



# The Solix Data Fabric: **A Multi-cloud Information Architecture**



**John Ottman,**  
Executive Chairman, Solix Technologies, Inc.



*As originally published by the Forbes Technology Council*

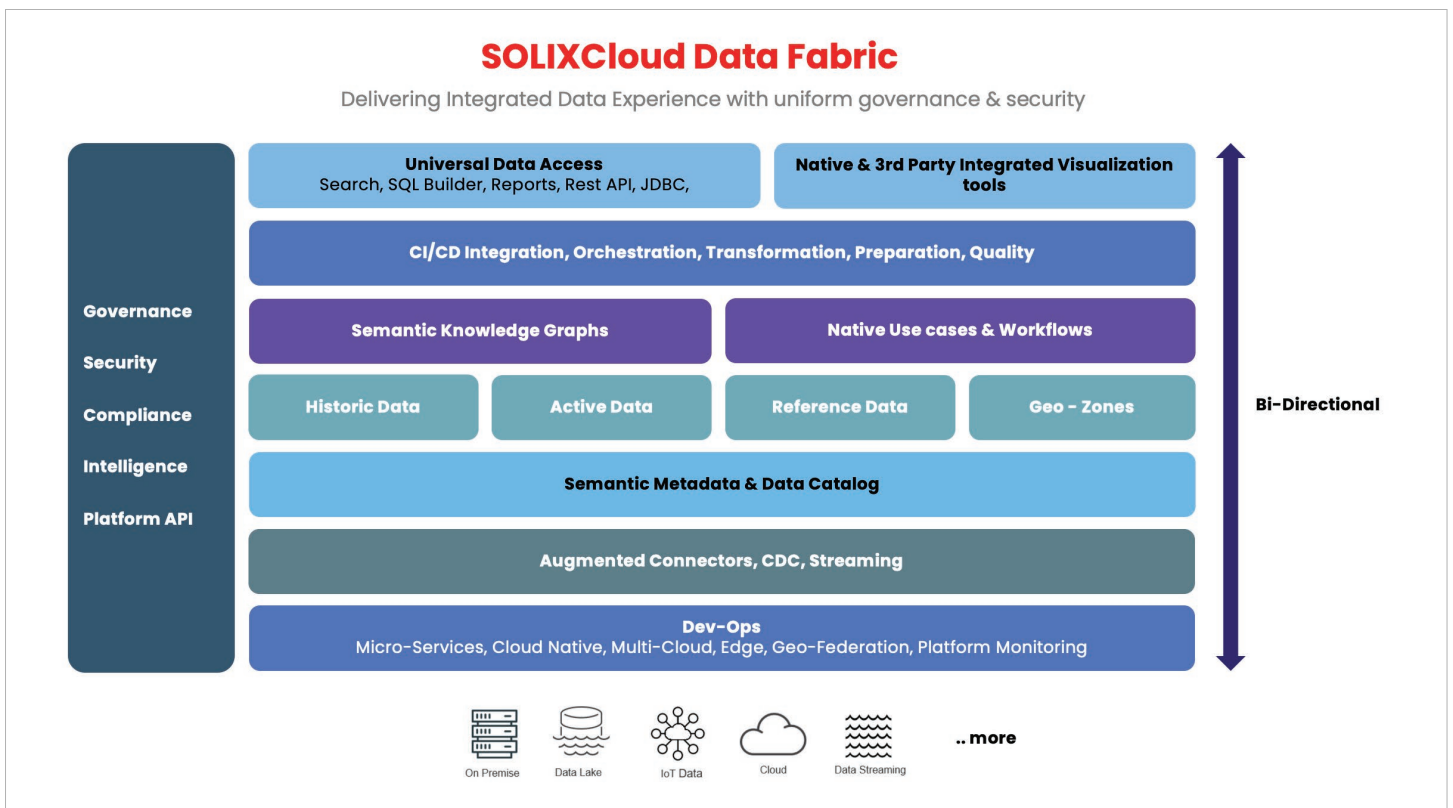
*The Data Fabric: Cloud Data Management Essentials You  
Need To Know*

# The Solix Data Fabric:

## A Multi-cloud Information Architecture

Digital transformation is driving organizations as fast as they can to implement data-driven solutions that transform business results such as advanced analytics, artificial intelligence and machine learning. But achieving digital transformation isn't so easy for many organizations because a new information architecture is required called the data fabric. The data fabric is a trusted, multi-cloud information architecture that automates data management functions end to end.

