

CS573 - Final Project - Find Your TED

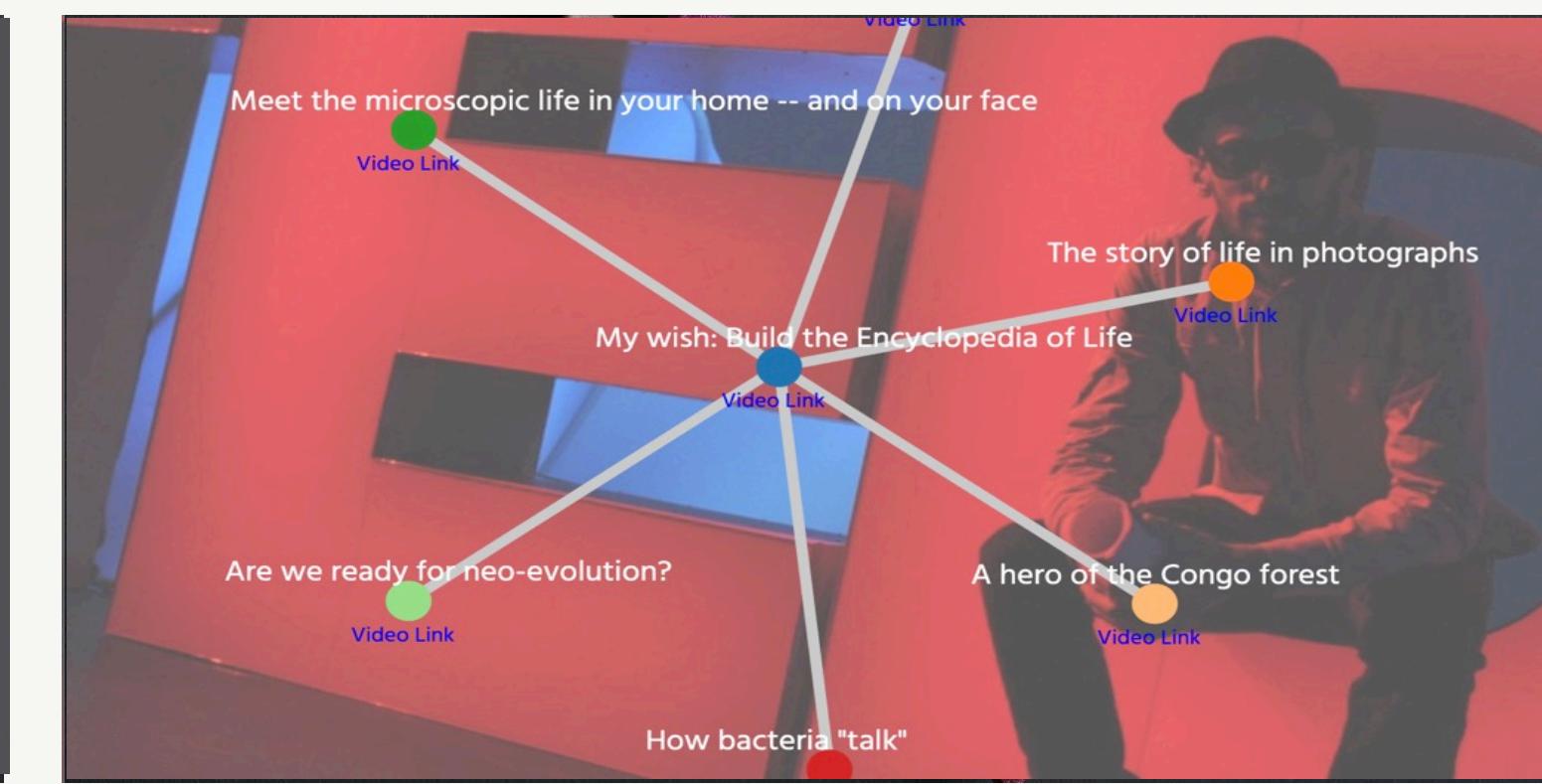
