

# Midterm Presentation

Jingheng Pan

jingheng.pan@studium.uni-hamburg.de

Adrian Lindloff

adrian.lindloff@studium.uni-hamburg.de

Lijunnan Bai

lijunnan.bai@studium.uni-hamburg.de

# Structure

1. Recap
2. Comparison to Project Plan  
→ changes
3. New features & concepts
4. Demo
5. Future Work
6. Milestones

## What is the “Adaptive Storyfinder” Project?

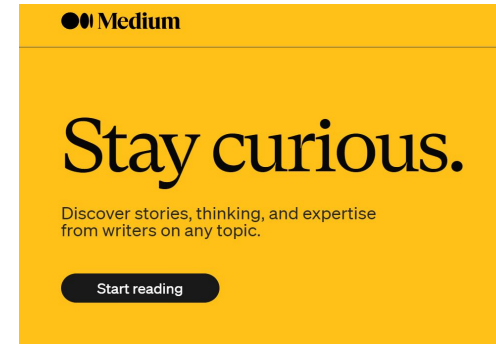
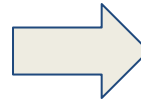
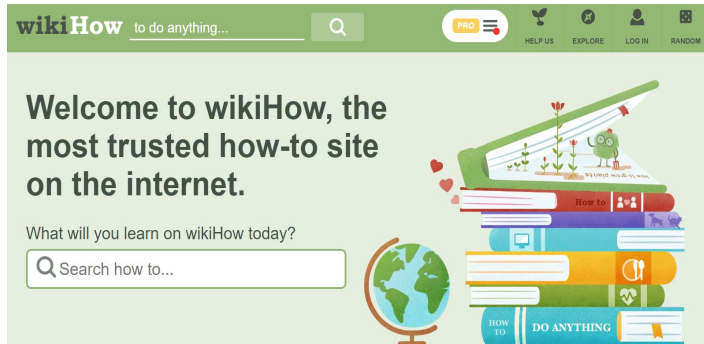
- Personalized content recommendation system
  - Create user preferences
  - Focus on re-ranking articles

# Comparison to the Project Plan

## Comparison: Article Database

## Article database:

- From WikiHow API & dataset to a medium dataset



# Comparison: Technology

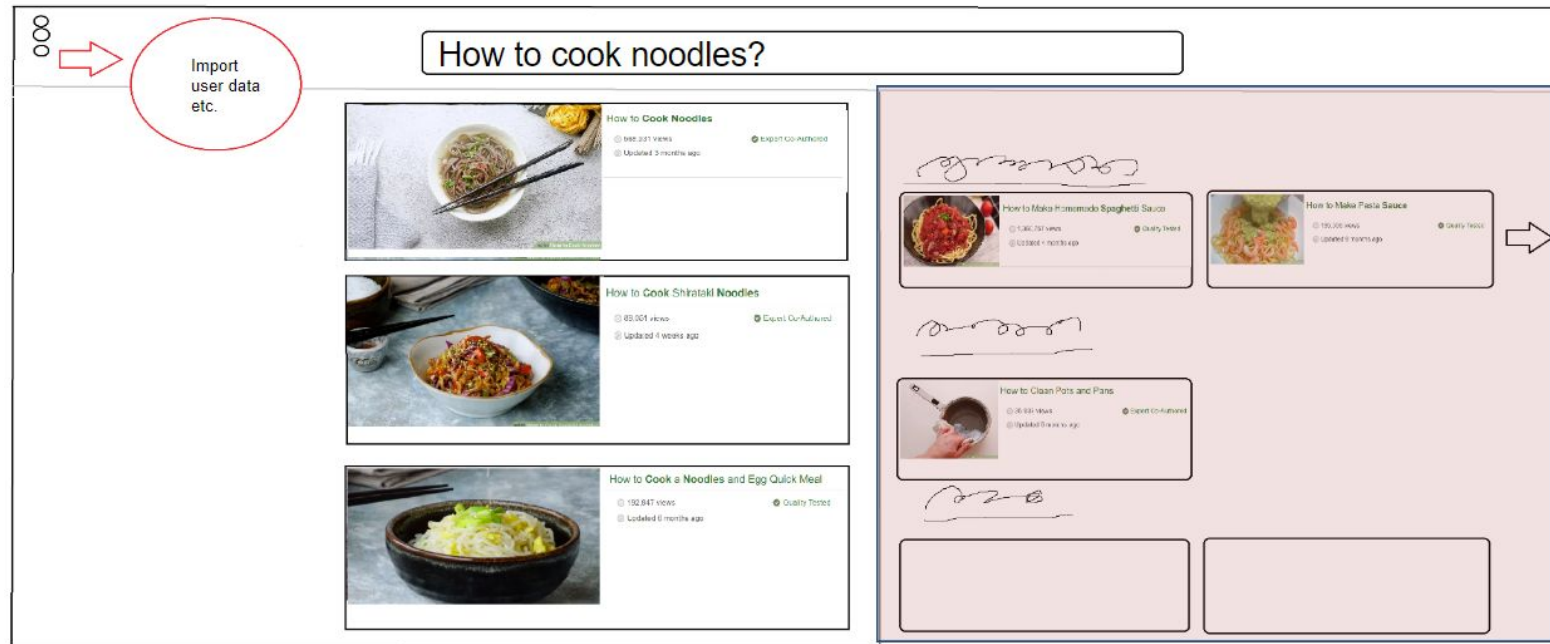
## Backend:

