

Service Name

AWS Data Exchange

Service Overview

AWS Data Exchange is a comprehensive data marketplace service that enables organizations to easily discover, subscribe to, and use third-party data in the AWS Cloud. The service facilitates secure data sharing between data providers and subscribers, offering over 3,500 data sets from more than 300 providers. Data can be accessed through files, APIs, or Amazon Redshift queries and delivered directly to data lakes, applications, analytics, and machine learning models. AWS Data Exchange eliminates the need for providers to build and maintain data delivery, entitlement, or billing infrastructure while providing subscribers with a centralized platform to find, evaluate, and consume external data sets across various industries.

Key Use Cases

- **Financial Services:** Access market data, ESG ratings, consumer spending patterns, and portfolio analytics for risk management, capital markets analysis, and compliance reporting
- **Healthcare and Life Sciences:** Accelerate research with medical imaging data, clinical trial information, claims data, and genomic test results to improve patient care and quantify health outcomes
- **Marketing and Advertising:** Enhance campaigns with consumer behavior data, demographic information, location intelligence, and social media insights to improve targeting and ROI

Features

- **Multi-format Data Support:** Supports Files, API, Amazon Redshift datashares, Amazon S3 data access, and AWS Lake Formation data permissions
- **Automated Data Delivery:** Automatic export of new or updated data revisions to Amazon S3 buckets with structured storage
- **Subscription Management:** Centralized platform to track and manage all data grants and AWS Marketplace subscriptions
- **Data Discovery:** Browse catalog with over 3,500 data sets across multiple industries with filtering and search capabilities
- **Security and Compliance:** End-to-end encryption, subscription verification, IAM-based access control, and compliance with data privacy laws
- **Flexible Pricing Models:** Support for subscription-based, pay-as-you-go, private offers, and custom terms with consolidated billing

Value Proposition

AWS Data Exchange provides unparalleled advantages by eliminating the complexity of third-party data procurement and management. Customers benefit from simplified contracts, consolidated secure billing, and unified data ingestion across multiple providers:

- The service offers native integration with AWS analytics and machine learning services, reducing time-to-insight and operational overhead
- Data providers eliminate the need to build delivery infrastructure while reaching millions of AWS customers, while subscribers gain access to premium data sets with automated updates and built-in security controls, significantly reducing the total cost of ownership for external data initiatives

Service Integration

AWS Data Exchange integrates deeply with core AWS services including Amazon S3 for data storage and export, Amazon API Gateway for API-based data products, Amazon Redshift for direct database queries without ETL processes, and AWS Lake Formation for data lake permissions. The service leverages AWS Marketplace for product discovery and billing, AWS IAM for access control, Amazon Athena for data querying, Amazon SageMaker for machine learning workflows, AWS Glue DataBrew for data preparation, Amazon QuickSight for visualization, and Amazon EMR for big data processing, creating a comprehensive ecosystem for data consumption and analytics.

Onboarding and Offboarding

Getting started with AWS Data Exchange requires an AWS account and appropriate IAM permissions. Subscribers can immediately browse the catalog through the AWS Management Console or AWS Marketplace, evaluate products using data samples and dictionaries, and subscribe with a few clicks:

- Data providers must first register as AWS Marketplace sellers, meet eligibility requirements, and undergo a qualification process
- Both roles can access comprehensive documentation, workshops, and expert guidance
- The service provides step-by-step getting started guides for subscribers and providers, with automated notification systems for subscription updates and new data availability

Getting Started Guide: <https://docs.aws.amazon.com/data-exchange/latest/userguide/setting-up.html>

Data Removal

When subscriptions end, access to AWS Data Exchange data sets is immediately terminated. However, subscribers retain access to previously exported files in their own Amazon S3 buckets. Data providers can unpublish products from the catalog while existing subscriptions remain active until expiration. Subscribers must review their Data Subscription Agreement to understand requirements for deleting exported data. The service provides clear guidance on data retention obligations and removal procedures.

Backup/Restore and Disaster Recovery

AWS Data Exchange relies on underlying AWS service infrastructure for backup and disaster recovery capabilities. Data stored in Amazon S3 benefits from S3's built-in durability and availability features. The service operates across multiple AWS regions with automatic replication of metadata and subscription information. However, AWS Data Exchange does not provide specific backup services and relies on customers to implement their own data protection strategies for exported data.

Backup Capabilities

AWS Data Exchange does not have native backup capabilities and does not integrate directly with AWS Backup. The service depends on the durability and availability of underlying AWS services like Amazon S3, which provides 99.99999999% (11 9's) durability. Customers are responsible for implementing backup strategies for data exported to their own AWS accounts using standard AWS backup services and best practices.

Disaster Recovery

AWS Data Exchange does not directly support AWS Elastic Disaster Recovery integration. The service achieves resilience through its multi-region architecture and dependency on highly available AWS services:

- Disaster recovery is primarily handled through AWS's infrastructure resilience, with data sets and subscriptions accessible across supported regions
- Customers must implement their own disaster recovery strategies for exported data using appropriate AWS services

Recovery Objectives

AWS Data Exchange helps customers meet RPO and RTO objectives through its distributed architecture across multiple AWS regions and high availability of underlying services. The service provides immediate access to subscribed data sets and supports automated exports to reduce manual recovery processes. However, specific RPO/RTO targets depend on customer implementation of backup and recovery strategies for exported data rather than inherent service capabilities.

Service Constraints

Service Quotas: <https://docs.aws.amazon.com/data-exchange/latest/userguide/limits.html>

FAQ: <https://aws.amazon.com/data-exchange/faqs/>

Technical Requirements

Service Documentation: <https://docs.aws.amazon.com/data-exchange/latest/userguide/what-is.html>

User Guide: <https://docs.aws.amazon.com/data-exchange/latest/userguide/>

API Reference: <https://docs.aws.amazon.com/data-exchange/latest/apireference/welcome.html>

Pricing Overview

Please see the AWS UK G Cloud 15 Pricing Document accompanying this service in the Digital Marketplace.

Public Pricing: <https://aws.amazon.com/data-exchange/pricing/>

Cost Optimization

AWS Data Exchange offers unique cost optimization through its tiered fulfillment fee structure for data providers and consolidated billing for subscribers managing multiple data subscriptions. Providers can optimize costs by leveraging automatic data export features, reducing manual delivery overhead:

- The service supports various pricing models including subscription-based and pay-as-you-go options, allowing customers to align costs with usage patterns
- Integration with AWS native services enables customers to minimize data transfer costs by keeping data within the AWS ecosystem and leveraging services like S3 storage classes and data lifecycle policies

Savings Considerations

Major cost savings come from eliminating the need to build custom data delivery infrastructure, reducing operational overhead through automated data management, and leveraging consolidated billing across multiple data providers. Customers save on integration costs through native AWS service compatibility and reduce time-to-market for data-driven applications:

- The pay-as-you-go model for certain products eliminates upfront commitments, while subscription models provide predictable costs

- Additional savings emerge from reduced data transfer fees by keeping operations within AWS, optimized storage costs through S3 integration, and elimination of manual data procurement and management processes