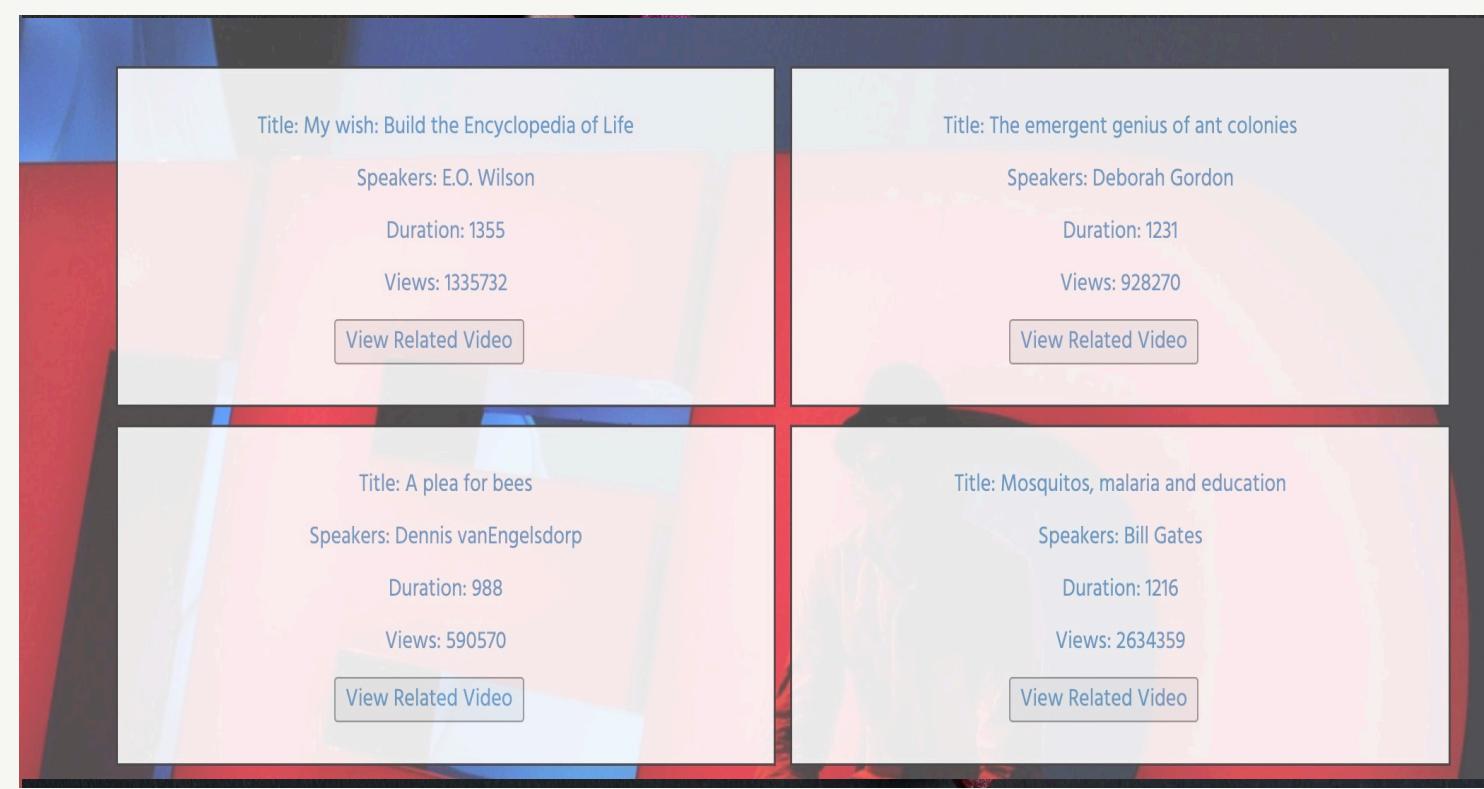