- ~~Django (Python web framework)~~ → FastAPI
- Docker
- PyTorch for ML + HuggingFace

## Frontend:

- Vue.js with Vuetify & Bootstrap

# Comparison: Appearance



# Comparison: Re-ranking

## Old concept:

- Build user preferences with website features like f.E. the amount of images
- Re-rank with user preferences based on behavioral data

## New concept:

- Re-rank with user **histories & websites visited** (with the plugin)

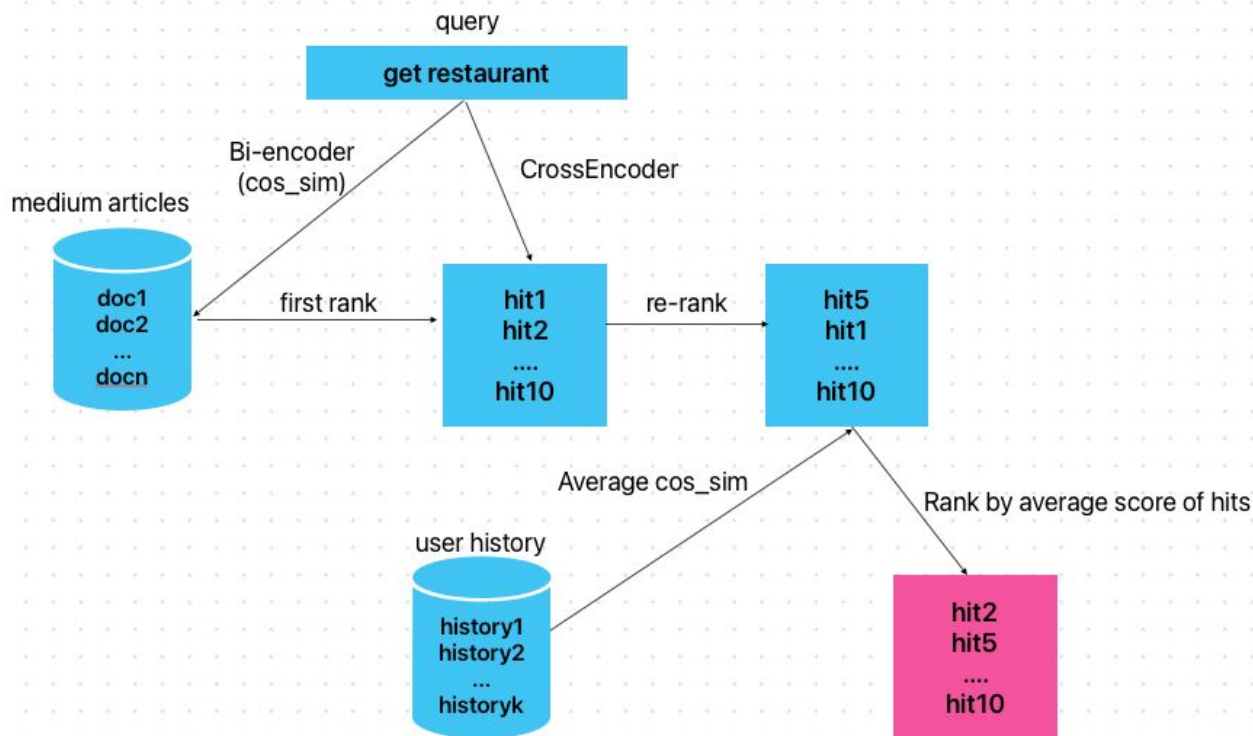


# Comparison: Functionality

- Re-ranking: Article with the highest similarity to the preferences is at the top
- Focus of the Re-ranking on content now
  - instead of behavioral data usage of the text of one's browser history
- Tag recommendation with self-generated tags for each article
  - instead of the “you might also like ...” section

