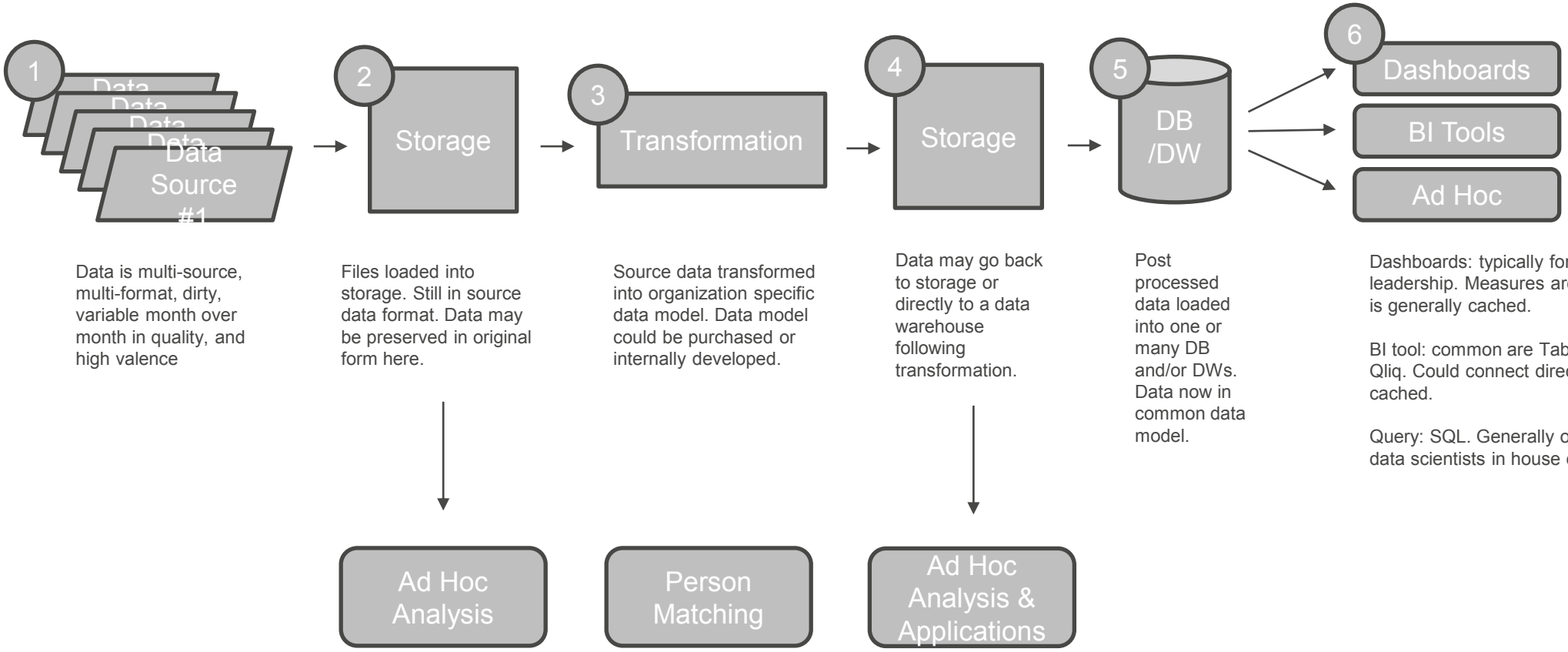
Generalized HHS Analytics Workflow



Data is multi-source, multi-format, dirty, variable month over month in quality, and high valence

Files loaded into storage. Still in source data format. Data may be preserved in original form here.

Source data transformed into organization specific data model. Data model could be purchased or internally developed.

Data may go back to storage or directly to a data warehouse following transformation.

Post processed data loaded into one or many DB and/or DWs. Data now in common data model.

Dashboards: typically for organization leadership. Measures are pre-defined and data is generally cached.

BI tool: common are Tableau, MicroStrategies, Qliq. Could connect directly into DW or may be cached.

Query: SQL. Generally only in orgs that have data scientists in house or on contract.

amazon web services