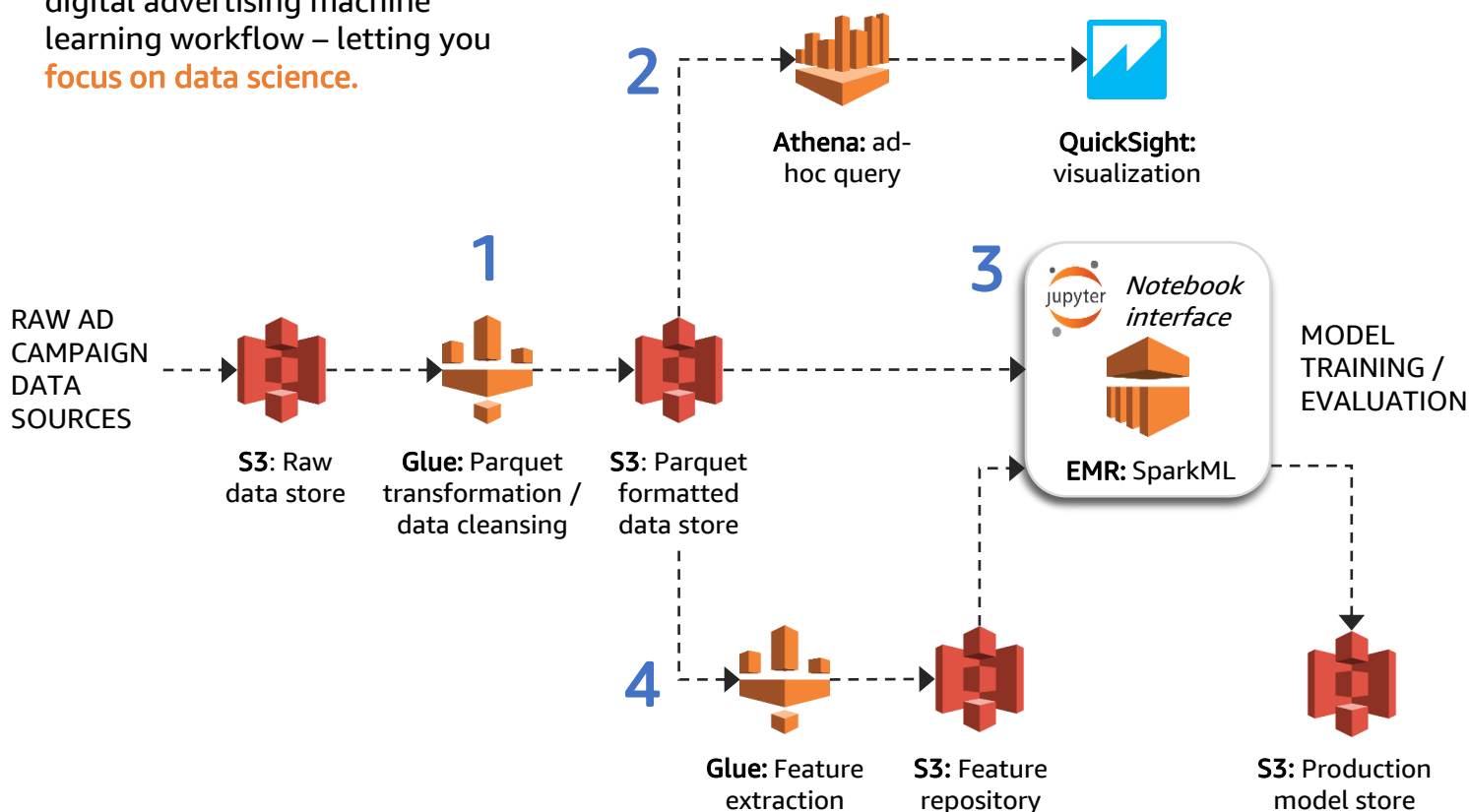


Machine Learning for Digital Advertising: reference architecture

Offload the undifferentiated data pre-processing tasks in the digital advertising machine learning workflow – letting you **focus on data science**.



1 Transformation/cleansing: schedule transformation of incoming campaign data into Parquet format and automate routine data validation and cleansing tasks in AWS Glue

2 Exploration: better understand your data with ad-hoc queries in Athena and visualize outputs in QuickSight – with a simple SQL interface, your BI team can also self-serve queries

3 Train/evaluate: EMR infrastructure running SparkML can size to handle even the largest digital advertising time series datasets

4 Productionize: automate creation of production-grade data features with Glue and store in an S3 feature repository for use in your productionized model training process

Learn more about using AWS for digital advertising:

aws.amazon.com/digital-advertising