la integración SQLAlchemy (detección de queries lentas) y el filtro de saneado PII aplicado en la capa

de logging. La tasa de profiling está al 1% de transacciones trazadas para equilibrar coste y señal.

4.4 Monitorización de infraestructura

Las health probes (`/livez` , `/readyz`) reportan en el chequeo automático de Fly. Readiness pinguea la base de datos (dependencia dura) y Redis (suave, fail-open). La integración con página de estado está en la hoja de ruta.

4.5 Detección de anomalías

El middleware de rate-limit emite headers estándar `X-RateLimit-{Limit,Remaining,Reset}` y registra anomalías (picos súbitos de una sola IP) para revisión. La integración con un SIEM está en la hoja de ruta una vez SOC 2 Type II esté firmado.

5. Respond — Respuesta a incidentes

5.1 Clasificación de incidentes

Severidad	Ejemplos	Respuesta
P0	Brecha de datos confirmada; caída total de plataforma > 30 min	Paginar on-call inmediatamente; actualización de página de estado en 15 min; war-room
P1	Caída parcial; evento de seguridad sin afectación de datos	On-call paginado; actualización de estado en 30 min
P2	Degradación del servicio; actividad sospechosa bajo investigación	Triaje en horario laboral
P3	Vulnerabilidad reportada bajo triaje; sin explotación activa	Acuse de recibo en 48h

5.2 Checklist de los primeros 30 minutos

Una checklist lista para pegar para el operador vive en `docs/runbooks/disaster-recovery.md` §5:

1. Confirmar el impacto vía `/readyz` .
2. Probar el estado de máquinas Fly y los logs recientes.
3. Categorizar la clase de fallo (DB / Redis / bug de app / infra Fly).
4. Publicar actualización en página de estado antes de arreglar — fija expectativas del cliente.
5. Mitigar según la sección correspondiente del runbook.
6. Comunicar directamente a los tenants afectados.

5.3 Notificación al cliente

Cuando lo exige la ley para violaciones de datos personales bajo el Artículo 33 del RGPD, la notificación al cliente se realiza dentro de las **72 horas desde la confirmación**. La notificación incluye la naturaleza de la brecha, las categorías y número aproximado de interesados, las consecuencias probables y las medidas adoptadas o propuestas para abordarla.

5.4 Revisión post-incidente

Se redacta un post-mortem sin culpa en 5 días laborables para cada incidente P0 y P1, capturando línea temporal, causa raíz, factores contribuyentes y seguimientos concretos. Los post-mortems se archivan internamente y se comparten con clientes afectados bajo NDA.

5.5 Divulgación de vulnerabilidades

Publicamos un `security.txt` en `/.well-known/security.txt` con un contacto de reporte (security@devaisuite.com) y una ventana de divulgación de 90 días. Los investigadores que actúen de buena fe no están sujetos a acción legal bajo nuestro lenguaje de safe-harbor.

6. Recover — Recuperación ante desastres y continuidad

6.1 Postura de backups

Las snapshots automatizadas diarias de PostgreSQL se conservan durante 5 días. El RTO objetivo para una restauración full-stack es 30 minutos; el RPO es de hasta 24 horas de escrituras (la recuperación Point-in-Time sub-diaria requiere migración a Fly Managed Postgres, en la hoja de ruta).

6.2 Procedimiento de restauración

Documentado en detalle en `docs/runbooks/disaster-recovery.md` §1. Resumen:

1. Detener la API para prevenir escrituras split-brain.
2. Identificar el ID de snapshot anterior a la ventana de corrupción.
3. Restaurar como una nueva app Postgres (NO sobrescribir in-place — preserva el original para análisis forense).
4. Validar contra baselines de métricas de negocio.
5. Reapuntar `DATABASE_URL` al cluster restaurado.
6. Devolver la API a online.
7. Smoke-test vía `/readyz` y los endpoints de salud por proveedor.

6.3 Recuperación de R2 (object storage)

Cloudflare R2 es de región única. "Failover" para el bucket global de plantillas significa regenerar desde el código fuente vía `publish_global_library`; para el bucket de docs por tenant requiere que el versionado R2 esté habilitado (verificado en despliegue).