Not only is the data fabric based on open standards, but its design also scales horizontally across virtualized, commodity infrastructure components. Large-scale data consumption can drive up data storage requirements, but object storage served in S3 buckets not only provides acceptable performance for most data-driven workloads, but it also makes an ideal solution for low-cost, bulk data storage.



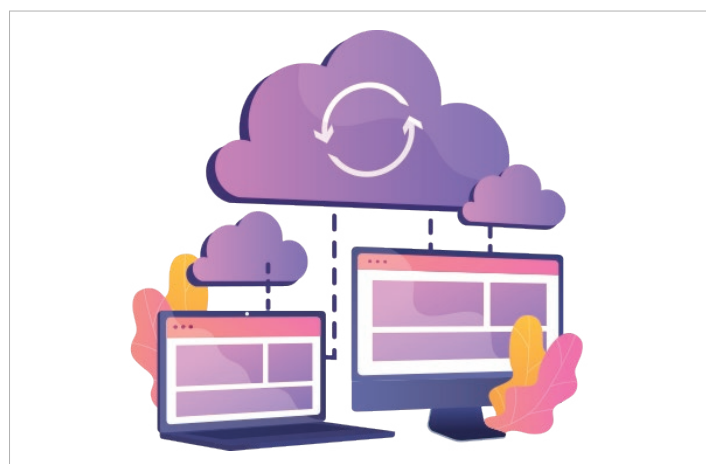
New infrastructure requirements are required to support the data fabric. Real-time, event-driven processing across multi-cloud environments is the goal, and the data fabric relies on cloud native, W3C standard software for end-to-end integration of microservices, containerization and orchestration.

Cloud data management is an application framework that delivers four essential services to the data fabric.

## The Solix Data Fabric: A Multi-cloud Information Architecture

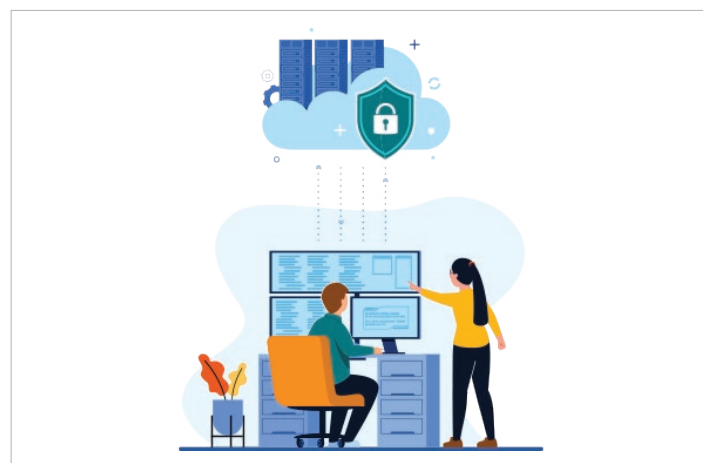
### Connect

The data-driven enterprise relies on data pipelines to feed downstream applications, and cloud data management systems are responsible for collecting and moving data from any source to any target. Source data may be streaming or reside as files located anywhere, and very often, data is sourced from enterprise applications, mainframes, factory floors and other SaaS systems. The ability to connect to any data source and not only move large volumes of data but also transform and prepare the data for use is a core competency for the data-driven enterprise.



### Govern

With so much data under management, proper data governance is essential to ensure data security, compliance and privacy. Information lifecycle management (ILM) classifies and manages data based on policies and business rules throughout the entire data lifecycle, from collection and ingestion to data retention and, ultimately, to shredding.



