

NO RETWEET BDT

A DEMO PROJECT EXHIBITING THE POWER OF BIG DATA TOOLS

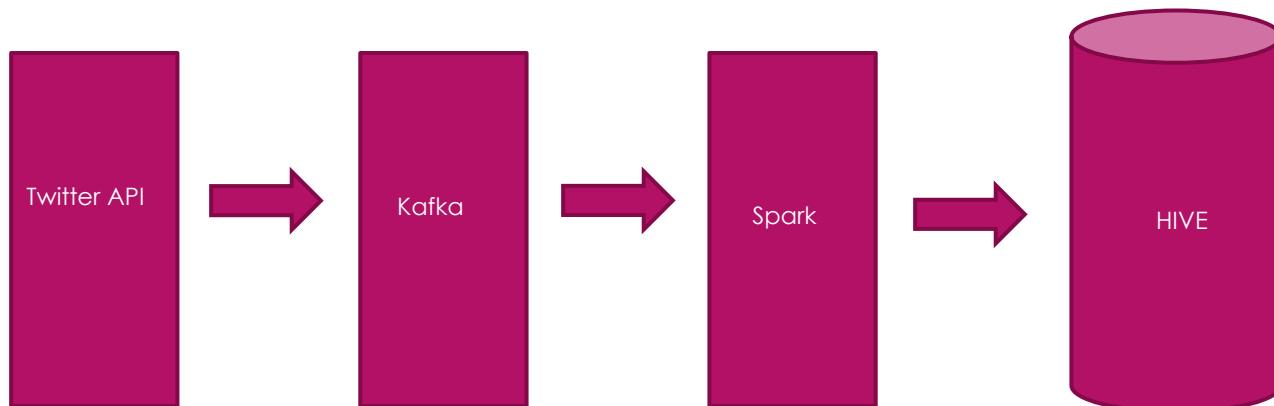
PAUL KIGOZI, CS523

TOOLS USED

- ▶ Spring Integration
- ▶ Twitter Streaming API
- ▶ Kafka + KaDeck
- ▶ Spark 3.0
- ▶ Hive

FLOW DIAGRAM

▶ Streaming data pipeline



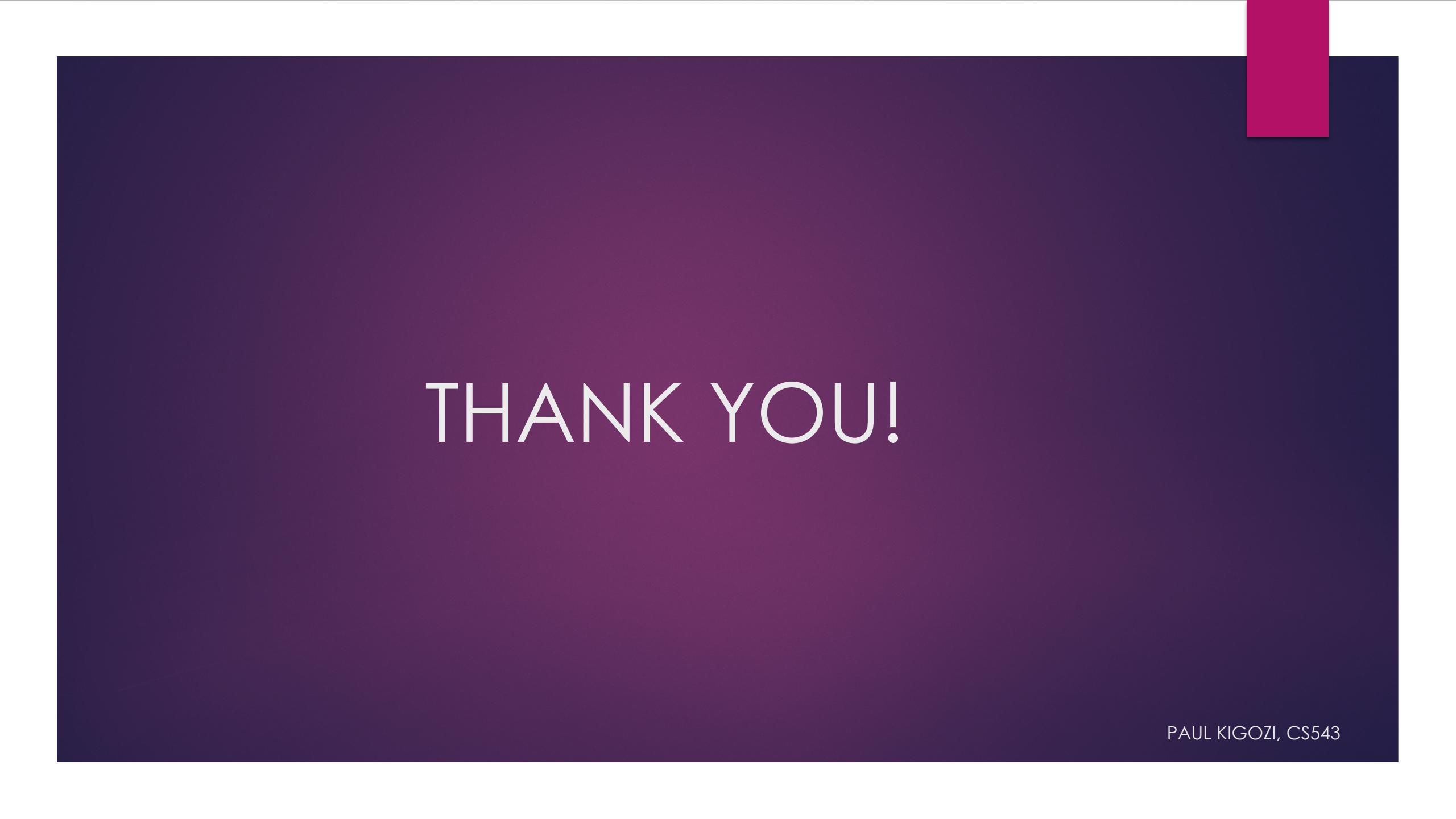
I/P & O/P

- ▶ I/P is A stream of tweets from Twitter API
Configured with personal twitter account setup as twitter developer.
- ▶ O/P is tweet texts stored in Hive (Without any RTs)

Commands

- ▶ Go to project root and run:
- ▶ mvn spring-boot:run

Note: Using embedded spark and hive.



THANK YOU!

PAUL KIGOZI, CS543