

# Fast-track your data lake transformation with AWS Data Lake Accelerators

Businesses have always relied on data to solve complex problems and drive growth. What's different now is the unfathomable explosion in the amount of data that's being generated every day. According to IDC, the quantity of data companies will manage is expected to grow by 400% between 2020 and 2025.

With the right planning, design and governance, an AWS data lake can help you process data in any format, and at petabyte scale. We'll help you imagine the possibilities, identify use cases and accelerate time-to-value with reference architectures and accelerators. And we can deliver a proof-of-concept in just three weeks.

## Business challenges

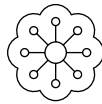
Hadoop based on-premises data lake systems often struggle to deliver value with storage and data co-located and difficult to integrate with other clouds and new technologies

Organizations want high-performing platforms that can easily and cost-effectively scale to support modern use cases

Maintenance of existing data lake architectures consumes IT resources, capacity planning and expansion require time

Facing compliance and security challenges

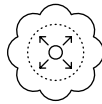
## AWS data lake



Unify your data architecture, including hybrid or multicloud options, to achieve a complete view of your business with a seamless user experience



Handle data of any scale to unlock modern analytics use cases, reduce TCO and move to an operational cost model



Unlimited scale for your data lake with governance and ecosystem expansion to support machine learning and analytics workloads



Easily maintain strong cutting-edge security with fine-grained identity and access management controls

## Use Case:

### Consumer Goods Enterprise

#### Challenge

Our customer wanted its smart toothbrush to identify oral health conditions and report its status to users, but this customer was facing challenges managing rapidly growing amounts of data from varied sources.

#### Solution

Onica by Rackspace Technology™ built a data lake to ingest, store and process imagery and other data. This provided flexibility to run different types of analytics on the data, such as training a neural network to identify oral health conditions based on the toothbrush camera.

#### Results

Onica® was able to help this customer:

- Scale its data infrastructure to allow data collection
- from many users
- Train a neural network
- that identifies oral health conditions for users
- Deliver a sophisticated upgrade to its smart toothbrush



Amazon Simple Storage Service (Amazon S3)



AWS Key Management Service (AWS KMS)



AWS Identity and Access Management (IAM)



Amazon Kinesis Video Streams



Amazon Kinesis Data Firehose



AWS Glue



Amazon Athena



Amazon SageMaker



Amazon QuickSight



Amazon CloudWatch



AWS CloudTrail



Amazon GuardDuty

Onica by Rackspace Technology provides data architecture and engineering services that accelerate your path to leveraging a highly scalable, flexible and agile AWS data lake.

### AWS Mini Data Lake

A fully operational, non-production data lake that ingests one data source and applicable data pipeline.

### AWS Serverless Data Lake

A full production and staging environment leveraging the AWS Serverless Data Lake Framework (SDLF).

Duration	3 weeks	6 weeks
<b>Use case</b>	<ul style="list-style-type: none"> <li>• Proof of concept</li> <li>• Environment for testing and experimentation</li> <li>• Security and governance not a priority</li> </ul>	<ul style="list-style-type: none"> <li>• Analyze large amounts of data from multiple sources</li> <li>• Support multiple users and user groups</li> <li>• Compliance requires data to be auditable</li> </ul>
<b>Features</b>	<ul style="list-style-type: none"> <li>• Supports single data source</li> <li>• Basic 2 zone design with notional data transformation</li> <li>• Data secured with KMS throughout</li> <li>• Access governed by AWS Lake Formation</li> <li>• Two user groups to show segregated data and resource access</li> <li>• Illustrative QuickSight dashboards</li> <li>• Basic CI/CD pipeline</li> </ul>	<ul style="list-style-type: none"> <li>• Traceability and version control</li> <li>• Scalability and reproducibility</li> <li>• AWS best practices enforced within the framework</li> <li>• Data quality validation</li> <li>• Pre-built and tested Terraform templates</li> <li>• Increased automation of repeatable tasks</li> <li>• Augmented modularity and flexibility, less hard coding and reduced coupling between components</li> <li>• Illustrative QuickSight dashboards</li> <li>• Multi-environment CI/CD pipelines</li> </ul>
<b>Deliverables</b>	<ul style="list-style-type: none"> <li>• Functional pilot environment</li> <li>• Solution documentation</li> <li>• Architectural diagrams</li> <li>• Data pipeline</li> <li>• Transformation code</li> <li>• Configurations and specs</li> </ul>	<ul style="list-style-type: none"> <li>• Production and staging environment</li> <li>• Detailed design document</li> <li>• Architectural diagrams</li> <li>• Automation code</li> <li>• Data pipeline</li> <li>• Transformation code</li> <li>• Schema, SQL, and data catalog</li> <li>• Runbooks</li> </ul>
<b>Technologies</b>	S3, Lambda, Glue, Athena, CodeCommit, CodePipeline, QuickSight, IAM and KMS	S3, DynamoDB, Lambda, Glue, CloudTrail, SQS, CodeCommit, CodePipeline, Athena, CloudWatch, IAM, KMS and QuickSight

### About Onica by Rackspace Technology

- 2,500+ AWS accreditations
- AWS Service Delivery Partner for Database Migration Services, RDS, DynamoDB and Redshift
- AWS 2021 Migration Partner of the Year (U.S. and Canada)
- 13 AWS Service Delivery Designations
- Managed Service Provider

### Expertise Across 15 AWS Competencies

- Data & Analytics
- DevOps
- Education
- Financial Services
- Healthcare
- Industrial Software
- IoT
- Machine Learning
- Microsoft Workloads
- Migration
- Oracle
- Retail
- SaaS
- Storage
- Travel & Hospitality



### Take the Next Step

Learn more about how Onica can fast-track your data transformation. And register for your free data discovery session.

**Call:** 1-800-961-2888

**Email:** [onica@rackspace.com](mailto:onica@rackspace.com)

**Visit:** [www.rackspace.com/lp/data-discovery-session](http://www.rackspace.com/lp/data-discovery-session)