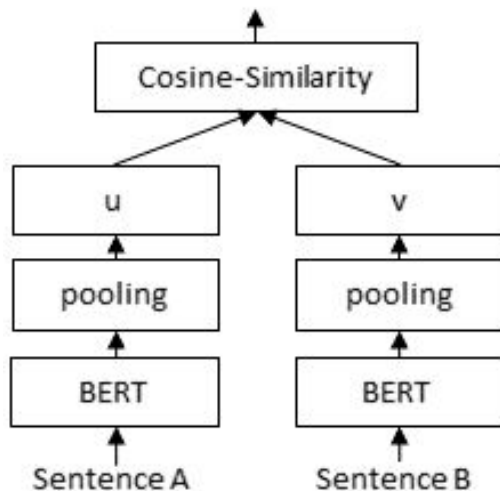
# Functionality & new features

# Re-ranking

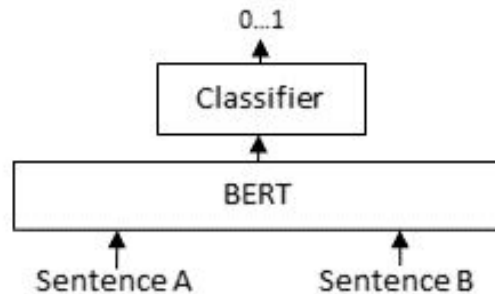


# Sentence-transformers

## Bi-Encoder



## Cross-Encoder



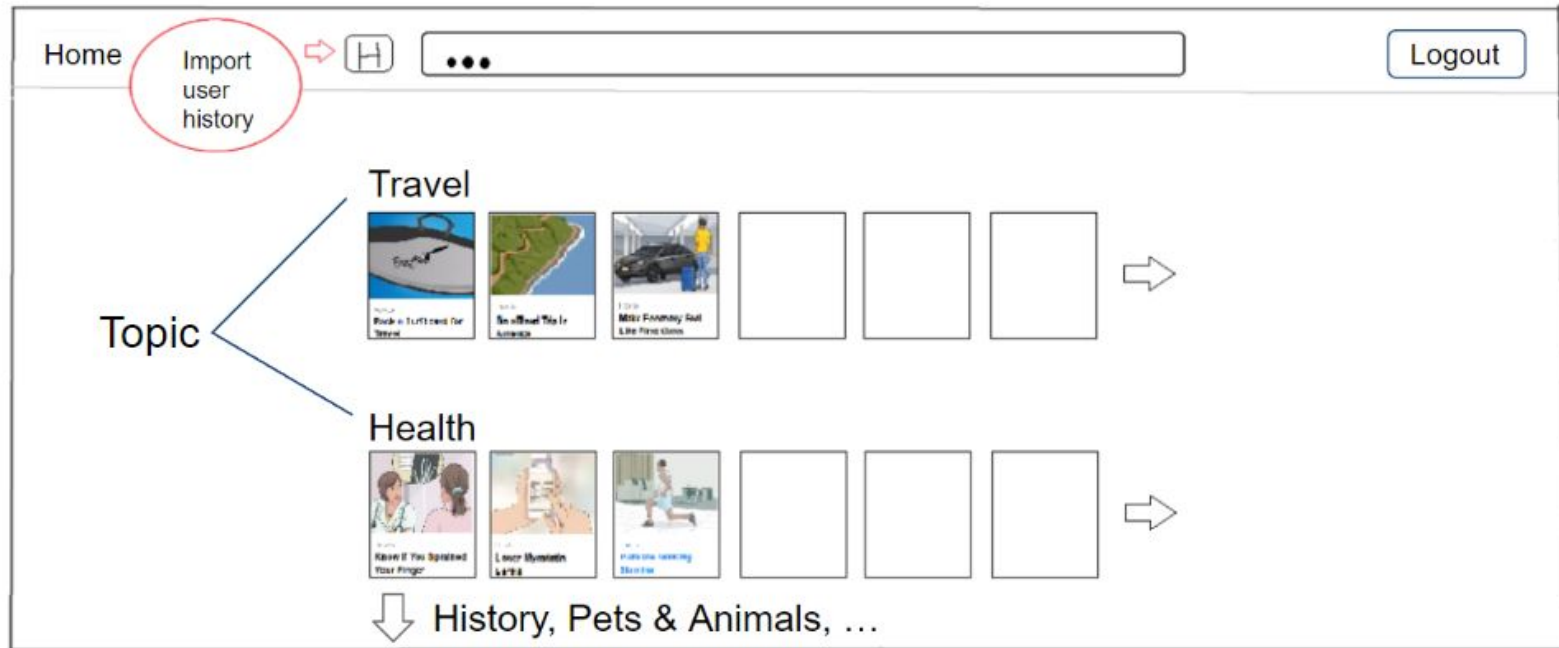
Bi-encoder:

