

1GB Množina správ



Ruby skript, ktorý simuluje prúdiace správy zo sociálnej siete Twitter

Miešané rovnomerné zoskupenie správ

{téma:"tweet",
msg: "...text
..."}

Kafka komunikačný systém

Predspracovanie prúdu

Programovací rámec Storm

Prameň správ



{tweet -> "...text #auto..."}

Extrakcia identifikačných slov (hashtag)



{hashtag -> "auto",
tweet -> "...text
#auto..."}

Filtrovanie prednastaveným filtrom



Extrakcia dát



Extrakcia užitočných informácií zo správy, napríklad text a meno autoa

Uloženie do Redis



Spracovanie v reálnom čase

Post-sprac.