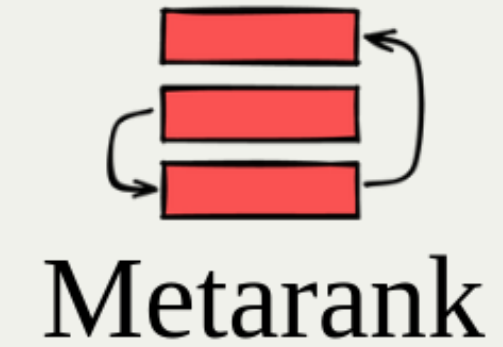


co:here



Semantic recommendations

A demo for the lablab.ai **semantic search** hackathon

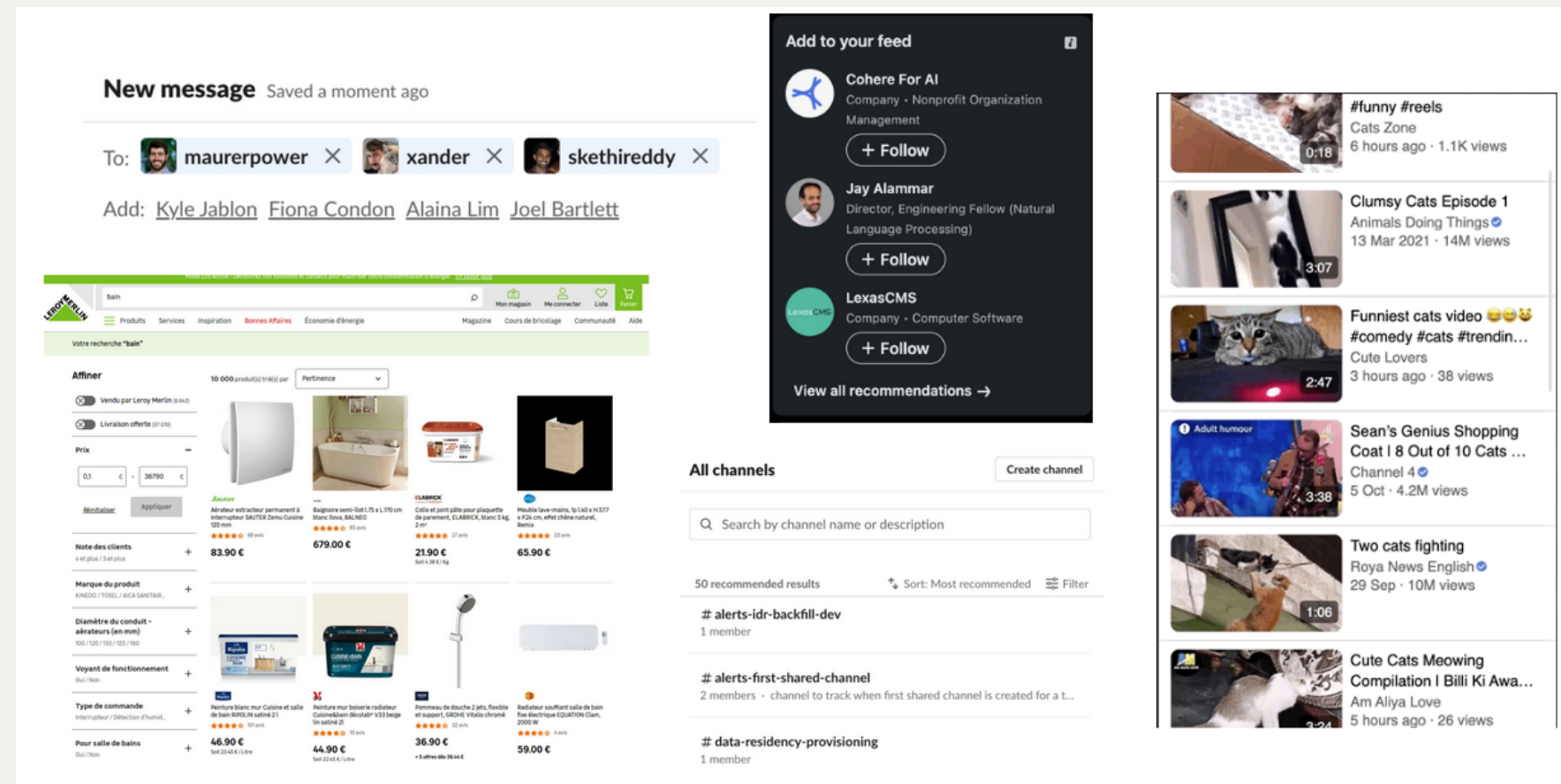
Team **metacrank**, 2023

The metacrank team of one



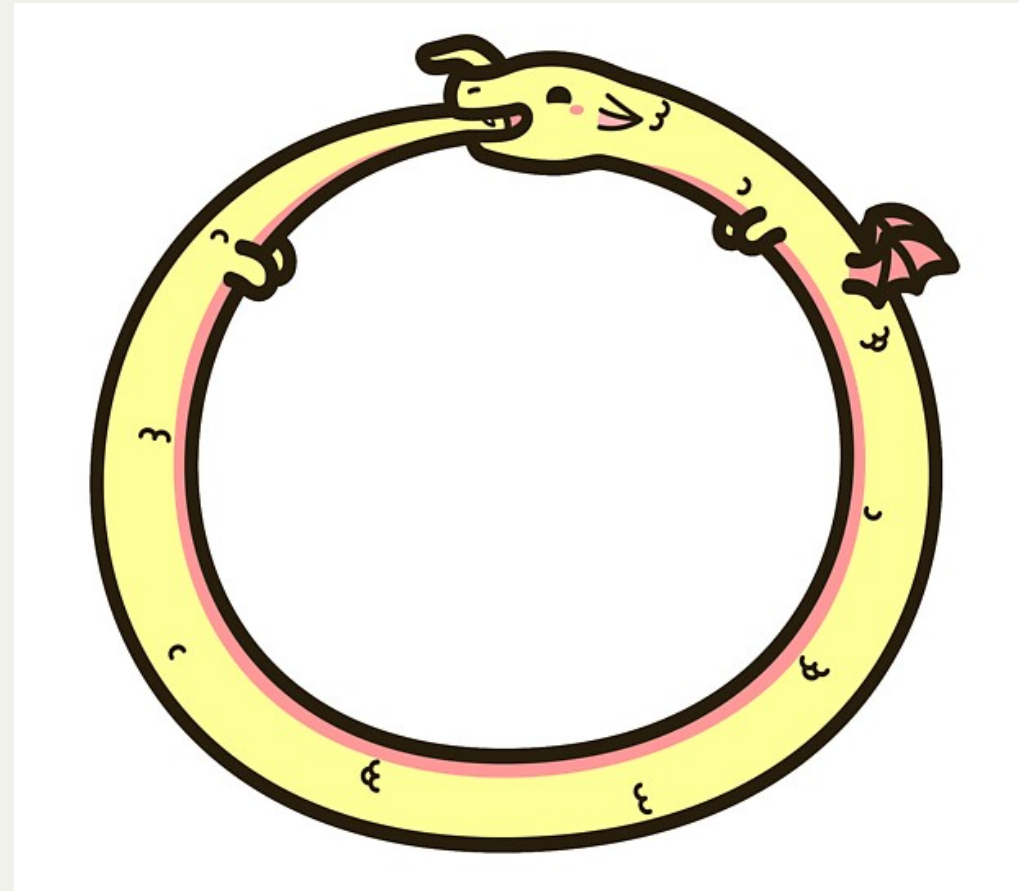
- Long ago: PhD in CS, quant trading, credit scoring
- Past: Search & personalization for ~7 years
- Now: ~~Unemployed Open-source contributor~~ Principal DS

Recommendations 101



- **Similarity** between a current context and a set of candidates
- **Collaborative filtering**: interactions of other people
- SVD/ALS, BERT4rec/SASRec, LightFM, ...

Recommendations cold-start problem



- To have good recs, you need **visitor feedback**
- For the feedback, you need to show **good-enough** recs
- **But you can't**, as you don't have the feedback yet!