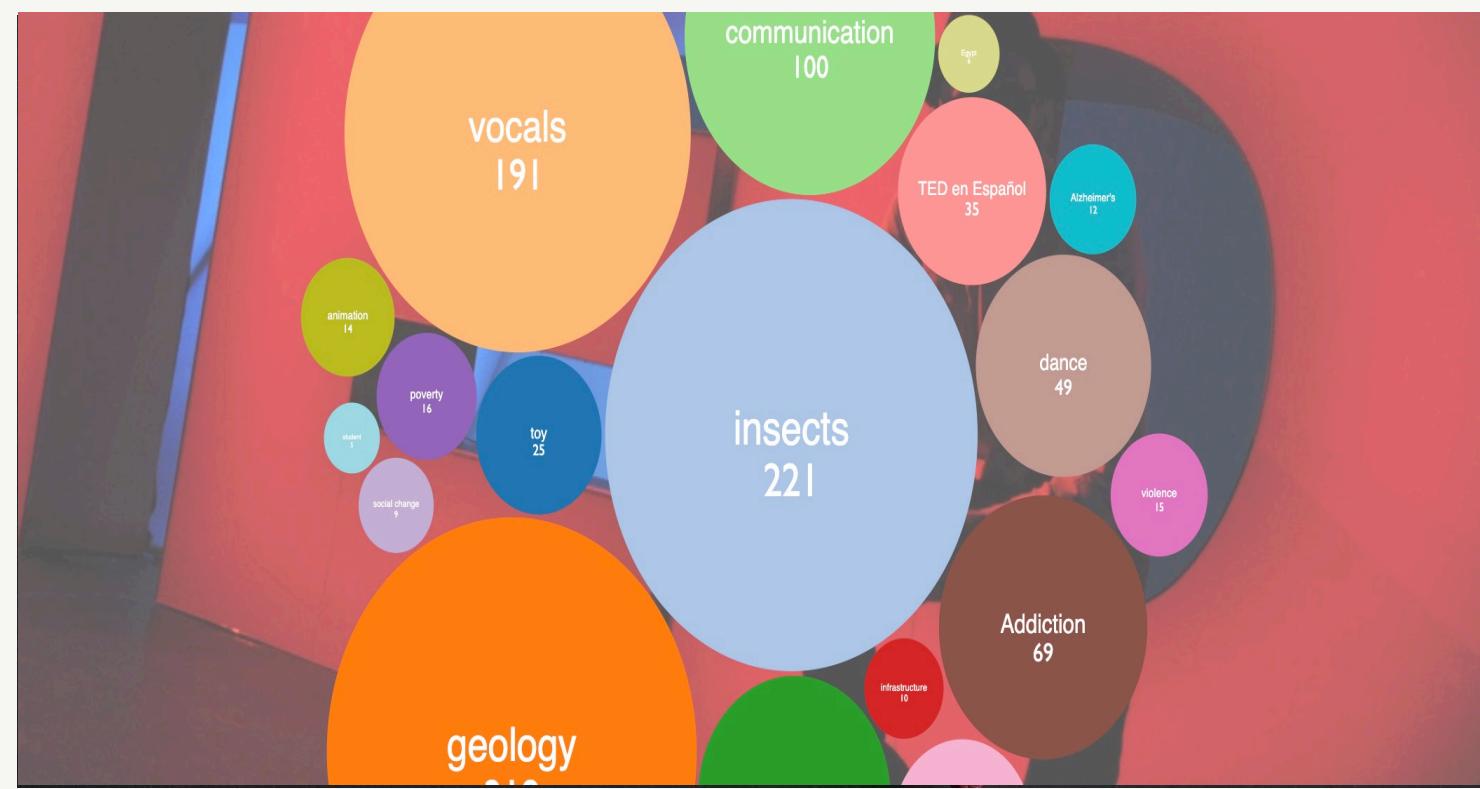
Mingjie Zeng & Yufei Lin

