Automatically Discover And Surface Data In Your Data Lake, Reducing Time To Analysis

You have thousands of growing datasets with millions of distinct data fields stored in Microsoft Azure. Manually discovering and documenting the data isn’t an option.

Waterline’s AI-driven Data Catalog automates data discovery and tagging of your data leveraging AI and machine learning at petabyte scale and velocity. With Waterline’s AI-driven Data Catalog, business users can easily find data; compliance teams can ensure data is secured, and IT teams can rationalize the data to identify redundancies as well as security and privacy risks.

Speed Data Discovery
Less time spent trying to find data means more time for actual analytics and gathering insights.

Reduce Risk of Fines
Respond quickly to evolving data regulations, and ensure you are always in compliance using Waterline Data’s patented AI technology that auto-tags every data set at petabyte scale.

Avoid a Data Swamp
Redundant or overlapping data is easily identified across Azure Data Lake Store and Azure Blob Storage to prevent clutter from accumulating and minimize data storage costs.

Close Security Gaps
Automatically identify and secure sensitive data using your existing data security solutions.
Harness The Power Of Your Data On Microsoft Azure

Simplify data curation, governance and compliance using Waterline AI-Driven Data Catalog

**BUILT FOR THE ENTERPRISE**

Waterline’s AI-driven data catalog is a multi-purpose solution, unlike embedded data catalogs that service only a single-purpose data application.

- **Simplify your IT infrastructure** with one solution that powers multiple use-cases, including self-service analytics, data compliance and security, and data rationalization.

- **Accelerate digital transformation** using Microsoft Azure while eliminating duplicate or overlapping data sets, so your Azure Data Lake Store and Azure Blob Storage remain uncluttered.

- **Perform at petabyte scale** with a big data technology foundation that leverages your Hadoop stack on Azure, using HDInsight with Apache Spark and Apache Solr for compute scalability.

- **Increase efficiencies** by integrating with existing business processes across your software ecosystem, including data analytics, security, data management and embedded catalogs.

- **Improves data accuracy** using a combination of artificial intelligence, machine learning and best-in-class crowdsourcing.

> Traditionally, compliance is seen as an overhead, but with Waterline Data we see compliance as an opportunity to ensure we’re doing our job really, really, well.

*Alasdair Anderson*

**EVP and CDO**

**Nordea Bank**

**ENTERPRISE GRADE**

**Built for Scale**

Built on big data technologies using Apache Spark and Apache Solr for compute scalability.

**Enterprise Wide**

Supports all data platforms—on-premise and cloud including Microsoft Azure. Elastic scalability to support variable workloads.

**Secure and Compliant**

Avoid costly fines by allowing those with access rights to view and use sensitive data. All users are authenticated with Azure Active Directory.

CONTACT US TODAY TO SCHEDULE A DEMO

REQUEST A WATERLINE DATA SANDBOX FOR AZURE TODAY

---

**Power BI**  
**Paxata**  
**TRIFACTA**

**Data Security**

Azure Active Directory

Apache Ranger

**Apache Spark On Azure HDInsight**

**Azure Storage**