Content-based recommendations to the rescue



- But we do have content: title and description!
- Embedding of context \sim embedding of candidates













Semantic movie recommendations



- Take MovieLens/TMDB as a source of clicks and metadata
- Train MF ALS for a baseline CF recs
- [co:here/sbert](https://github.com/alexanderlindner/sbert) embeddings of title+description, Qdrant for k-NN search

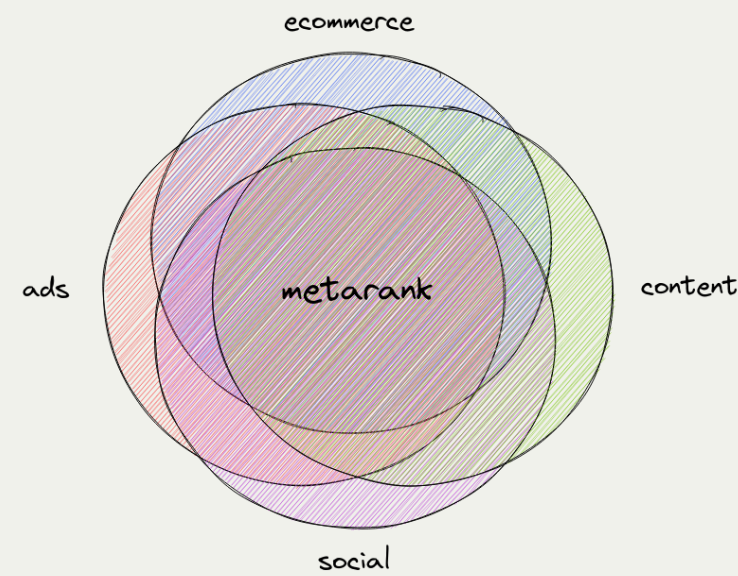
Demo

Metarank semantic movie recommendations Trending About

 <p>Pulp Fiction</p> <p>A burger-loving hit man, his philosophical partner, a drug-addled gangster's moll and a washed-up boxer converge in this sprawl...</p> <p>Select 154 mins</p>	 <p>The Matrix Reloaded</p> <p>Six months after the events depicted in The Matrix, Neo has proved to be a good omen for the free humans, as more and more human...</p> <p>Select 138 mins</p>	 <p>The Matrix Revolutions</p> <p>The human city of Zion defends itself against the massive invasion of the machines as Neo fights to end the war at another front...</p> <p>Select 129 mins</p>	 <p>Léon: The Professional</p> <p>Léon, the top hit man in New York, has earned a rep as an effective "cleaner". But when his next-door neighbors are wiped out by...</p> <p>Select 111 mins</p>	 <p>The Terminator</p> <p>In the post-apocalyptic future, reigning tyrannical supercomputers teleport a cyborg assassin known as the "Terminator" back to ...</p> <p>Select 108 mins</p>	 <p>The Butterfly Effect</p> <p>A young man struggles to access sublimated childhood memories. He finds a technique that allows him to travel back into the past...</p> <p>Select 113 mins</p>
 <p>Fight Club</p>	 <p>Pirates of the Caribbean: On Stranger Tides</p>	 <p>Back to the Future</p>	 <p>Harry Potter and the Philosopher's Stone</p>	 <p>The Matrix</p>	 <p>Inception</p>

<http://semrec.dfdx.me>

How it was made



Glueing things together with Metarank:

- Existing open-source project for recs/ranking
- PR: semantic recommendations support
- PR: qdrant knn-search support
- A simple Flask UI with Movielens/TMDB dataset

YAML ML FTW

```
cohere:
  type: semantic
  itemFields: [title, description]
  store:
    type: qdrant
    size: 4096
    distance: Cosine
    endpoint: "http://qdrant:6333"
    collection: cohere

encoder:
  type: csv
  path: /conf/cohere-large.csv
```

Thanks!



- Demo source: github.com/shuttie/lablab-qdrant-cohere-hackathon
- Metarank: github.com/metarank/metarank
- Demo: <http://semrec.dfdx.me>
- Slide deck: shuttie.github.io/lablab-qdrant-cohere-hackathon