Keboola

ENTERPRISE SECURITY DOCUMENTATION

# Security Whitepaper

Enterprise-grade security, controls, and deployment flexibility
for regulated and large-scale data operations

### Audience

IT security and compliance teams, risk managers, data platform owners, and data engineers evaluating Keboola for enterprise environments.

### Scope

Security architecture, identity and access management, data protection, operational controls, secure development lifecycle, deployment models, and inherited cloud and warehouse security.

**About this paper:** This white paper summarizes how Keboola is designed and operated to protect customer data and support enterprise security and compliance requirements. Claims are grounded in Keboola's published documentation and public security resources, and (where applicable) mapped to independently audited controls such as SOC 2 Type II (available under NDA).

Updated January 2026

# Contents

# 1. Executive Summary

Keboola is a cloud-native data integration, orchestration, and analytics platform delivered across multiple cloud regions and providers. The platform is built around strong tenancy boundaries ("stacks"), secure-by-default data handling, and enterprise deployment options that meet the most stringent regulatory requirements.

Keboola supports multi-tenant managed environments as well as single-tenant stacks dedicated to one enterprise customer, and hybrid models where the customer brings their own data warehouse (Snowflake or BigQuery). This flexibility allows organizations to choose the deployment model that best aligns with their security policies and compliance requirements.

## 1.1 Core Security Principles

Keboola's security model is built on four foundational pillars that work together to provide comprehensive protection for enterprise data operations:

| **Defense in Depth** | **Least Privilege** |
|---|---|
| Multiple security layers from network to application | Minimal access rights by default |
| **Secure by Default** | **Continuous Monitoring** |
| Security controls enabled automatically | Real-time visibility and alerting |

## 1.2 Key Security Capabilities

| **Capability** | **Description** |
|---|---|
| Encrypted Data Storage | All data encrypted at rest using AWS KMS with AES-256 encryption and in transit via TLS 1.2+ protocols |
| Auditable Support Access | Customer-controlled approval workflows with complete audit trails for all support interventions |
| Token-Based Authorization | Granular, scoped API tokens with configurable expiration and permission boundaries |
| Secure Workspaces | Isolated execution environments with explicit input/output mapping to prevent data leakage |

# 2. Security Architecture Overview

Keboola runs as a set of services within independent deployment units called **stacks**. Each stack is an independent instance of the Keboola platform in a specific cloud provider and region, identified by its own domain (URL). This architecture provides natural isolation boundaries and enables compliance with data residency requirements.

## 2.1 Three-Tier Security Architecture

The Keboola platform uses a three-tier architecture with layered access controls:

| Tier | Components |
|---|---|
| **Users and Tools** | MFA, SSO, SSO/MFA |
| **Keboola Stack** | UI, APIs, Job execution |
| **Customer Data Warehouse** | Snowflake, BigQuery |

## 2.2 Stack Architecture and Isolation

Multi-tenant stacks operate as shared environments with logical isolation between customers, while single-tenant stacks are dedicated to one customer and use domains in the form connection.CUSTOMER_NAME.keboola.com

Each stack maintains its own:

| Component | Isolation Details |
|---|---|
| Compute Resources | Dedicated job runners and transformation engines per stack |
| Storage Layer | Isolated storage buckets with encryption keys per tenant |
| Network Boundaries | VPC isolation with controlled ingress/egress rules |

## 2.3 Deployment Models at a Glance

| Model | What Keboola Manages | What Customer Controls | Typical Fit |
|---|---|---|---|
| Fully Managed (Multi-tenant) | Platform services, compute infrastructure, storage, backups, and underlying resources | User access, project configuration, data governance policies, component selection | Fast onboarding; standard enterprise workloads |
| Multi-tenant with BYODB | Platform services, orchestration, transformation execution | Customer's Snowflake/BigQuery account, data residency, warehouse sizing | Data residency requirements; custom data hosting |
| Single-tenant | Dedicated stack with isolated infrastructure for one enterprise customer | Complete isolation boundaries, optional customer cloud environment, custom networking | Strict isolation; regulated industries; custom compliance |

