Accelerate data preparation and speed time-to-insight with Paxata on Microsoft Azure

Businesses are migrating growing volumes of data to big data stores to support digital transformation and analytical initiatives. Data lakes are full of data collected without a clear sense of how it will be used.

To gather insights from the data, however, it must be clean, complete, and contextual. About 80 percent of analytics effort and machine learning modeling is spent bringing together data, and cleansing, shaping, and preparing data with specialized IT resources. Getting data ready for analysis is tedious, time-consuming and expensive. As a result, many organizations never realize the value of all the data they have collected.

**PAXATA SELF-SERVICE DATA PREPARATION SOLUTION ON MICROSOFT AZURE**

With Paxata self-service data preparation running on Microsoft Azure, businesses can empower their business analysts and subject matter experts—those most familiar with the context of data—to interact and prepare data without IT support. They can work with data from any source, including data from Azure Data Lake Store and Azure Blob Storage. They can publish it anywhere, including Microsoft Power BI, in hours, versus the weeks or months it would take with a traditional process.

Paxata runs natively on your Hadoop stack on Azure. Use it with HDInsight—or on Cloudera on Azure. Paxata offers powerful algorithmic intelligence, an intuitive user experience, and a smart distributed architecture that supports large enterprise data volumes.

Paxata is the pioneer that intelligently empowers all business consumers and citizen data scientists to find, ingest, profile, clean, and transform raw data instantly and automatically, with an enterprise-grade, self-service data preparation application. Our Adaptive Information Platform weaves data into an information fabric from any source and any cloud to create trusted insights. Business consumers use clicks, not code to achieve results in minutes, not months.

Paxata, a DataRobot company, has offices in Redwood City, California, New York, Ohio, Texas, and Singapore.

Microsoft enables digital transformation for the era of an intelligent cloud and an intelligent edge. Microsoft Azure is an open, flexible, enterprise-grade cloud computing platform. Wherever your data lives, Azure services will help you unlock its potential. Services include HDInsight, a fully managed cloud Apache™ Hadoop and Apache™ Spark™ service backed by a 99.9% SLA.
FASTER INSIGHTS FOR NON-TECHNICAL USERS

Minimize analytics and machine learning bottlenecks and gain insights from data faster. With the Paxata Self-Service Data Preparation application, any business user can turn raw data into trustworthy information with clicks, not code.

- Excel-like user interface requires minimal training, presenting a visual and interactive experience.
- Smart suggestions help users detect joins, overlaps, and misspellings, so you can standardize and unify data very quickly regardless of volume, variety, and source.
- Algorithmic intelligence works behind the scenes to accelerate ad hoc, operational, and predictive analytics.

MAKING SENSE OF BIG DATA

If you’re running Apache Spark™ on HDInsight or Cloudera on Azure, you are already processing massive amounts of data. And while it’s relatively simple to get data into a Data Lake Store or other Blob storage, you probably know that it’s difficult to get it out and make sense of it—until Paxata.

- Land data without concern for schema and use Paxata to make sense of it when you are ready. As business users pull data, the intelligent ingest capability detects file type, including whether it is a semi-structured, nested file or an electronic document, and automatically flattens the file into a tabular format.
- Support massive data volumes using Azure Data Lake Store and Azure HDInsight for enterprise-scale processing.

GOVERNANCE WITH SELF-SERVICE

Paxata governance controls enable you to control access to the system and deliver traceability.

- Paxata self-documents data preparation activities, providing auditability, repeatability, and traceability for all data transformations, and preserves data lineage.
- Paxata also provides secure authentication, authorization, encryption, and usage tracking.
- Paxata’s centralized and governed data catalog enables business analysts to search and find certified answer sets or reuse data preparation projects with new data sets.

“Azure customers can now leverage Paxata’s governed, self-service data preparation solution to leverage the intelligence, hybrid and enterprise-grade capabilities of Microsoft Azure – harnessing the full power of Azure’s Big Data offerings with a single click on HDInsight.”

Jason Zander, Corporate Vice President, Microsoft Azure

ENTERPRISE GRADE


Built for the cloud: Elastic scalability to support variable workloads. Deploy from the Azure Marketplace in one click, without opening any ports, and run using native Azure services.

Secure and compliant: All requests are routed through the secure gateway on HDInsight and users also are authenticated with Azure Active Directory.
WHY PAXATA

Faster time to value

A 150-year old, multi-national bank, operating over 70 different markets, was able to accelerate reporting on financial crimes and compliance and reporting, reducing processing time from 21 days to 1, while improving the quality of reporting by 52%.

Cost efficiencies in data cleanup

A large telecom company migrated its ERP in 3 months, saving approximate $13 million in staffing and consulting costs. Paxata’s automatic validating and one-click profiling processes enabled the company to clean the data as they migrated it.

More value from data

A healthcare company used Paxata to reduce the process to uncover insurance fraud from 300 days to 90, with the help of Paxata’s interactive data exploration and unique filtering capabilities to find anomalies and outliers.
PAXATA IS CENTRAL TO YOUR MODERN DATA ESTATE

Paxata as part of an Azure solution architecture. Business users can pull data from any source in any format, and prepare data for analysis and consumption using the Hadoop stack of their choice.

GET STARTED TODAY

With Paxata, you're just a quick install away from accelerating your time to insight in a secure, scalable, data preparation environment running on Microsoft Azure.

Deploy with one click from Azure Marketplace

TRY PAXATA FREE FOR 30 DAYS: www.paxata.com/get-started
VISIT PAXATA.COM/ALLIANCES/MICROSOFT