Viktor, the AI employee for Slack and Microsoft Teams

Viktor security at a glance

Viktor is an AI employee that works inside your Slack or Microsoft Teams workspace and the Viktor web app. This page summarises how the platform protects your data and keeps every action under your control. Audit reports and test summaries are available in the Trust Center at trust.zetalabs.ai; the DPA is public at viktor.com/legal.

SOC 2	Type I attestation completed; Type II audit in progress with an independent audit firm. Report and status in the Trust Center.
GDPR	Data Processing Agreement with Standard Contractual Clauses, published at viktor.com/legal
Penetration testing	Independent tests performed regularly, covering the web app, API, Slack app and the AI layer. Executive summary in the Trust Center.
Security contact	security@viktor.com and trust.zetalabs.ai

1 How Viktor is built

Each AI employee runs in its own isolated environment and never holds a credential.

- One isolated micro-VM per AI employee for skills, memory, files and code execution; nothing is shared with other AI employees or customers.
- Every external action is an API call through a tool gateway that attaches credentials at call time, enforces the permission level, applies approvals and writes the audit event.
- Credentials live only on the backend, encrypted. They are never written to the sandbox, a log or a model context.
- Public endpoints sit behind a web application firewall with DDoS protection; inbound Slack and Teams events are signature-verified. Hosted in the United States (US-East) on audited cloud providers.

2 Your data

Your data stays yours, is encrypted, and never trains a model.

- TLS 1.2+ in transit, AES-256 at rest; tokens and secrets additionally encrypted at application level.
- No training: prompts, outputs, code, tool traces and memory are not used to train viktor.com or third-party foundation models. LLM providers (Anthropic, Google, OpenAI) are called per request under no-training terms.
- You remain the controller; viktor.com processes as a processor under the DPA.
- Retention: conversation history for the life of the AI employee, audit events 365 days, operational logs 95 days. Deletion on written instruction at any time and within one month of termination.
- Regulated data such as PHI or cardholder data is out of scope of the standard service unless a separate written addendum is signed.

3 Who can use Viktor and what it can reach

Access follows your identity provider and the accounts you connect.

- Web app sign-in through your identity provider (SAML/OIDC, Okta supported); MFA is enforced by your IdP. Directory sync (SCIM) for provisioning and deprovisioning is available for enterprise workspaces.
- Administrators control integrations, permission levels, approvals, access lists, the audit log and stop/disconnect controls. An access list restricts who may talk to an AI employee at all; everyone else gets no answer. Viktor only sees channels it has been added to.
- Integrations connect via OAuth (2.1 with PKCE or client credentials) or a customer-issued token. Every action runs as the account that consented, so Viktor can never exceed that account's permissions; dedicated service accounts are recommended.
- In Slack and Teams, Viktor is an app with a bot user, never a human account. Internal systems connect as custom integrations (MCP server or REST API).

4 Controls on what the AI employee may do

Every tool is automatic, requires approval or forbidden. Deny wins and a no is final.

- 'Requires approval' pauses the run and posts a preview of what will happen, where and with which data. A human approves or rejects before anything executes; rejected or forbidden actions never run and automated runs cannot self-approve.
- Integration write tools and high-impact actions (permissions, access lists, spend) require approval by default. The exact split is agreed per deployment.
- Stop at any time: stop a running task in-thread, pause or delete an automation, disconnect an integration, remove a person from the access list, revoke API keys, or uninstall the app. Revoking a credential at the source stops Viktor at its next call because the sandbox caches nothing.

5 Memory isolation and prompt injection

Content Viktor reads is treated as untrusted; controls do not rely on the model spotting an attack.

- Memory is a set of files inside the AI employee's own sandbox. No other sandbox, AI employee or customer can read it, and there is no shared vector store built from customer data.
- Credential theft is not possible because credentials never enter the sandbox. Cross-account access is bounded by the connected accounts' own permissions.
- A second adversarial-content check runs before any approved action executes. What an AI employee has learned is visible in the app and can be forgotten or reset by an administrator.

6 How viktor.com secures itself

Independent testing, gated change control and continuous control monitoring.

- Independent penetration tests of the web app, API and Slack app, including prompt injection, instruction leakage and abuse of model-integrated actions. Executive summary in the Trust Center.
- Code reaches production only through a reviewed pull request with CI, automated code review, dependency and secret scanning, then staged deployment (dev, staging, production).
- Vulnerability remediation SLAs: critical 30 days, high 60 days, medium 120 days; actively exploited issues patched out of band.
- Documented incident response plan with defined roles and severities, tested at least annually. Personal data breaches notified within 72 hours of confirmation, per the DPA.
- Company devices under MDM; SSO and MFA on critical systems; sub-processors assessed before onboarding and reviewed annually.

7 Shared responsibility

Area	You	viktor.com
Identity	You run the IdP, MFA policy and access lists	We enforce SSO, SCIM, access lists and roles
Integration accounts	You create and scope service accounts, revoke at source	We store tokens encrypted, attach at call time, revoke on disconnect
Permissions	You set the automatic / approval / forbidden split and approve previews	We enforce it, fail closed and audit every decision

Trust Center: trust.zetalabs.ai

Public documents

- Data Processing Agreement with Standard Contractual Clauses: viktor.com/legal
- Sub-processor list and Privacy Policy: Trust Center and viktor.com/legal

Available under NDA through the Trust Center

- SOC 2 Type I report and Type II status letter
- Penetration test executive summary

Security questions: security@viktor.com.