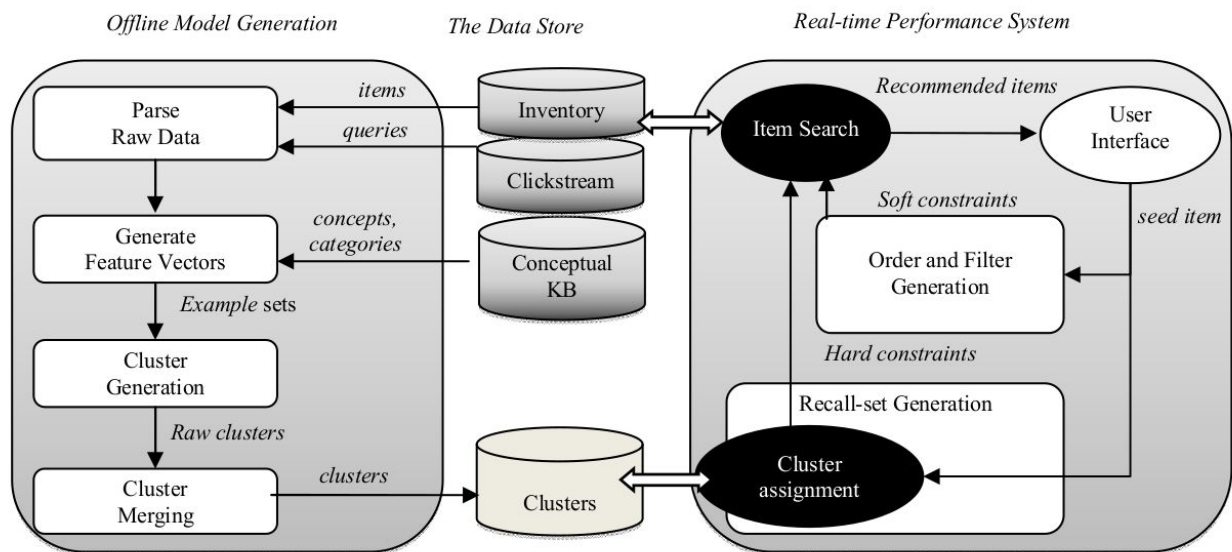


Please summarize and try to assess the approach described in this paper
<http://bit.ly/2oYesuo>.



What is the motivation?

- Recommending similar items to increase user engagement
- Dealing with short-lived items
- Scalability

What are the key ideas?

- Trade-off between similarity and quality
 - After placing a bid: more specific results (similarity)
 - Coming from a search result: broader results (more weight to quality)
- Short-lived items, but long-term clusters
- Use user queries to learn how they group items
- A separate clustering process can run for each user query and therefore the algorithm is highly parallel

How does the approach differ from collaborative filtering or a “naive information retrieval System”?

- Deals better with highly dynamic inventory of short-lived items
- Enables a balance between quality and similarity

What are possible shortcomings of the approach?

- Clustering based on occurrence of terms doesn't capture some semantic similarity.
- Clusters might get outdated if a large number of new items appear on a short period.

Can you think of possible extensions?

- Use topic modeling to replace fixed-term clusters with term distributions
- Use an incremental clustering approach to keep clusters updated without the need of expensive model re-training.