## 2.4 BYODB Technical Architecture

**Bring Your Own Database (BYODB)** allows customers to maintain full control over their data warehouse while leveraging Keboola's orchestration capabilities. In this model:

- Data remains in the customer's own Snowflake or BigQuery account at all times
- Keboola connects via secure, customer-provisioned service accounts with scoped permissions
- Read-only roles can be configured for external BI connections through Data Gateway
- Customer retains full administrative control and audit logging capabilities

# 3. Data Protection and Privacy

Keboola's data protection approach is built around three core principles: (1) encryption everywhere, (2) controlled and auditable access, and (3) clear data lifecycle handling.

Keboola's Data Processing Agreement (DPA) states that Keboola implements technical and organizational measures under GDPR Article 32 and references adherence to the SOC 2 framework.

## 3.1 Encryption Standards

| Category | Encryption at Rest | Encryption in Transit |
|---|---|---|
| Standard | AWS KMS managed encryption keys with AES-256 | TLS 1.2+ for all customer-facing endpoints |
| Storage | Encrypted EBS volumes for all storage operations | Certificate management via automated renewal |
| Metadata | Encrypted metadata stores and configuration data | Internal service-to-service encryption |
| Files | Storage API files encrypted using cloud-provider capabilities | Secure connections to external data sources |

## 3.2 Secure Workspaces and Input/Output Mapping

Keboola uses a **workspace pattern** for all transformations and jobs, providing isolation between execution environments and source data. This architecture is critical for preventing accidental data exposure and ensuring data lineage.

### How Workspace Isolation Works

| Step | Phase | Description |
|---|---|---|
| 1 | Input Mapping Phase | Data is copied from storage into a temporary, isolated workspace based on explicit input mapping configuration. Only specifically declared tables and columns are available. |
| 2 | Transformation Execution | Code runs within the isolated workspace with no direct access to production storage. Temporary credentials are scoped to the specific operation. |
| 3 | Output Mapping Phase | Only explicitly defined outputs are written back to storage after successful completion. Undeclared data never leaves the workspace. |
| 4 | Workspace Cleanup | Temporary workspace is destroyed after job completion, ensuring no residual data remains in the execution environment. |

## 3.3 Data Retention and Deletion

Keboola’s DPA provides explicit lifecycle commitments aligned with privacy regulations:

| **30-Day Deletion Window** | **Customer-Initiated Deletion** |
|---|---|
| Personal data deleted or returned within 30 days after service termination (subject to applicable law) | Data can be deleted upon customer instruction at any time during the contract period |

# 4. Identity, Access Management, and Support Access

Enterprise security depends on both strong identity controls and safe operational access patterns. Keboola combines organization/project administration, role-based access, and token-based authorization to provide comprehensive identity management.

## 4.1 Authentication Methods

| Method | Description |
|---|---|
| SSO Integration | SAML 2.0 and OAuth 2.0 support for enterprise identity providers |
| Multi-Factor Auth | MFA enforcement at organization level for enhanced security |
| API Tokens | Scoped tokens for programmatic access with configurable permissions |

## 4.2 Token-Based Authorization

Keboola's APIs use tokens for authorization, enabling controlled automation while maintaining security boundaries. Tokens support:

- **Granular Permissions:** Tokens can be scoped to specific buckets, components, or operations
- **Configurable Expiration:** Time-limited tokens for temporary access scenarios
- **CI/CD Integration:** Safe storage in GitHub secrets for automated workflows
- **Audit Trail:** All token usage is logged for compliance and troubleshooting

## 4.3 Controlled Support Access

Keboola implements a customer-controlled support access model. When support engineers need to access a customer project:

| Step | Process |
|---|---|
| 1 | Support engineer submits an access request through the platform |
| 2 | Project administrators receive notification and can approve or reject the request |
| 3 | If approved, time-limited access is granted with full audit logging |
| 4 | Organizations can disable auto-join to require explicit invitations for all access |

