Event-Driven Senty
Final Presentation

St. Gallen, 25. May 2023
Group 1 – Johannes, Luka, Philipp

From insight to impact.

Agenda

1. Senty Overview

2. Live Demo

3. Topology

4. Development Steps

   ◦ Ex 7: Stream Processing with Kafka Streams I
   ◦ Ex 8: Stream Processing with Kafka Streams II
   ◦ Ex 9: Stream Processing with Kafka Streams III

5. Learnings + Q&A

Senty

Use case: Sentiment Analysis tool

Live Demo

Comment your thoughts and emotions ☺

The 10 Myths of Entrepreneurship

277.492 Aufrufe • vor 10 Jahren

Topology

```
                          StreamsBuilder
                                │
                             stream()
                                │
                        CommentBatchStream
                                │
                    flatMapValues() & selectKey()
                                │
                           Comments
                                │
                    removeEmojisFromComments()
                                │
                     Comments without Emojis
                                │
                            transform()
                                │
                      Deduplicated Comments
                                │
                            transform()
                                │
                      Comments with Language
                         ╱              ╲
                   branch()            branch()
                      ╱                    ╲
                     │              Non-English Stream
                     │                      │
                     │                   filter()
                     │                      │
                     │            Translatable Comments Stream
                     │                 ╱          ╲
                     │            filter()      transform()
                     │              ╱               ╲
              English Stream  Untranslatable   Translated Stream
                     │         Comments             │
                 merge()                         merge()
                      ╲                            ╱
                          Merged Stream
                                │
                         filterComments()
                                │
                      Filtered Comment Stream
                                │
                            transform()
                                │
                    Sentiment Classified Stream
                                │
                            groupBy()
                                │
                         Grouped Stream
              ╱           │           │            ╲
        aggregate()    count()    aggregate()    aggregate()
          ╱              │           │               ╲
  Languages KTable  Comment Count  Recent Comments  Total Sentiment
                      KTable          KTable           KTable
```
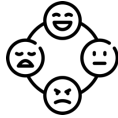
Development Steps
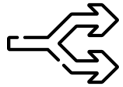
Ex 7: Stream Processing with Kafka Streams I

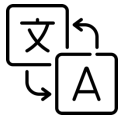**Stateless stream operations implemented:**

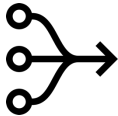Filtercomments -> filtering for Blacklisted Keywords

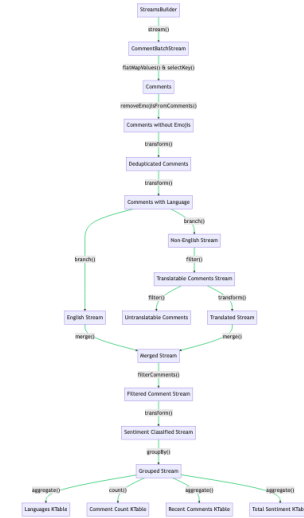SentimentAnalyzer -> filtering the sentiment for comments

flatMapValues -> breaking comment batches into single comments

CommentTranslator -> translates the comment into english

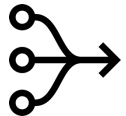merge(translated_stream) -> merging the translated stream back

Ex 8: Stream Processing with Kafka Streams II

Streams and Tables implemented:

- Language KTable

- Total Sentiment KTable

- Recent Comments KTable

- Comment Count KTable

Data from multiple streams:

Merging Translated and English stream

Ex 8: Stream Processing with Kafka Streams II

Interactive Queries:

- Retrieve the number of comments in each language for a list of video IDs
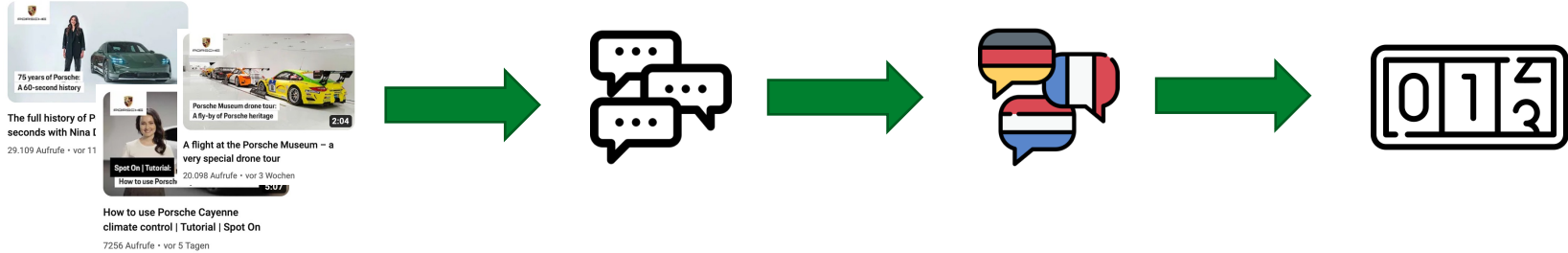
- Retrieving Comments regarding their sentiment for a specific Video(todo)
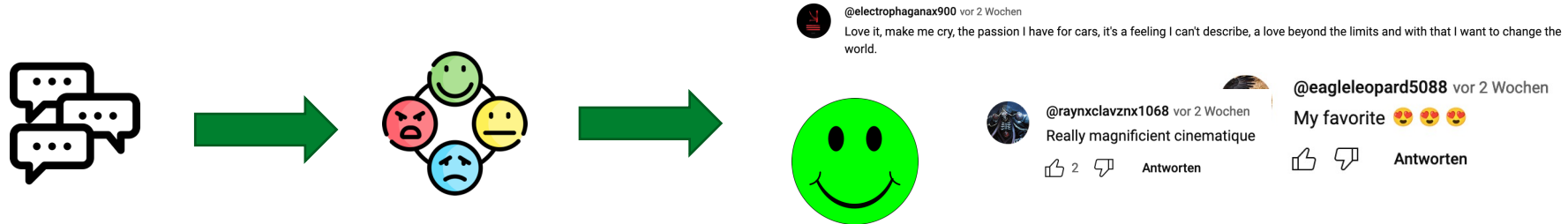
Ex 8: Stream Processing with Kafka Streams II

