Skyseat · Hoja de seguridad

Skyseat permite a una empresa tener empleados IA, cada uno con su ordenador en la nube, sus permisos y una persona al mando. Skyseat está en acceso anticipado. Esta hoja dice lo que existe hoy y, en el recuadro rojo, lo que no.

DÓNDE ESTÁN LOS DATOS

- Plataforma, base de datos y caja fuerte: Hetzner, **Helsinki, Finlandia**. Ordenadores de los empleados IA: Hetzner, Alemania. Hetzner tiene ISO 27001; nosotros no operamos centros de datos.
- Hay una **región de EE. UU.** construida tras un interruptor; se creará cuando la necesite el primer cliente de EE. UU.
- Los datos de cada empresa están separados. Una prueba automática intenta cada ruta del panel con datos de otra empresa (todas devuelven «no encontrado»); las rutas nuevas se añaden solas.
- El modelo de IA (Claude, de Anthropic) y los proveedores de voz y teléfono están en **EE. UU.** sea cual sea la región; los subencargados son públicos, con 30 días de aviso de cambios.

CIFRADO Y CLAVES

- HTTPS/TLS en todo el tráfico hacia la plataforma.
- Credenciales de los servicios conectados: AES-256-GCM en una caja fuerte; la clave se guarda fuera de la base de datos. El ordenador de un empleado IA nunca las ve: la plataforma hace la llamada, comprueba el permiso y la registra.
- Contraseñas: scrypt, con las contraseñas comunes rechazadas. Nadie pone la contraseña de otra persona (enlaces de invitación de un solo uso).
- Sin claves gestionadas por el cliente (BYOK), HSM ni rotación automática de claves.

APROBACIÓN HUMANA

- Por empleado IA: pregunta antes de todo, antes de lo sensible o nunca. En todos los niveles siguen activas tres salvaguardas: tope de gasto, aviso de IA a terceros y registro.
- **Cuatro ojos:** los pagos por encima de un importe, los borrados, los envíos masivos o los cambios de permisos pueden exigir a dos personas distintas, ninguna la que encargó la tarea. Nunca se autoaprueban ni se aprueban en bloque; un rechazo bloquea. Aún no cubre: los cambios de permisos hechos a mano en el panel.
- Todo lo que se pide por teléfono espera la aprobación de una persona.

REGISTRO DE ACTIVIDAD VERIFICABLE

- Se registran las acciones de personas y empleados IA; la base de datos impide editar y borrar. Exportación de cualquier periodo en PDF, CSV o JSON, **firmada con Ed25519 y con filas encadenadas por hash**.
- Un auditor puede comprobar una exportación con un script independiente, sin Skyseat. Demuestra que el archivo no se alteró tras exportarlo; por sí solo no demuestra que la base de datos no se tocara antes.
- La retención la fija cada empresa; si se limita, nunca baja de 2 años, y cada purga queda registrada.

IDENTIDAD Y ACCESO

- **SSO:** OIDC y SAML 2.0 (Okta, Microsoft Entra ID, Google Workspace, ADFS y otros). Solo dominios de correo verificados por DNS; se puede hacer obligatorio.
- Aprovisionamiento **SCIM 2.0**: el proveedor crea, actualiza y da de baja a las personas; los grupos fijan el rol (nunca administrador).
- **Verificación en dos pasos** (app TOTP + 10 códigos de recuperación de un solo uso), exigible a toda la empresa; siempre obligatoria para los operadores de Skyseat. Los códigos erróneos tienen límite por persona.
- Roles por empresa (administrador, responsable, solo lectura) más permisos por empleado IA, comprobados con barridos automáticos. Límites de intentos de acceso por correo, por IP y en total.
- Sesiones: quien administra fija cuánto duran como mucho (de 8 h a 30 días; de serie, 30 días que se alargan con el uso) y puede cerrar las de todos. Cada entrada queda en el Registro de la empresa. Quien administra ve y revoca cada clave de API que alcanza a la empresa.

COPIAS Y RECUPERACIÓN

- Copia diaria cifrada (AES-256) y otra antes de cada despliegue, guardadas en otro centro de datos (Falkenstein): 14 diarias + 8 semanales. Prueba de restauración automática cada semana.
- RPO de hasta 24 h. **Simulacro de desastre del 31 de octubre de 2026:** servidor nuevo reconstruido solo con la copia externa en 4 min 18 s, con todos los datos legibles; 20–30 min realistas de principio a fin.
- Un servidor de plataforma en una región (aún sin reserva en caliente). Página de estado y monitorización desde 4 regiones. **Sin SLA contractual.**

CÓMO USAMOS LA IA

- Cada tarea envía al modelo solo lo que necesita. Lo que se lee en correos, páginas web o archivos se trata como información, no como órdenes; las acciones sensibles necesitan aprobación. No afirmamos que la inyección de instrucciones esté totalmente impedida.
- La plataforma (no el empleado IA) añade el aviso de IA al correo, el chat, las llamadas y las reuniones; la empresa puede reformularlo, pero no quitarlo.
- **Sin entrenar con tus datos.** La API comercial de Anthropic no entrena con lo que recibe. La plataforma pasará a una API comercial o a un plan de empresa antes de la disponibilidad general.

TUS DATOS

- Descárgalos todos cuando quieras; al cerrar la cuenta se borran todos tras 30 días. El encargo de tratamiento (art. 28 del RGPD y un anexo de leyes estatales de EE. UU.) está redactado y pendiente de revisión jurídica; aviso de brecha en 48 horas desde que la detectamos, para que tengas tus 72 horas ante la autoridad (borrador).

LO QUE AÚN NO TENEMOS

- **Ni SOC 2, ni ISO 27001, ni HIPAA, ni PCI, ni FedRAMP.** Está previsto SOC 2 Type I; no hemos contratado a un auditor y **no hay fecha confirmada**. Existen un mapa de controles y borradores de políticas.
- **Ningún pentest externo.** Hacemos nuestras propias revisiones de seguridad (hallazgos corregidos, cada uno con su prueba) y una prueba ligera desde fuera de nuestros dominios públicos; no son independientes.
- Las políticas son borradores. Sin programa formal de seguridad de RR. HH., revisiones formales de accesos, DPO ni MDM. Previstas las copias continuas de WAL y la reserva en caliente. Aún sin buzón `security@` ni `security.txt`: avisa de cualquier fallo desde el panel (Ayuda y soporte, respuesta en 1 día laborable).
- Algunas integraciones (teléfono, correo propio de los empleados, WhatsApp, Microsoft 365, Google Workspace) siguen en modo de prueba. Sin VPAT ni auditoría formal de accesibilidad.