6.4 Cadencia de simulacros DR

Simulacros trimestrales contra el Postgres de staging (`devai-api-stg-pg`):

Trimestre	Foco
1.º	Restauración Postgres (simulacro completo §1.5 sobre staging)
2.º	Rotación de secretos (rotar un proveedor de extremo a extremo)
3.º	Regeneración de bucket R2 desde código fuente
4.º	Simulacro de incidente sobre papel (tabletop)

Un simulacro fallido dispara un P1 para actualizar el runbook o la configuración subyacente antes de la próxima cadencia de simulacro.

6.5 Continuidad de negocio

Las funciones críticas de negocio (facturación, soporte al cliente, rotación on-call) están documentadas en playbooks separados de la plataforma misma, de modo que la indisponibilidad de la plataforma no bloquee los esfuerzos de recuperación del cliente.

7. Gestión de proveedores y subencargados

7.1 Incorporación

Cada nuevo subencargado se revisa por:

- Residencia de datos (dónde almacenan nuestros datos de cliente).
- Postura contractual: DPA, SCCs (para procesadores no EEE que manejen datos personales UE), términos de responsabilidad.
- Postura de seguridad: certificaciones publicadas (SOC 2, ISO 27001), páginas públicas de seguridad, historial de brechas.
- Dependencia operativa: criticidad para el servicio, facilidad de cambio.

La decisión se registra en el registro de proveedores; la revisión de renovación es anual.

7.2 Lista actual de subencargados (extracto)

Proveedor	Propósito	Región	Cumplimiento clave
Fly.io	Compute, Postgres gestionado	UE (FRA primaria)	SOC 2 Type II
Cloudflare	DNS, R2 storage, Turnstile	Edge global	SOC 2 Type II, ISO 27001
Anthropic	Inferencia LLM (Claude)	EE.UU.	SOC 2 Type II
OpenAI	Inferencia LLM (GPT)	EE.UU. (opt-in)	SOC 2 Type II
Groq	Inferencia LLM (baja latencia)	EE.UU.	SOC 2 Type II
Voyage AI	Embeddings	EE.UU.	DPA en vigor
Stripe	Facturación y pagos	EE.UU., UE	SOC 1, SOC 2, PCI DSS Level 1
Resend	Correo transaccional	EE.UU.	SOC 2 Type II
Sentry	Seguimiento de errores (PII saneado)	EE.UU.	SOC 2 Type II, ISO 27001
Inngest	Orquestación de jobs background	EE.UU.	DPA en vigor
RunPod	Serving LLM (adaptadores por tenant)	EE.UU. (multi-región)	DPA en vigor
HuggingFace	Almacenamiento de artefactos de adaptadores	EE.UU.	DPA en vigor

La lista completa siempre actualizada con números de versión y enlaces al DPA de cada proveedor está publicada en `/subprocessor-list.html` . La notificación al cliente de cambios de subencargado sigue la cadencia del DPA del cliente (típicamente 30 días de antelación para cambios materiales).

7.3 Revisión de cambios de subencargado

Los clientes son notificados de cambios de subencargado según la cadencia del DPA y pueden objetar dentro de la ventana contractual. Cuando la objeción del cliente no puede reconciliarse con el nuevo subencargado, el contrato permite la rescisión según el DPA.

8. Ciclo de vida seguro de desarrollo

8.1 Control de fuente y revisión

- Repositorio canónico único en GitHub.
- Branch protection en `main` : revisiones PR obligatorias, status checks obligatorios, force-push deshabilitado.
- Hooks pre-commit: lint, type check, escaneo de secretos.
- CI en cada PR: suite completa de tests (actualmente ~3.000 tests en niveles smoke y full).

8.2 Pruebas

Tipo de prueba	Alcance	Cadencia
Tests unitarios	Lógica de servicio, validadores	Cada PR
Tests de integración	Rutas + DB + middleware vía FastAPI TestClient	Cada PR
Tests smoke	Subconjunto de 200 tests de mayor leverage	Cada PR (CI nivel-PR)
Sweep completo	Los 3.000 tests	Push a `main` / `staging`
Gate de drift de esquema	Spec OpenAPI generada contra Python commiteado	Cada PR
Vitest (frontend)	Componentes React, hooks	Cada PR
Type checks	mypy (Python) + tsc (TypeScript)	Cada PR
Regresiones de seguridad	Archivos de test específicos para RLS, RBAC, auth, billing	Cada PR (nivel smoke)

8.3 Gestión de dependencias

- GitHub Dependabot habilitado; PRs semanales para advisories de seguridad.
- Revisión manual de cada actualización de dependencia; los parches relevantes para seguridad se aplican en 7 días para advisories de severidad alta, 30 días para media.
- Sin frameworks mayores obsoletos: Python 3.11 (canónico CI), Node 20+, FastAPI / Pydantic / SQLAlchemy actuales.