Process Book

May 2, 2022

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin



Find Your TED

Find Your TED is a visualization dashboard that can give you suggestions about finding your favorite TED talks through the related networks between the videos. We hope this project can be useful and convenient to use by different users.

This is the project link: [Find Your TED](#)

This is the screen-cast link: [Find Your TED](#)

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Contents

I. Overview and Motivation

II. Related Work

III. Question

IV. Data

I) Source

II) Pre-process

V. Exploratory Data Analysis

VI. Design Evolution

I) Inspiration

II) Sketch

VII. Implementation

(I) Tryout

(II) Layout

(III) Interaction

(IV) Improvement

VIII. Evaluation

I) Learn from data

II) Future work

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Overview and Motivation

In application of data visualization, we would like to find a place that is greatly in need of such a solution. Then, we came across a TED talk dataset in Kaggle, where it shows a series of topics of a particular video as well as a list of correlated videos with a given video. Then, we are thinking of targeting this project towards people with identity as scholars, people with free time and children. It is because these are the group of people that may have the potential to want to grow and have enough free time to spend on watching videos. Although in general TED talks are aimed at targeting a huge variety of different groups of people, our application needs to first start from a more narrower place and generalize the usage gradually.

TED talks are popular for the inspirations of its speakers and lecturers' experiences. In order to provide a good solution for people to get more inspired by more TED talk as well TED talk topics, we would like to provide a product containing data visualizations that explores the interconnection between each different TED talk as well as topics. After considering the dataset from Kaggle, which provides detailed information about TED talks, we have put together a web application using D3.js to proceed with exploration of such relationships within the topic and video data.

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Overview and Motivation

In need of providing our end users with a well functioning application. We would like them to first view the distribution of existing topics in TED talks. It is because, with different topics, users with different interests may want to explore differently. Therefore, showing the topics beforehand would be necessary. Thus, we need to come up with a proper solution for visualizing the topic distribution. It would be professional if we choose to use bar charts for showing this or a table. However, making the visualization fun and interesting would motivate our users to use more of our product. Thus, a bubble chart without an axis becomes our first choice. Then, since we have over 400 existing topics, we have chosen to only limit the display to 20 topics to avoid overwhelming our users.

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Overview and Motivation

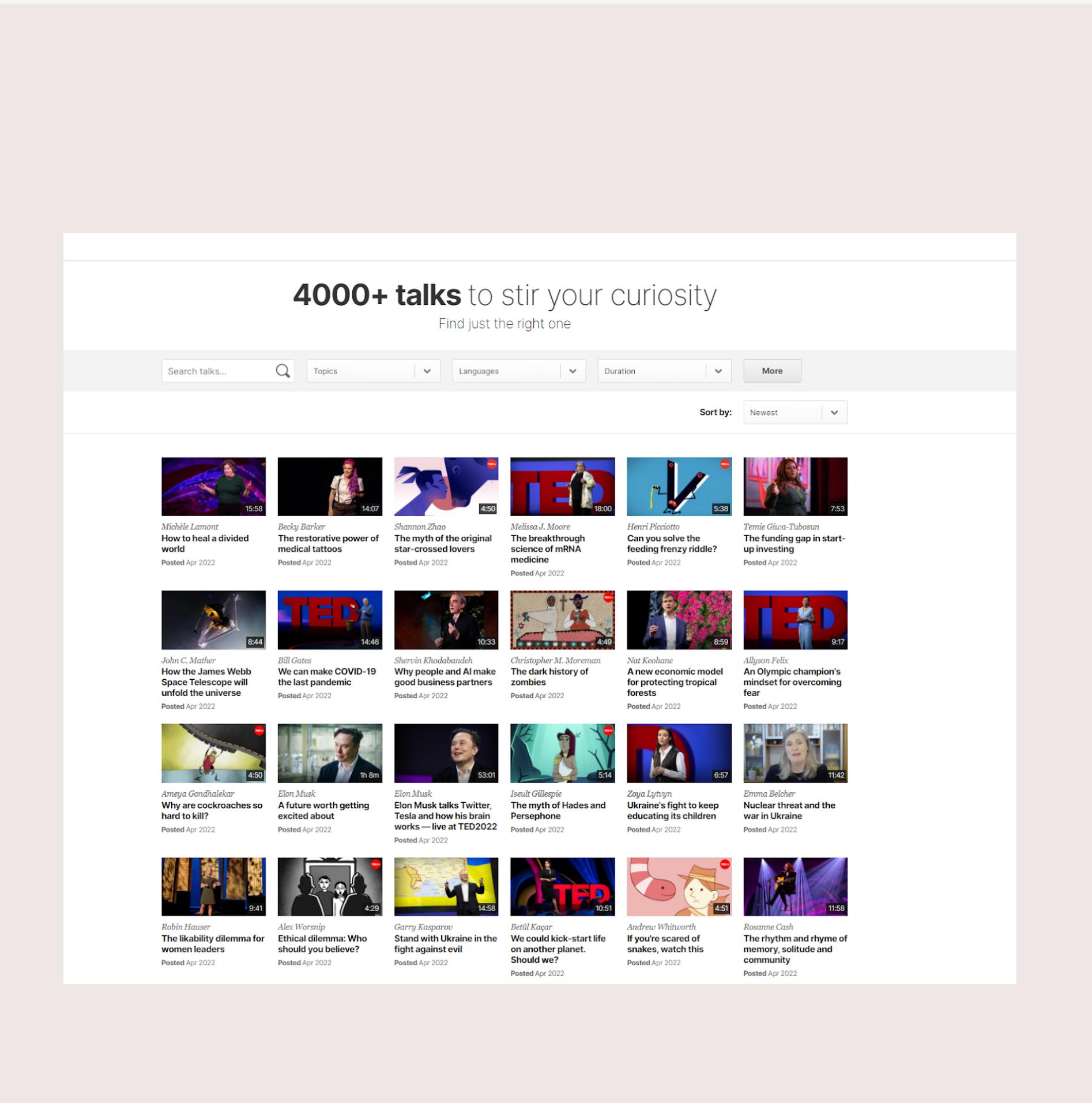
After exhausting all existing choices of visualization tools, we have decided to use network graphs to visualize the relationship between video and video. In a general study, when people try to view a video, they would turn to related videos or similar recommended videos for further exploration. Thus, using network graphs provides us with the advantage of providing our end users with a solution to look into related videos and interconnections. Moreover, there is an idea named “knowledge graph,” the graph that provides people a way to trace their minds from one node to another. network graphs are born by this notion. Thus, it makes sense that we connect this idea of visualization with our idea of providing end users a click away from their related videos. Furthermore, once a user lands onto a related video, he or she needs to explore more about the new video, we then need to design a feature of refreshing the network graph after clicking on one of the videos.