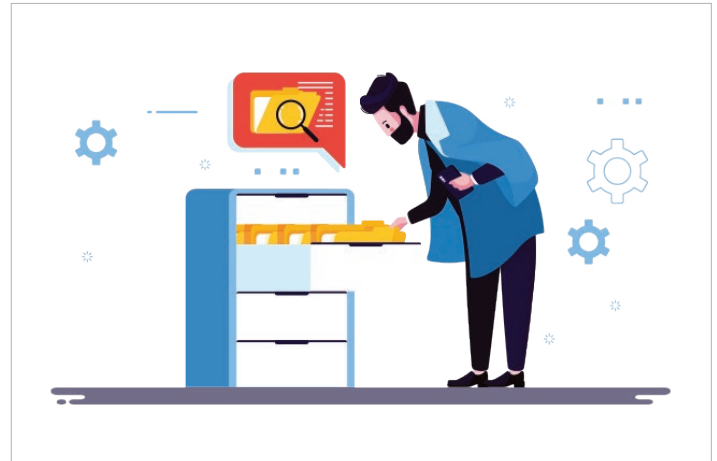
# The Solix Data Fabric: A Multi-cloud Information Architecture

## Discover

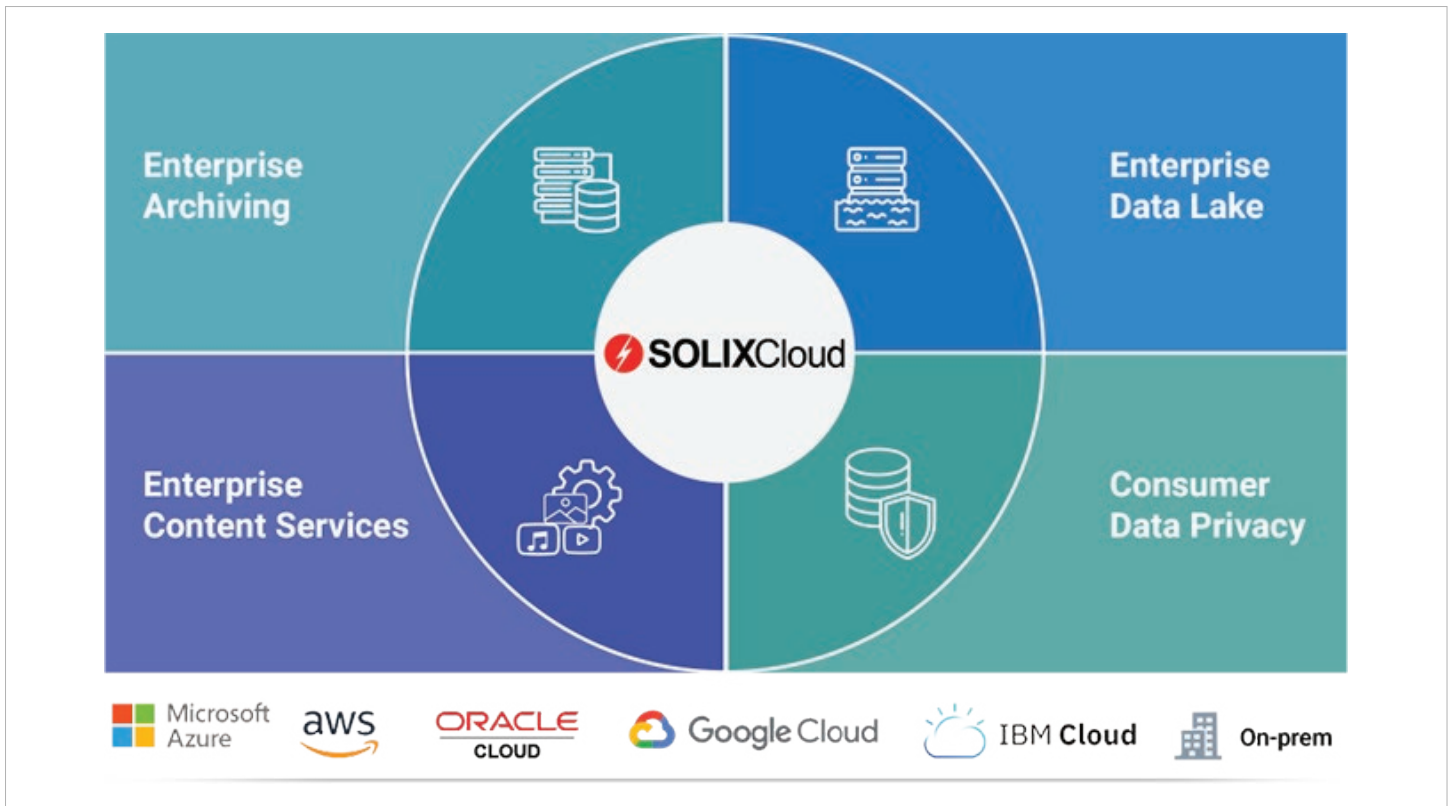
In order to explore the entire data landscape, data access must be available not only through API but through text search, ad hoc queries and structured reports. Role-based access controls granting the "least privilege necessary" ensure universal access to the data without compromising data privacy, security and compliance.