8.4 Disciplina de migraciones

Los cambios de esquema de base de datos fluyen mediante migraciones Alembic, revisadas e idempotentes. Producción ejecuta `alembic upgrade`

`head` como pre-flight del comando de release. Los rollbacks están probados; las migraciones destructivas llevan un gate `--confirm`.

8.5 Análisis estático

- El type checking es la herramienta principal de análisis estático (mypy + tsc strict).
- `flake8` para estilo Python.
- ESLint para TypeScript.
- Una integración SAST futura (Semgrep o similar) está en la hoja de ruta SOC 2.

8.6 Validación de configuración en arranque

`validate_production_config(settings)` se ejecuta en el lifespan startup de FastAPI. La aplicación rechaza el boot cuando:

- Falta un secreto requerido en producción.
- Está activo `STRIPE_LIVE=true` pero la clave Stripe no es `sk_live_` (mismatch de modo live).
- Otros invariantes requeridos para evitar estados inseguros.

Esto detecta la deriva de configuración antes de que pueda servir una sola solicitud.

9. Manejo de datos

9.1 Ciclo de vida de los datos

Etapa	Manejo
Recolección	Iniciada por el cliente vía la UI de la plataforma, API o carga de documento. No raspamos ni compramos datos de clientes.
Almacenamiento	Cifrado en reposo, aislado por RLS, fijado por región a UE (FRA) para la base de datos primaria.
Uso	Acotado por contexto de tenant en cada solicitud. Las llamadas de inferencia IA están claramente delimitadas y el proveedor IA queda registrado en el log de auditoría de la org.
Retención	Según el DPA del cliente. Defecto: artefactos retenidos durante la duración del contrato + 7 años para datos regulados; datos generales retenidos durante la duración del contrato + 1 año.
Eliminación	Eliminación de org self-service vía la UI de admin (Sprint 8 D3 — ventana de gracia de 30 días + restauración, luego hard-delete vía cron diario de purga). La eliminación de usuario impulsada por DSR se procesa en 30 días.

9.2 Residencia de datos

La base de datos primaria y la capa de aplicación corren en la región Frankfurt (FRA) de Fly. Los buckets de Cloudflare R2 están en jurisdicción UE; las regiones de proveedor LLM varían y están documentadas en la lista de subencargados.

Para clientes con requisitos estrictos de residencia, hay despliegue on-premise / región dedicada disponible en planes Enterprise.

9.3 Derechos del interesado (RGPD / CCPA / CPRA)

El flujo DSR soporta:

- Derecho de acceso — exportación completa de los datos del cliente bajo solicitud.
- Derecho de eliminación — en 30 días, con confirmación al solicitante.
- Derecho de rectificación — gestionado vía la UI por administradores de org.
- Derecho de portabilidad — exportación JSON alineada con el mismo contrato de exportación que el acceso.
- Derecho de objeción / restricción — gestionado según la tabla de fines de tratamiento del DPA.

Las solicitudes se registran y trazan a través del log de auditoría.

9.4 Transferencia transfronteriza

Cuando los datos personales de interesados UE son procesados por subencargados con sede en EE.UU. (proveedores LLM, Stripe, etc.), la transferencia se rige por:

- Cláusulas Contractuales Tipo (SCCs) incluidas en el DPA.
- Medidas suplementarias cuando sea necesario (cifrado en tránsito, sin logging de texto plano en el proveedor).

La evaluación completa de impacto de transferencia es interna pero está disponible para clientes Enterprise bajo NDA.

10. Consideraciones específicas de IA

10.1 Qué se envía a los proveedores LLM

Las llamadas de inferencia envían:

- El prompt del usuario (system prompt + mensaje del usuario).
- El contexto RAG recuperado (solo documentos del propio usuario).
- Puntero al adaptador por tenant (solo router de inferencia Fase 3 — nunca un adaptador de otro tenant).

Lo que NO se envía:

- Datos de otros clientes.
- Secretos de aplicación, API keys, credenciales del sistema.
- Tokens de autenticación del cliente.