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Related Work

The first inspiration we go to is the official TED talk website (<https://www.ted.com/talks>). It is because this website provides all existing TED talk videos, and has carried out what the organizers of TED talk think the TED talk website should be looking like. However, this website is more of a collection of videos rather than interacting with the audience. In the website, all videos are listed in a card-like component with details like a cover image, speaker, title and date created. It still provides a very similar functionality which all video websites would provide: a header bar, a search engine with filter options, as well as signing in options.



CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Related Work

In addition to that, we examine the official website for how its topics are covered and depicted. On the upper right corner of the page, the website provides a button where we could explore all existing topics in a list view, sorted and visualized alphabetically.

Topics

Browse TED topics from A to Z

A

Activism	AI	Ancient world	Anthropology	Asteroid	Augmented reality
Addiction	AIDS	Animals	Archaeology	Astrobiology	Autism spectrum disorder
Africa	Algorithm	Animation	Architecture	Astronomy	
Aging	Aliens	Antarctica	Art	Atheism	
Agriculture	Alzheimer's	Anthropocene	Asia	Audacious Project	

B

Bacteria	Best of the Web	Biology	Biotech	Body language	Brazil
Beauty	Big Bang	Biomimicry	Birds	Books	Buddhism
Bees	Biodiversity	Bionics	Blindness	Botany	Bullying
Behavioral economics	Bioethics	Biosphere	Blockchain	Brain	Business

