

AWS Data Lake Accelerators

Fast track your data lake transformation journey.

Businesses have always leveraged data to solve complex problems and drive growth. What's different now is its unfathomable explosion. According to IDC, data is expected to grow by 400% from 2020 to 2025. With the right planning, design and governance, an AWS data lake can process data in any format at petabyte scale. We'll help you imagine the possibilities, identify use cases and speed your time-to-value with reference architectures and accelerators. With our expertise and experience, we can deliver a proof-of-concept in just three weeks.

Business challenges

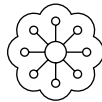
Hadoop based on-premises data lake systems often struggle to deliver value with storage and data co-located and difficult to integrate with other clouds and new technologies

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

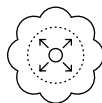
AWS data lake



Unify your data architecture, including hybrid or multi-cloud 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

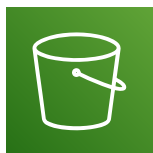
Use Case: Consumer Goods Enterprise

Problem: Customer wanted its smart toothbrush to identify oral health conditions and report status to their users but they were dealing with 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:

- Scalable data infrastructure for data collection from many users
- Trained neural network identifies oral health conditions for users
- Enables sophisticated upgrade to client's smart toothbrush offering



Amazon Simple Storage Service (Amazon S3)



AWS Key Management Service (AWS KMS)



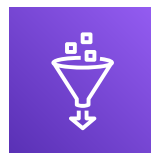
AWS Identity and Access Management (IAM)



Amazon KinesisVideo Streams



Amazon KinesisData Firehose



AWS Glue



Amazon Athena



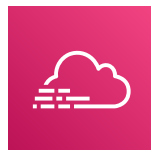
Amazon SageMaker



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).

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

- 2,700+ AWS technical certifications worldwide
- 3,200+ AWS accreditations worldwide
- 12 AWS service delivery designations including Database Migration Services, DynamoDB, EMR, QuickSight, RDS, Redshift
- Expertise across 15 competencies:
 - Data and analytics
 - DevOps
 - Education
 - Financial services
 - Healthcare
 - Industrial software
 - IoT
 - Machine learning
 - Microsoft® workloads
 - Migration
 - Oracle
 - Retail
 - SaaS
 - Storage
 - Travel & hospitality



Premier
Consulting
Partner

Next step

Get started with a complimentary data discovery session to learn how Onica by Rackspace Technology can fast-track your path to becoming a truly data-driven organization. Learn more at:

www.rackspace.com/lp/data-discovery-session.