Interactive Queries:

- Retrieving most recent average Sentiment per video

- Retrieving most recent 5 Comments
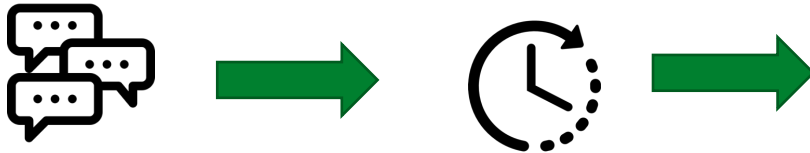
@electrophaganax900 vor 2 Wochen
Love it, make me cry, the passion I have for cars, it's a feeling I can't describe, a love beyond the limits and with that I want to change the world.

@eagleleopard5088 vor 2 Wochen
My favorite 😍 😍 😍

@raynxclavznx1068 vor 2 Wochen
Really magnificient cinematique

👍 2 👎   Antworten

👍 👎   Antworten

Ex 8: Stream Processing with Kafka Streams II

Interactive Queries:

- Retrieving amount of Comments for specified videos

Ex 9: Stream Processing with Kafka Streams III

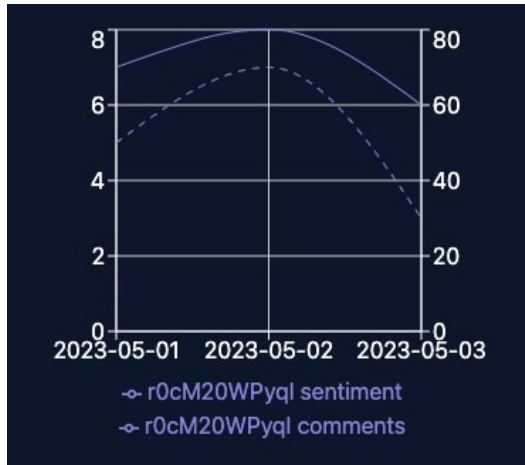**Windowed Operations**

Count of Languages
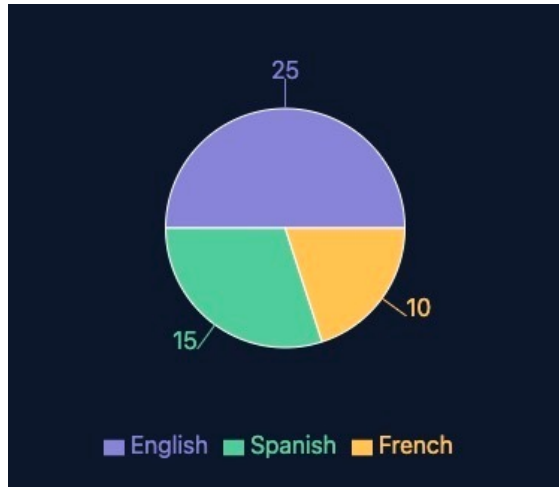
Most used words

Ex 9: Stream Processing with Kafka Streams III

Windowed Operations

Average Sentiment and amount
of Comments per day (todo)

Most recent comments
with sentiment

Architectural Decisions

Topologies for stream processing
ADR #0007

Decision:
We have decided to use a single, main topology in Senty for all our stream processing needs. This topology will encompass all necessary operations for our application's data processing pipeline

Consequences:
+++ Clear Data Flow, Consistency in Processing, Simplified Debugging

--- Scalability Concerns, Potential Bottlenecks

Stateless and stateful processing
ADR #0008

Decision:
We have decided to implement both stateless and stateful processing methodologies in our stream processing service, depending on the specific requirements of each processing stream.

Consequences:
+++ Flexibility in Processing, Independent Data Processing, Tailored Processing

--- Increased Complexity, Increased Resource Requirements, Potential Inconsistencies

Learnings

Learnings

- Getting the Topology right

- Complications with Stateful processing

- Designing Interactive Queries can get tricky

- Connecting all stream processes in one Topology

- Choosing the right kind of Windowed Operations

Learnings

Thank You