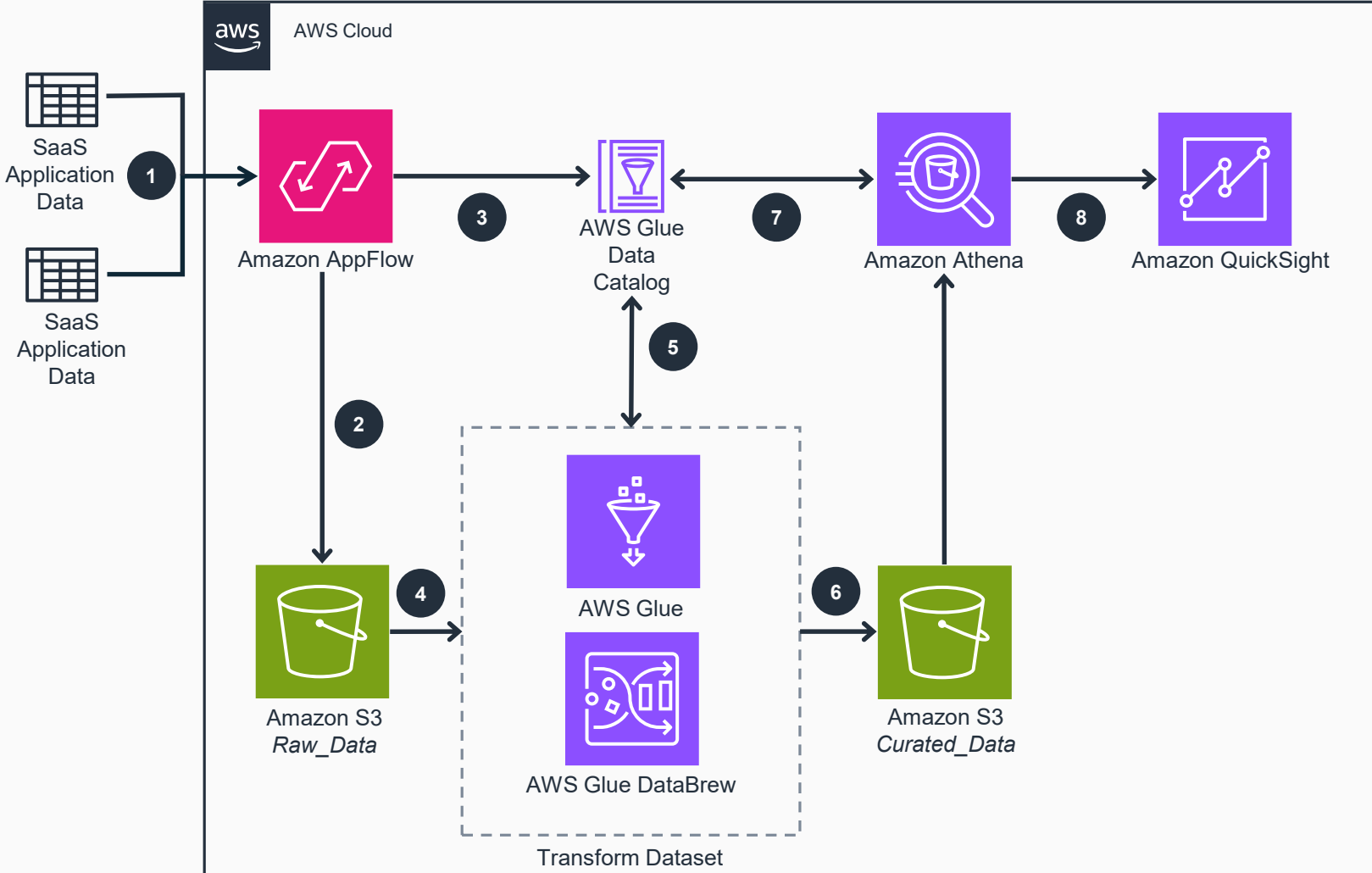


# Guidance for Integrating Third-Party SaaS Data using Amazon AppFlow

This architecture diagram shows how to integrate data from third-party software as a service (SaaS) providers into their data lake on AWS.



- 1 Amazon AppFlow is invoked to run either on-demand, or on a schedule, to pull data from software as a service (SaaS) applications such as SAP, Salesforce, ServiceNow and other SaaS applications.
- 2 Amazon AppFlow stores the raw data pulled from SaaS applications in Amazon Simple Storage Service (Amazon S3).
- 3 Amazon AppFlow integrates with AWS Glue Data Catalog to catalog and store metadata.
- 4 Using AWS Glue or AWS Glue DataBrew, the raw data is transformed and enriched using a low-code serverless data integration pipeline.
- 5 AWS Glue reads Data Catalog to access raw data, and also makes a new table of the enriched and transformed data.
- 6 Transformed data is stored in an Amazon S3 bucket (Curated\_Data) that can be analyzed further using business intelligence (BI) services or tools.
- 7 Amazon Athena integrates with Data Catalog to query curated data from the Amazon S3 bucket.
- 8 Data is loaded into Amazon QuickSight SPICE cache for developing responsive dashboards.

