

### Benefits

- Extract new value from core systems data
- Feed enterprise data warehouses and data lakes
- Support downstream system integrations

### Features

- Streaming access to all raw (uncurated)
- Guidewire InsuranceSuite data
- Secure access with intra-daily changes at a low latency
- Data available in easy-to-understand and use parquet file format
- Support for schema evolution
- Part of the Guidewire Cloud subscription

# Guidewire Cloud Data Access

Augment your enterprise data strategy with Guidewire Cloud Data Access.

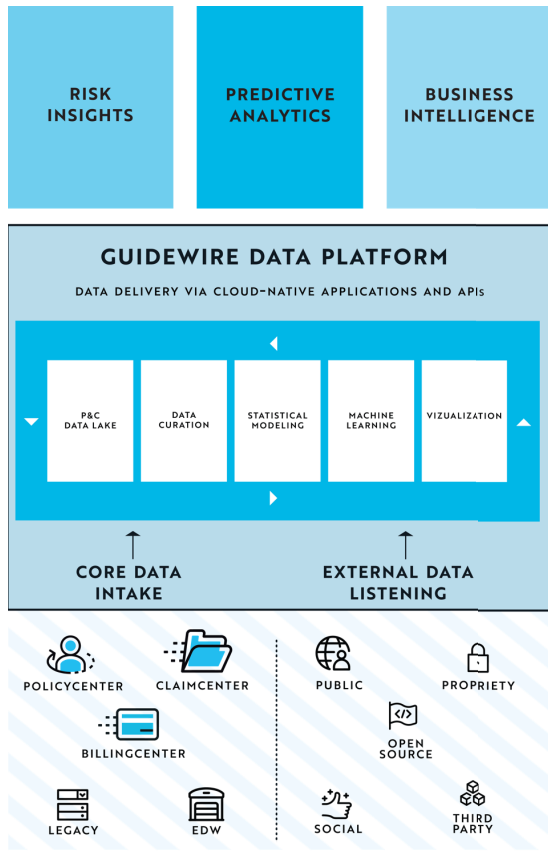
**Guidewire Cloud Data Access is a data service that gives Guidewire Cloud customers secure access to all raw (uncurated) Guidewire InsuranceSuite data, including all intra-daily changes at a low latency.**

## Unlock Unlimited Potential with Data

If you're an insurer, access to transactional core systems data is absolutely critical to augment and support your enterprise-wide data strategy. This is even more critical for core systems deployed in the cloud. Modern core systems generate a lot of data and must be able to share this data with downstream systems efficiently.

In the past, customers created database replicas, wrote custom queries, and performed complicated extract, transform, load (ETL) operations. This is often a very costly endeavor for most insurers. Business users today demand secure access to core systems data in near-real time and with all intra-daily changes. Looking at the status of a claim or a submission at the end of each day is no longer sufficient. In addition, data is needed for custom operational and financial reporting, for feeding enterprise data warehouses and data lakes, and for supporting downstream system integrations.

Guidewire Cloud Data Access is a data service that is enabled by Guidewire Data Platform and that provides secure access to all raw (uncurated) Guidewire InsuranceSuite data, including all intra-daily changes at a low latency. This service is included with the Guidewire Cloud subscription.