## The Cloud Data Management

Cloud data management runs either as a private cloud solution managed by in-house resources or as a SaaS subscription. Once data is captured and ingested into a common data platform, cloud data management applications provide tooling to pipeline, enrich,



transform and prepare the data for later use. IT managers gain increased control because common data platforms enable security, data governance and access control to be managed from a single point.





## The Solix Data Fabric: A Multi-cloud Information Architecture

---

The top three cloud data management applications for the data fabric are data lakes, archives and consumer data privacy.

### Top three Cloud Data Management Applications for the Data Fabric

**Data Lakes** are repositories of “current” data collected to feed data-driven applications. Advanced analytics, machine learning and artificial intelligence applications often require some data transformation, data cleansing or data enrichment in preparation for processing. Data lakes are the building blocks of data pipelines, and they provide a data preparation platform to make data “fit for use” by data-driven applications.

**Archives** are centralized repositories of “historical” data that improve data governance and compliance through the principle of data minimization, which states simply that data should be retained only for as long as minimally necessary, and only to fulfill a specific and defined purpose. By archiving historical data from production databases, file servers and email servers, organizations can improve the performance of their production applications, optimize infrastructure and reduce the high cost of production infrastructure because the data footprint is so much smaller.

Archives also enable application retirement and decommissioning of legacy systems that are no longer in use. Significant cost savings may be realized, and improved data compliance is achieved by rationalizing application portfolios and retiring outdated applications and data consistent with ILM and compliance policies.

**Consumer Data Privacy** regulations are emerging worldwide, sparked by the broad adoption of the EU’s GDPR initiative. New privacy laws and regulations are now enacted by over 100 countries and states, and

significant fines have been levied for non-compliance. Consumer data privacy applications establish “privacy by design” and equip data controllers with critical tools such as sensitive data discovery, data masking, legal hold, redaction and erasure in order to comply with regulations governing personally identifiable information (PII).

### The Conclusion

The data fabric is strategic to digital transformation, but moving large amounts of data from source to target also introduces a full stack of complexity. Proper planning over scope, objectives and approach eliminates most problems, but advanced SQL processing, machine learning and artificial intelligence consume large amounts of CPU power and network bandwidth, and when the data is not transformed or prepared properly, it may not produce the desired results. So plan carefully, and make sure your team is skilled in cloud native software, NOSQL, dev ops and the relevant cloud provider console.

Digital transformation requires a new information architecture called the data fabric to support data-driven applications, and cloud data management is the essential application framework that provides tools and use cases to collect, store and prepare data pipelines. Get started now and propel your business as a data-driven enterprise.

**The Solix Data Fabric:**  
A Multi-cloud Information Architecture

---



**Contact Us**

For more information contact us at:

**Solix Technologies, Inc.**

4701 Patrick Henry Dr., Bldg 20  
Santa Clara, CA 95054

**Toll Free:** +1.888.GO.SOLIX (+1.888.467.6549)

**Telephone:** +1.408.654.6400

**Fax:** +1.408.562.0048

**Email:** [info@solix.com](mailto:info@solix.com)

**URL:** <http://www.solix.com>

Copyright ©2022, Solix Technologies and/or its affiliates. All rights reserved.

Solix Technologies, Inc., is a leading big data application provider that empowers data-driven enterprises with optimized infrastructure, data security and advanced analytics by achieving Information Lifecycle Management (ILM) goals. Solix Common Data Platform offers an Information Lifecycle Management framework for Enterprise Archiving and Data Lake applications with Apache Hadoop as an enterprise data repository. The Solix Enterprise Data Management Suite (Solix EDMS) enables organizations to implement Database Archiving, Test Data Management (Data Subsetting), Data Masking and Application Retirement across all enterprise data. Solix Technologies, Inc. is headquartered in Santa Clara, California and operates worldwide through an established network of value added resellers (VARs) and systems integrators.