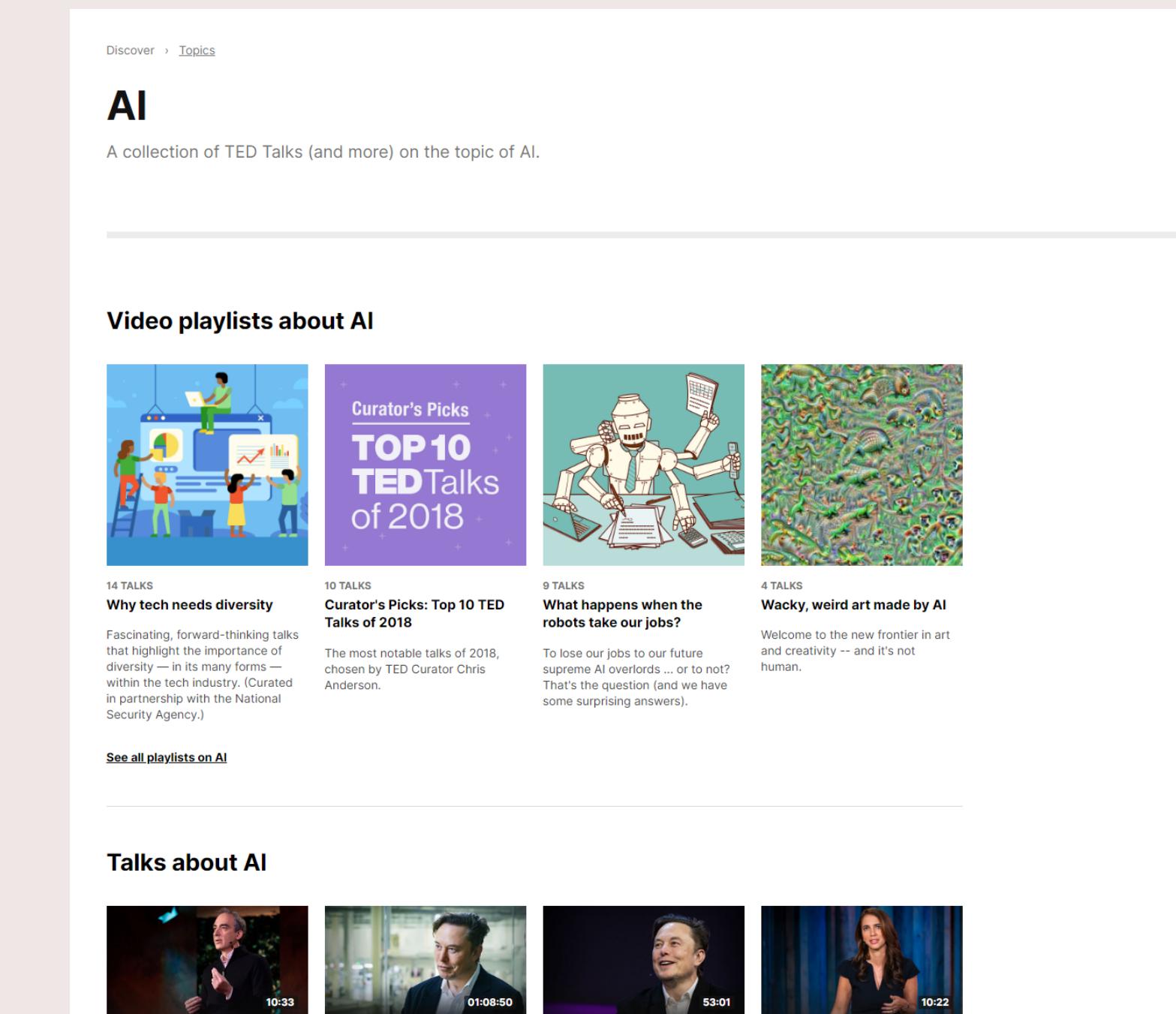
C

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Related Work

After entering each topic, the website is separated into two sections: all existing video lists about the topics, and all videos about the topic. All these designs are professional and useful for presenting to the end users about the contents TED talks have. Also, being clear about the contents help the users to navigate through the website easily and get what they need fast. But, this presentation missed the point of a website, where it should not only provide clarity, but also drag the audience into the topic and motivate them to explore more of the contents such that they could be inspired more.