10.2 Manejo de datos del proveedor

Las cláusulas de manejo de datos de cada proveedor LLM se revisan en la incorporación del proveedor. Los proveedores que retienen prompts/respuestas para entrenamiento están protegidos detrás de una flag opt-in por org (`ai_training_opt_in`); el defecto es opt-out.

10.3 Revisión de propuestas de IA

El AI Coordinator de la plataforma nunca ejecuta acciones destructivas en silencio. Cualquier cambio generado por LLM en datos del cliente fluye por el ciclo de vida AIProposal:

1. La IA genera una propuesta (fila FMEA, tile de dashboard, clasificación de documento).
2. La propuesta se almacena con `status=draft` y se expone al usuario.
3. El usuario revisa y acepta explícitamente; el rechazo también se registra con texto de corrección opcional para tenant-learning.

4. Solo en aceptación la plataforma commitea el cambio a los datos del cliente.

Esto previene la corrupción silenciosa de datos por una respuesta LLM alucinada y crea un audit trail para cada cambio impulsado por IA.

10.4 Aislamiento de tenant-learning

La fase 3 del track de IA introduce adaptadores LoRA por tenant entrenados sobre las correcciones de la org a las propuestas IA. Propiedades críticas de aislamiento:

- Los datos de entrenamiento de cada adaptador se obtienen exclusivamente de las filas `tenant_learning_events` de una org, acotadas por RLS.
- Los adaptadores entrenados se etiquetan con `org_id` y se almacenan en HuggingFace Hub bajo el namespace de la organización.
- El router de inferencia selecciona el adaptador activo para el `org_id` de la solicitud antes de despachar a RunPod / Together; la selección de adaptador cross-tenant es imposible por diseño (probado en la suite Phase 3c wiring).
- Las orgs demo (`is_demo=True`) y las orgs sandbox (`is_sandbox=True`) están excluidas de los corpora de entrenamiento.

11. Responsabilidades del cliente (modelo compartido)

La seguridad es compartida. El cliente es responsable de:

- **Gestión de usuarios:** añadir, eliminar y revisar usuarios autorizados de su org. Deshabilitar prontamente a usuarios que han dejado la empresa.

- **Higiene de credenciales:** contraseñas fuertes; MFA donde la política de la org lo permita; no compartir credenciales.
- **Configuración:** habilitar dominios de email permitidos, exigencia de MFA y los ajustes de residencia de datos apropiados a la postura de cumplimiento del cliente.
- **Clasificación de datos:** evaluar qué datos elige cargar a la plataforma. La plataforma es adecuada para datos típicos de APQP/manufactura; clientes con clasificaciones más estrictas (controlados por exportación, clasificados, etc.) deben consultar antes de cargar.
- **Validación de salidas:** las salidas generadas por IA deben revisarse antes de basarse en ellas para decisiones reguladas (envíos PPAP, entregables al cliente, etc.). Los indicadores de confianza del AI Coordinator son un punto de partida, no un sustituto de la revisión por SME.
- **Credenciales de integración:** rotar las credenciales propias para sistemas conectados (Salesforce, SAP, ERP, MES) según la política interna.

Una copia de esta lista de responsabilidades del cliente está en el Master Service Agreement y el DPA.

12. Contactos

Canal	Dirección
Consultas de seguridad	security@devaisuite.com
Divulgación de vulnerabilidades	/.well-known/security.txt
Privacidad / DPA / DSR	privacy@devaisuite.com
Página de estado	https://status.devaisuite.com
Ventas / contratos	sales@devaisuite.com

La clave pública PGP para correspondencia de seguridad está disponible bajo solicitud.

13. Control documental

Campo	Valor
Versión	1.0
Fecha de entrada en vigor	12 de marzo de 2026
Última actualización	29 de abril de 2026
Propietario	Responsable de seguridad
Cadencia de revisión	Anual + post-incidente
Distribución	Pública (esta versión); las versiones internas siguen los controles en marcha

Este white paper se revisa al menos anualmente y en cada cambio material en la postura de seguridad de la plataforma. La fuente

autorizada es `docs/security/WHITEPAPER_ES.md` en el repositorio fuente de la plataforma.

Una versión en inglés se mantiene en paralelo en `docs/security/WHITEPAPER_EN.md` y se renderiza en `/security-whitepaper.html`.