**what's app chat analysis — May change to another platform since we cannot download the data — tiktok** [**https://github.com/PolyAI-LDN/conversational-datasets/tree/master/amazon\_qa**](https://github.com/PolyAI-LDN/conversational-datasets/tree/master/amazon_qa)

Why? — Reveal hidden patterns and insights within conversations. This can translate to better customer engagement strategies or improved product feedback loops for businesses. In academic settings, chat analysis can shed light on social dynamics, communication styles, and linguistic trends.

How? — Sentiment analysis to analyze the emotional expression, topic modeling to find the keywords/topic in the conversation, and regression model to predict future message events.

TikTok Trend Analysis: <https://developers.tiktok.com/>

<https://github.com/JoeanAmier/TikTokDownloader>

- Utilize the TikTok API to collect data on trending hashtags, challenges, and content creators.

- Aggregate data on the number of views, likes, and comments for trending content.

- Implement a user-friendly dashboard that visualizes trends, their growth trajectories, and demographic information of the audience engaging with these trends.

- Provide insights into the duration of trends, identifying patterns and factors influencing their popularity.

D

Lisha Xia:

电影推荐系统：

1. 数据可以从TMDb的API获得，免费公开的API；可以获得包括类型，演员，评论，评分等信息。
2. 建模：Item CF：一种基于项目相似性的协同过滤方法，这种相似性是通过计算人们对这些项目的评分来确定的。

Why 使用Item CF

● 基于项目之间的相似度。如果一个用户喜欢某个项目，那么这个用户可能也会喜欢与这个项目相似的其他项目。

● 产品不会频繁变更，降低计算量

1. 训练模型：使用用户的观影历史来训练推荐模型。
2. 前端开发：开发一个友好的用户界面，让用户能够轻松提供反馈和查看推荐。

后端是推荐系统

Tianran Huang:

### **Generating content according to image**

[GitHub - dabasajay/Image-Caption-Generator: A neural network to generate captions for an image using CNN and RNN with BEAM Search.](https://github.com/dabasajay/Image-Caption-Generator)

A neural network to generate captions for an image using CNN and RNN with BEAM Search.

why procedure is a system implementing an existing data mining method to solve a real-life problem:

1. The usage of hashtags and attention-getting subtitles might help you reach out to the right people even more.
2. Natural language processing, such as text generation, automatic summarization, and pre-training models, only involve text, so how to make the model effectively interact with image and text information is very critical.
3. There are many open-source datasets that could be used to test the efficiency of the model.