- Pass sentence independently
- Output: embeddings

Cross-encoder:

- Pass sentence simultaneously
- Output: 0 and 1, indicating the similarity of the input sentence

# Topic recommendation



# Topic recommendation

- Generate new topics for the database articles
  - currently premade tags which are quite arbitrary
- Using Bert-based pre-trained model
  - News-category-classification (49 categories)
- An alternative model generated 572 topics
  - hard to generalize

# Topic recommendation

## News-category-classification example topics:



WELLNESS



PARENTING



HEALTHY LIVING



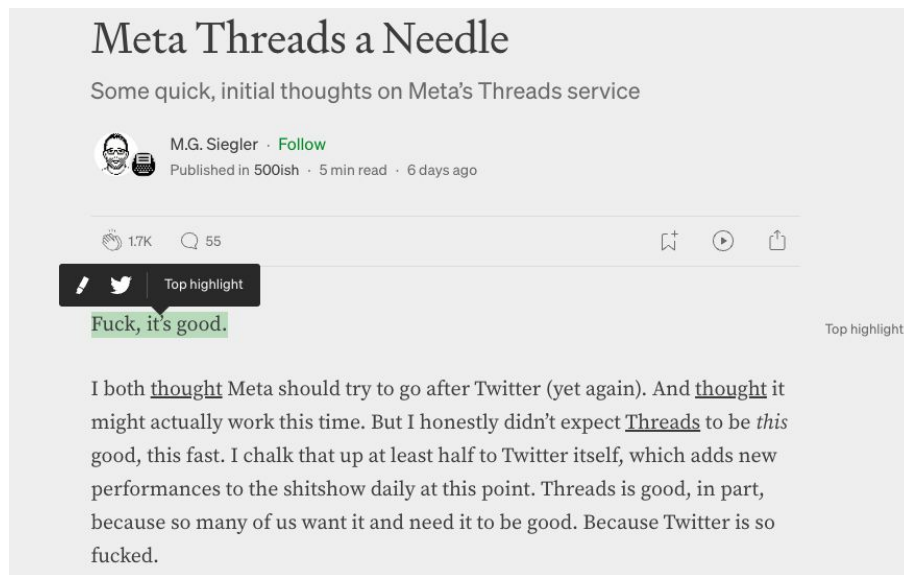
WOMEN

- Randomly take sample articles from every topic to display

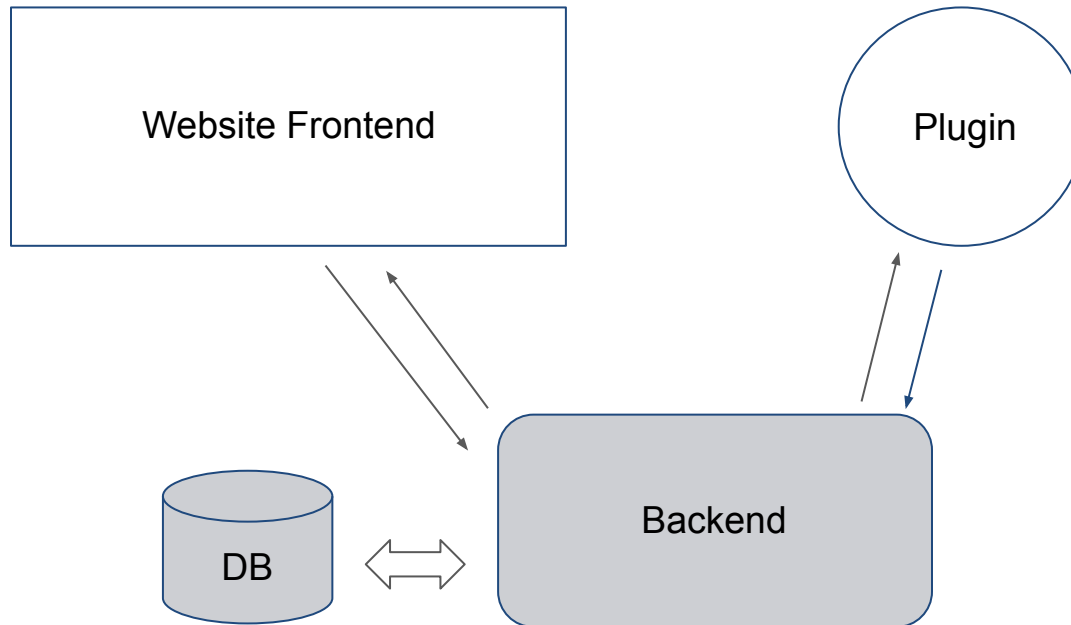
- Also show topics interesting to a specific user

# Section Highlighting

- Highlight parts of a website which might be especially interesting to a user  
→ Inject a highlighting style on a website (done by the browser plugin)







## **Backend deals with:**

- User Management

## **For the Website Frontend:**

- Re-ranking based on