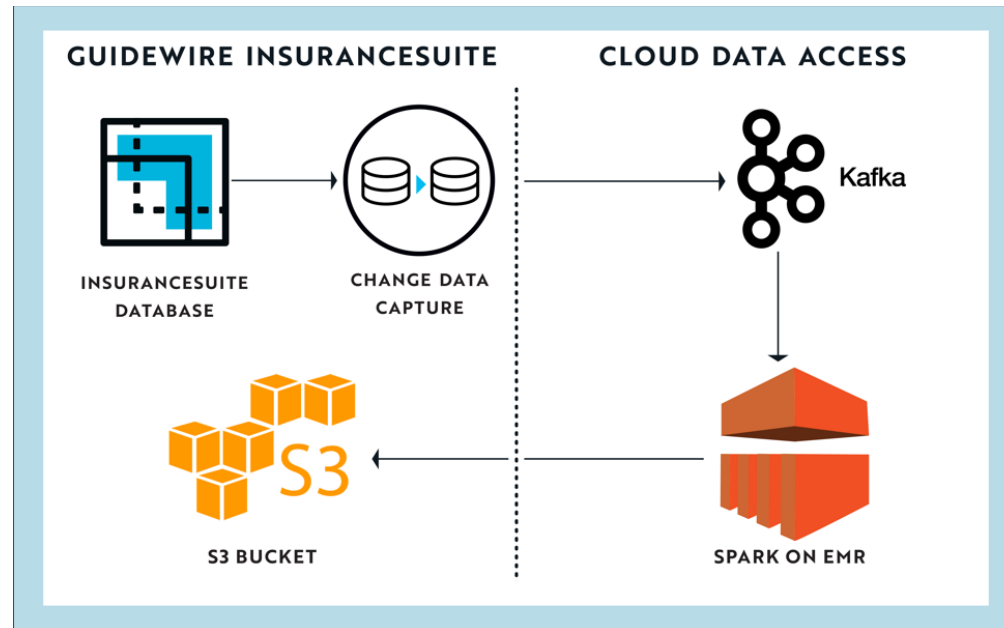
**Guidewire Data Platform**

### Extract New Value from Core Systems Data

Cloud Data Access is hosted in Guidewire Cloud on the Amazon Web Services (AWS) platform. Each instance communicates with a single InsuranceSuite database that is also deployed in Guidewire Cloud. Cloud Data Access executes an initial bulk load to capture existing historical data in the InsuranceSuite database and loads this data into Amazon S3. After this initial load, it captures incremental changes.

Data extraction and streaming occur in near-real time by enabling change data capture (CDC) against the source database tables. CDC is a common database feature that captures information about all database change activity. Cloud Data Access supports Microsoft SQL Server, Oracle Database, and PostgreSQL.

Guidewire Data Platform uses Apache Kafka to build a real-time data pipeline. It streams data (initial and incremental) into a Kafka topic. Additionally, schema changes are automatically identified and marked with metadata as the data streams into Kafka. Subsequently, Cloud Data Access automatically accounts for these changes.



**Guidewire Cloud Data Access**

## About Guidewire Data Platform

The Guidewire Data Platform is a P&C insurance-specific data repository and factory that continuously collects data from internal and external sources and provides analytical insights across the insurance lifecycle.

Cloud Data Access then executes an Apache Spark job on an Amazon EMR cluster to in an Amazon S3 bucket. Cloud Data Access guarantees that there are no duplicate records. It also supports additive, nondestructive changes to schema by identifying the schema change in the source database and assigning a unique schema fingerprint.

For each connector to an InsuranceSuite database, Cloud Data Access creates an S3 bucket with AES256 encryption. Cloud Data Access writes extracted data into S3 in the Apache Parquet format. The Parquet format is a machine-readable binary format that efficiently represents complex nested data structures in a column-oriented way that enables high performance in storage and processing.

For each table in the InsuranceSuite database, Cloud Data Access creates a table-specific folder in S3. In each table's folder, it groups data by the schema's fingerprint. In each fingerprint folder, it further groups data according to the time that it completed writing data into S3. For each incremental fetch, Cloud Data Access writes newly extracted data into new folders that are named with the time stamp when it completed writing this new data into S3.

Cloud Data Access also creates and maintains a manifest.json file at the same level as table-specific folders. This file contains metadata about the replication process to understand the S3 content that can be safely consumed. The manifest is continuously updated to reflect the successful delivery of new incremental data into an S3 folder.

After Cloud Data Access has successfully written data into an S3 folder, customers can download this data into their self-managed environments to support any downstream data integration use cases, such as reporting and data warehousing. Customers can use their own application to download the data or can leverage the Guidewire Cloud Data Access Client reference implementation, which has been made available as an open source GitHub repository.

Guidewire configures and manages the data pipeline, and customers can access only the contents of their own S3 bucket. Cloud Data Access secures the S3 bucket with Amazon Identity and Access Management (IAM) roles and bucket policies along with Amazon Key Management Service (KMS) to protect your data.

Guidewire Cloud Data Access gathers InsuranceSuite data at the right scale and tempo to best position insurers in today's competitive environment. It provides secure data access at a low latency to augment your enterprise data strategy and extract new value.

**Guidewire is the platform P&C insurers trust to engage, innovate, and grow efficiently. We combine digital, core, analytics, and machine learning to deliver our platform as a cloud service. More than 500 insurers, from new ventures to the largest and most complex in the world, run on Guidewire. For more information, contact us at [info@guidewire.com](mailto:info@guidewire.com).**