

Twitter hashtags: Joint Translation and Clustering



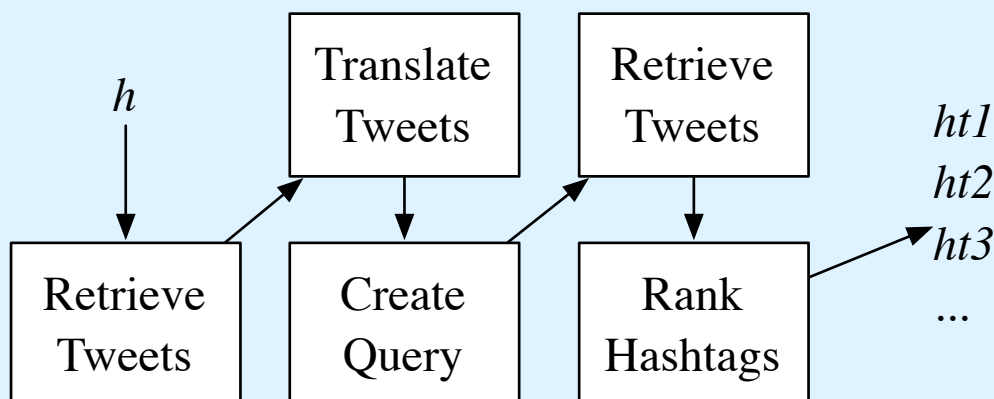
Simon Carter, Manos Tsagkias, Wouter Weerkamp
University of Amsterdam, ISLA



Task Find translations for a given source-side hashtag

Problems 1. Intra and 2. Inter language plurality + variability in hashtag use for the same topic

Method Retrieval based. Given a foreign hashtag h , the algorithm works as follows:



Example Source hashtag: Query Terms: Retrieved hashtags:

33mineros

rescue

#Chile

rescued

#chilemine

miners

#fuerzamineros

miner

#CNN

trapped

#miners

capsule

#chileanminers

Conclusions

1. Proof-of-concept
2. Large-scale experiments
3. Full system-integration