# data.world

## Security Overview

April 2021

# Security Overview

## Your data is very important to you. We take your trust seriously.

**data.world is not just cloud-first, but also security-first.**
We've designed our service from the ground up to ensure that we can support your unique combination of internal and external compliance needs. Part of the responsibility of a fully managed SaaS provider is not only to minimize the total cost-of-ownership of your data platform investment, but also to ensure a superior security and compliance posture for your organization.

We regularly audit our configurations against the latest CIS Benchmark for AWS, created by a large community of cybersecurity experts. We also engage a third-party security firm to perform annually a detailed code-level analysis and penetration testing.

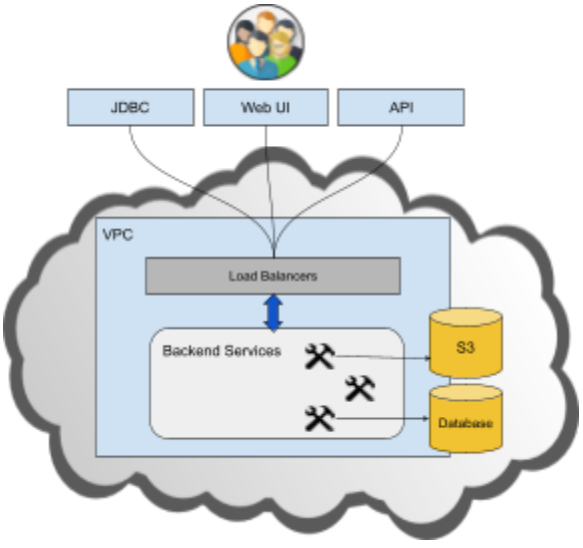
Our comprehensive security program applies current tools and practices to network connectivity, access control, data handling, and incident management.

## Network Security

data.world is hosted on Amazon Web Services in the us-east-1 region across three redundant data centers. All access to data.world via the Internet is handled using secure protocols, protected by TLS 1.2+ encryption. All data transfers to and from data.world use one of these secure endpoints.

We utilize the Virtual Private Cloud (VPC) service to create a private network space that is designed to protect internal resources. AWS Security Groups and Network ACL are configured to provide only the minimal necessary connectivity required by the application.

Additionally, data.world employs a Content Distribution Network (CDN) that provides improved global application performance and protection from some classes of network threats such as Distributed Denial of Service (DDoS) attacks.




## Access Control

### User Authentication

data.world account

User accounts on data.world are stored in our database with hashed and salted passwords. Users can create accounts manually, or through

OAuth providers such as Google, Facebook, Github and Twitter.

Single Sign On (SSO)

An organization on data.world may choose to require authentication against an existing directory of approved employees, often referred to as Single Sign On (SSO). data.world supports any SAML2 enabled directory such as Active Directory, Okta, OneLogin, JumpCloud, and more. Once configured, access to data.world requires both local account authentication (username/password) and a successful SAML assertion from an identity provider.

API Authentication

Authentication to the API is made with a generated token specific to a user account. Accounts that require SAML are able to generate tokens. Existing API tokens are invalidated if the SAML configuration changes for the account.

### User Authorization

data.world provides a multi-level role-based access control (RBAC). Users can be granted job specific roles within the site, such as Viewer, Contributor or Admin. These roles can be granted at different levels of granularity - dataset, project or entire organization - to individual users or to other organizations.

### System and Admin Authentication

All backend authentication at data.world is configured within AWS Identity and Access Management (IAM). This includes employee access to systems, system to system authentication, and authentication between tiers of the application - such as between a web service and a database. The IAM profiles - for employees and systems - are configured using the principle of least-privilege, meaning the minimum set and scope of permissions are granted for the use case. Additionally, any keys granted for employee access are temporary and expire within an hour of issuance.

## Data Handling

### Encryption Everywhere

All data transmitted over the Internet is encrypted with TLS 1.2+. Additionally, all customer data stored on disk (at rest) is encrypted. This includes files uploaded by customers, our application database, search indexes, and any locally cached customer data.

### Private data.world

data.world is a multi-tenant service that provides secure isolation at several levels. Customer data is stored in separate directories and access can be managed by users and organization administrators. However some customers require additional levels of isolation due to the sensitivity of their data or for compliance requirements. For such customers, data.world offers a single-tenant instance in a dedicated AWS account. Under this option no sharing of any cloud resources, databases, or encryption keys occurs.

## Incident Detection and Management

### Security Information and Event Monitoring (SIEM)

data.world continuously gathers information from the AWS control plane (CloudTrail), our server system logs, and application logs. It then inspects them for suspicious activity. All events that are identified as security-relevant are automatically prioritized, and high risk events are reviewed immediately. All security events are stored indefinitely for future forensic review.

### File Integrity Monitoring (FIM)

As part of our SIEM solution, we closely monitor critical system files and customer data files for modification or permission changes. Any such change triggers a high severity escalation and investigation.

### Physical Security

We are hosted in Amazon Web Services in data centers that are certified ISO 27001. AWS employs top-tier physical security controls in their data centers; full details are available here: https://aws.amazon.com/compliance/data-center/controls/

Storage media is decommissioned following the guidelines outlined in NIST 800-88

## Assurance Programs

data.world works with certified third-party auditors to validate our security program according to these assurance programs:

- SOC 2, Type II: data.world has completed attestation and audit for 2021.
- HIPAA: data.world has obtained an affirmative HIPAA opinion. Contact legal@data.world to obtain our form BAA