Twitter Stream Processing Pipeline

Stream Ingestion

Start Data Flow

Producer

Select Input Data

Pick a dataset

Add Metadata

Organize tweets with key attributes

Publish to Kafka

Stream tweets in real-time

Kafka-Broker

Define Kafka topic

Split data into partition

Send Data for Processing

Consumer

Processing Component

Retrieve hashtags

Perform sentiment analysis

Extract geographic coordinates

Build complete indexes

Store Indexed Data

Storage

Design Elasticsearch schema for optimization

Save processed data (hashtags, sentiments, locations)

Send Processed Data

---

web application

Visualize sentiment analysis

Search Engine Component

Map geo-tagged data