## 4.4 Enterprise Access Control Summary

| Control Area | Enterprise-Oriented Capability |
|---|---|
| Access Boundaries | Independent stacks (multi-tenant or single-tenant) with separate endpoints, domains, and authentication contexts |
| Customer-Controlled DWH | BYODB for Snowflake/BigQuery ensures data never leaves customer's account; Keboola uses customer-provisioned service accounts |
| Least Privilege Access | Read-only roles for external BI tools, scoped tokens for automation, role-based project access |
| Support Access Governance | Explicit request/approval workflow, configurable auto-join policies, complete audit trail of support activities |

# 5. Secure Development Lifecycle and Operational Controls

Keboola provides documented mechanisms to manage configuration and lifecycle through Git-based workflows and CI/CD integration. This supports enterprise change management practices including reviewable changes, separation of environments, and complete traceability.

## 5.1 Git-Based Configuration Management

Keboola's developer documentation describes GitHub Actions workflows that synchronize project configuration between Keboola and Git repositories. This enables:

| Capability | Description |
|---|---|
| Version Control | Complete history of all configuration changes with ability to diff, review, and rollback |
| Bi-directional Sync | Pull/push/validate workflows to keep Git and Keboola in sync automatically |
| Environment Promotion | Dev/staging/prod lifecycle patterns via Git branches with controlled promotion |
| Secrets Management | GitHub secrets for tokens and credentials—no plaintext secrets in repositories |

## 5.2 Operational Transparency

Keboola maintains transparency through multiple channels for operational visibility:

| Channel | Description |
|---|---|
| Public Status Site | Real-time service health, incident history, and scheduled maintenance notifications at status.keboola.com |

### 5.3 What to Look for in Enterprise Deployments

| Area | What to Look for in Enterprise Deployments |
|---|---|
| Change Control | Git-managed configuration with mandatory code review, approval workflows, and auditable commit history |
| Environment Separation | Dev/staging/prod lifecycle patterns via Git branches; controlled promotion through pull requests and CI gates |
| Secrets Handling | GitHub secrets for tokens and credentials; encrypted variable storage; no plaintext secrets in repositories or logs |
| Service Visibility | Public status site for incident visibility; webhook notifications for status changes |

# 6. Inherited Security from Cloud and Data Warehouse Partners

Keboola deployments are anchored in leading cloud providers (AWS, Azure, and Google Cloud) and integrate tightly with modern cloud data warehouses such as Snowflake and BigQuery. This shared responsibility model means enterprise security programs can leverage the extensive compliance certifications of underlying infrastructure.

## 6.1 Shared Responsibility Model

| Cloud Provider Responsibility | Keboola Responsibility |
|---|---|
| Physical data center security | Application security |
| Network infrastructure | Identity and access controls |
| Hypervisor and host OS | Data encryption configuration |
| Hardware maintenance | Operational procedures |

## 6.2 Cloud Platform Deployments

| Platform | Description | Key Features |
|---|---|---|
| AWS | Primary infrastructure provider with multiple region options | US, EU, and APAC regions; KMS encryption; VPC isolation |
| Google Cloud | Native BigQuery integration with rapid single-tenant deployment | Private environments; BigQuery BYODB; EU data residency |
| Azure | Microsoft enterprise integration and Azure-native deployments | Azure AD integration; European regions; Enterprise support |

## 6.3 Data Gateway for Secure Warehouse Access

Keboola's **Data Gateway** component provides a more secure replacement for direct access patterns to Snowflake workspaces:

- **Key Pair Authentication:** Eliminates password-based access in favor of cryptographic keys
- **Read-Only Access:** BI tools connect through read-only roles to prevent accidental modifications
- **Connection Auditing:** All connections through the gateway are logged for compliance

# 7. Compliance and Assurance

Keboola's compliance posture is designed to support enterprise risk assessments, vendor due diligence, and regulated environments. The compliance framework includes independent audits, contractual commitments, and operational documentation.