  - user histories
  - additional data from plugin
- Topic recommendation from
  - user histories

## **For the plugin:**

- Highlighting of paragraphs

# Demo

# More features & thoughts

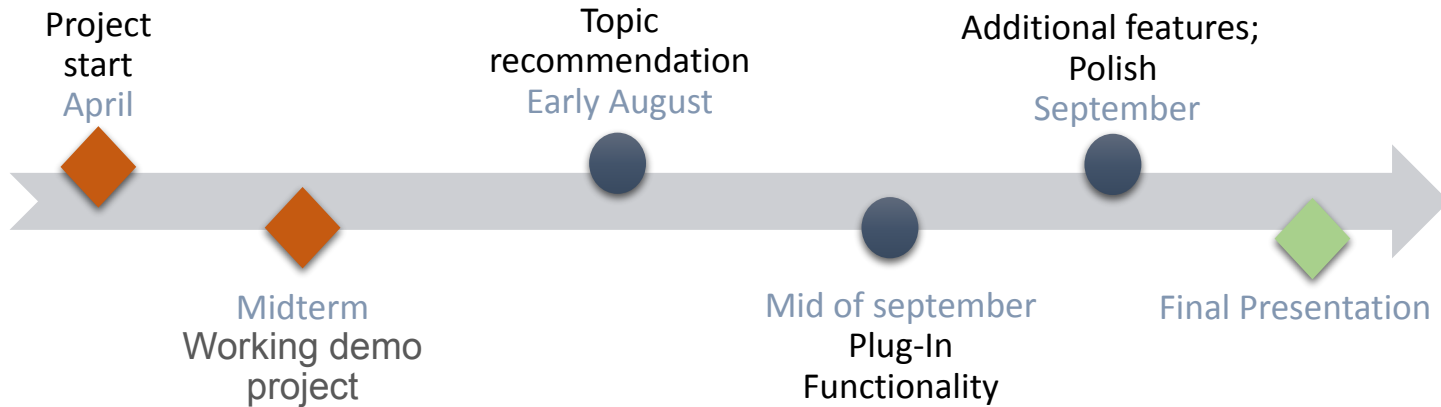
- History management → give users more control
  - Summarize an article
  - Detect informative history links
- 
- Include other sources of information beyond the medium dataset  
→ expand dataset with more sources & collected data by the plugin

# More features & thoughts

- Recommended highlights from other users
- Create scores for predefined features(f.E. amount of images etc.) of each presented article → further refine the re-ranking
- Implement user/ data privacy or at least inform the users about the scope of data usage

1. Personalized topic recommendation:
  - Create new, meaningful topics for the dataset (*currently*)
  - Display topics which a certain user might like
  
2. Plugin:
  - Highlighting of paragraphs on websites
  - Collect visited websites for each user (can be turned off)

# Project Outline



# Any Questions?

# Project Outline