## 7.1 Certifications and Frameworks

| Framework | Coverage |
|---|---|
| SOC 2 Type II (Annual Audit) | Independent assurance over security and availability controls over an extended audit period. The report covers: Security controls and policies; Availability and uptime commitments; Processing integrity controls; Change management procedures |
| GDPR Compliance (Data Processing Agreement) | Keboola's DPA defines comprehensive data protection commitments aligned with EU regulations: Article 32 technical measures; Confidentiality obligations; Data deletion/return terms; Sub-processor management |

## 7.2 Assurance Artifacts

| Assurance Artifact | What it Enables for Customer Programs |
|---|---|
| SOC 2 Type II Report (under NDA) | Independent audit evidence for security and availability controls; supports vendor risk assessments and regulatory examinations |
| Data Processing Agreement (public) | GDPR-aligned commitments, explicit confidentiality obligations, data subject rights support, and defined deletion/return procedures |
| Public Documentation | Comprehensive operational and architecture documentation for security reviews, implementation governance, and developer onboarding |
| Status Site | Real-time visibility into incidents, historical uptime data, and more |

# 8. Enterprise Readiness: Typical Requirements Mapping

Security, compliance, and platform teams commonly evaluate data platforms against a consistent set of controls. The following comprehensive mapping shows how Keboola addresses typical enterprise security requirements across multiple control domains.

| Requirement | How Keboola Addresses It | Evidence Sources |
|---|---|---|
| Isolation and Residency | Independent stacks per region/cloud; single-tenant stacks available; multi-cloud deployment options across AWS, Azure, and GCP | Keboola deployment docs; BYODB documentation; stack architecture guides; regional availability matrix |
| Encryption | AES-256 encryption at rest via AWS KMS; TLS 1.2+ for all data in transit; encrypted metadata stores and configuration | Security improvements changelog; storage documentation; TLS configuration guides |
| Access Governance | Approval-based support access with audit trail; scoped API tokens; RBAC with project/bucket granularity; SSO/MFA support | Support access documentation; token management guides; BYODB access patterns |
| Auditability | Git-based configuration with full history; operational status transparency; API logging; support access audit trail | Developer docs; GitHub Actions integration; status API documentation |
| Privacy Program Alignment | DPA with GDPR Article 32 measures; defined data deletion procedures; sub-processor list; SOC 2 framework reference | Public DPA document; privacy documentation; sub-processor register |

## 8.1 Getting Started with Security Assessment

For organizations conducting vendor security assessments, Keboola recommends the following approach:

1. Review this white paper and public documentation for architecture understanding
2. Request the SOC 2 Type II report through your Keboola account team (available under NDA)
3. Complete the standard vendor security questionnaire provided by Keboola

> 4. Schedule a security architecture review call for specific technical questions

# 9. References

The following resources provide additional detail on Keboola's security architecture and compliance posture:

## 9.1 Documentation Resources

- **Keboola User Documentation:** Overview and deployment options, stacks and terminology (help.keboola.com)
- **Keboola Developers Documentation:** API endpoints and stacks; CI/GitHub integration workflows (developers.keboola.com)

## 9.2 Legal and Compliance

- **Data Processing Agreement:** GDPR Article 32 measures, SOC 2 framework reference, deletion/return terms (keboola.com/dpa)
- **SOC 2 Type II Report:** Available under NDA via Keboola Trust/Security process

## 9.3 Operational Resources

- **Keboola Status:** Security improvements notice and public status (status.keboola.com)
- **Keboola Changelog:** Data Gateway component announcement and feature updates (changelog.keboola.co

> **Note:** Some assurance artifacts (e.g., SOC 2 Type II report) are shared under NDA. This white paper is designed as a public summary and does not reproduce confidential audit report content. Contact your Keboola account team to request access to restricted materials.

Enterprise-grade security for your data operations

Last Updated: January 2026