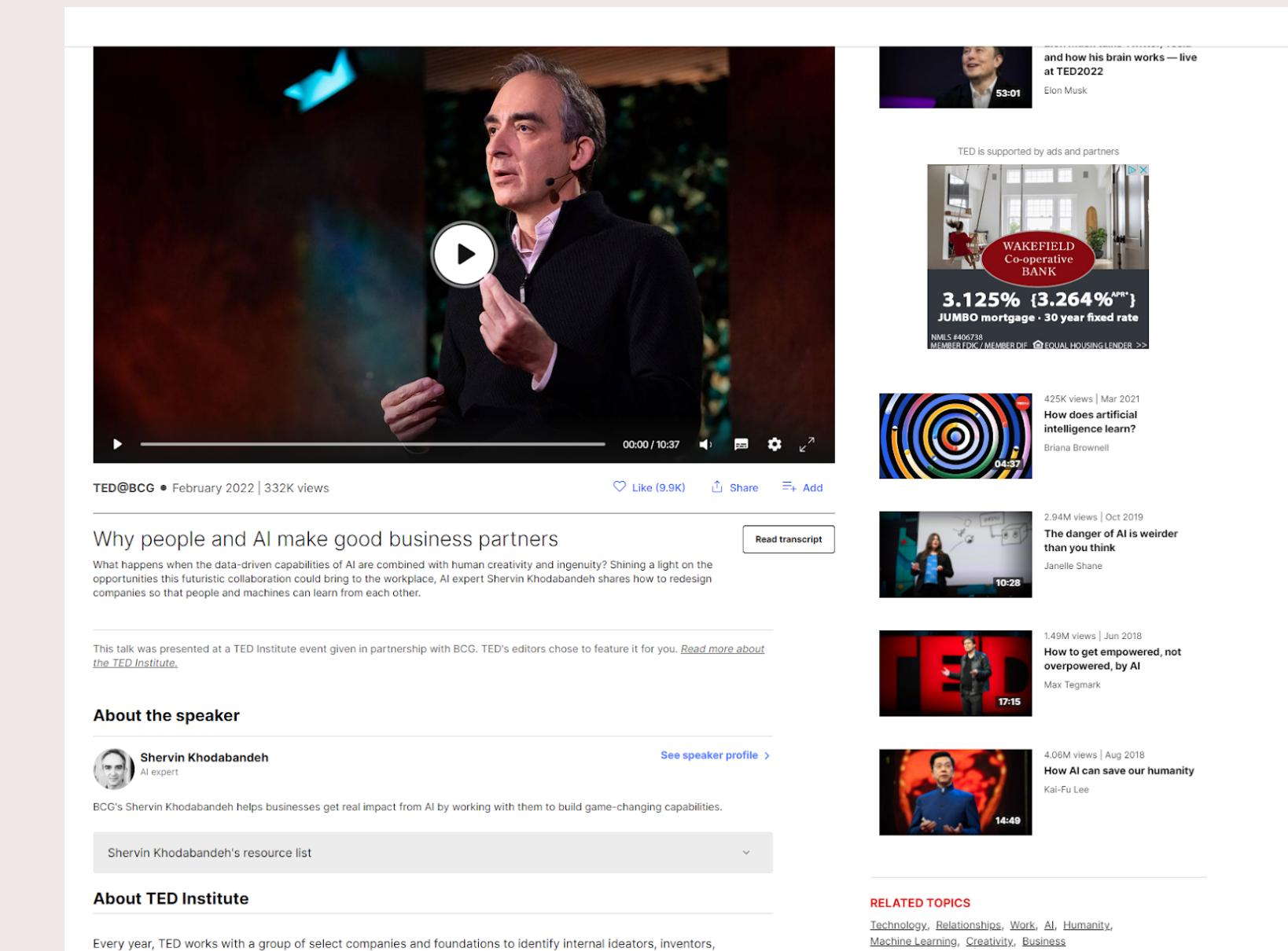


CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Related Work

Then, going into a particular video, there would be a video page and a list of related videos, as well as a related topics section, just as one would find in a normal video website. The video page contains information about the video title, video description, speaker information as well as likes for that video. This distribution helps people understand the video better before watching it.



CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Related Work

Then, in order to inspect a proper visualization to inherit existing features of the official website as well as taking it to a next level, we have examined the idea of how network graph helps people think about different perspectives. There are already existing literature explaining that using network graphs could be very influential on helping people make decisions upon peer selection ([Using Social Network Graphs as Visualization Tools to Influence Peer Selection Decision-Making Strategies to Access Information About Complex Socioscientific Issues](#)). Thus, we could consider a similar approach when people try to select videos. Since selecting a video to watch would be similar to finding an appropriate friend to talk to. By clearing out the network relationship between each different video, we could improve the choices and decisions made by our videos in selecting a proper next video to watch.

Furthermore, a bubble chart could be used effectively in describing comparisons between multiple fields within same category ([APPLICATION OF INTERACTIVE CHARTS IN THE EVALUATION OF SOCIO-ECONOMIC DEVELOPMENT OF REGIONS; THE CASE OF POLAND](#)), especially in economics. However, video selection could also be viewed as such a selection. It is because we could see we are all selecting a particular field from the large category of all existing topics when we are choosing a particular video to watch.

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Questions

- **What questions are you trying to answer?**

In order to complete the idea of providing people with a better website for exploration of TED Talks, we need to spot what current TED Talk website is lacking, and what we could do in addition to the existing version of the website to motivate people more to explore the talks. We from there have spotted that there's no very good solution for topic exploration, and need to find a proper way to present all existing topics and the importance of the topics in TED talks to our end users.

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Questions

- **How did these questions evolve over the course of the project?**

After learning about different tools as well as reading more papers during the course session, we have discovered that there would be some solution to the essential question of presenting an interesting page for end users of TED talk audience, which is to embed data into interactive visualizations written by D3.js. However, there are more explorations awaiting. Like how we could enable proper interactivity, and how much interactivity would be just the right amount for us to include in the web application.

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Questions

- **What new questions did you consider in the course of your analysis?**

After obtaining the data from Kaggle about TED Talks, we have ran into the obstacle, where the data is not properly formatted. Using it for data visualization is going to be time and resource intensive. Thus, a posing question would be what is the best format for us to transform the data into.

Since we have decided to answer the most essential question with data visualization, we would like to find out the best choices of visualization tools we need to use as well as knowing the design choices we are going to make, for instance color schema, size of the visualizations, position of the graphs, etc. Another posing question related to visualization is that code needs to be written to implement our ideas. A proper coding language as well as a library is in need. Then, using that language properly would be an important step for us to take.

Therefore, the project requires us to read more papers and explore more existing tools to provide solutions.

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Data

- **Source**

The main data we need is the basic information about TED talk videos, including the titles, the speakers, the published date, and the number of views, etc. This part is for the information display.

In addition to the basic information data, we also need information about the topic of the video in order to better classify the videos and display the visualizations. To show the related networks between videos, we'd better find the information about related videos.

Consider all our initial needs about the data, we found this dataset can basically cover all the information we are seeking for. Here is the link of the dataset: <https://www.kaggle.com/datasets/rounakbanik/ted-talks>.

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Data

- **Pre-process**
 1. This dataset contains a lot of information that we don't need and we need to simply make information selections from it. We have selected 7 columns of the dataset, including title, main speaker, related talks, tags, url, views, and published date.
 2. In this dataset, one video can have multiple topics. There is one column called tags that contains all the topics that the video covers. We extracted all the topic words and counted the number of all videos involving the topic.
 3. To form the related videos network, for each video, we extracted all the titles of its related videos.
 4. We convert all the required data into the data structure needed for the graph used.

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Exploratory Data Analysis

We performed a preliminary analysis of the dataset using Colab.

For the topic part, we found a total of 416 different topics and found that a video will have about five or so related topics. And also, these topics are also very representative, as can be seen from the number of videos they contain.

For the related videos part, a video will also have about five or so related videos.

	description	topic
0	Sir Ken Robinson makes an entertaining and pro...	children
0	Sir Ken Robinson makes an entertaining and pro...	creativity
0	Sir Ken Robinson makes an entertaining and pro...	culture
0	Sir Ken Robinson makes an entertaining and pro...	dance
0	Sir Ken Robinson makes an entertaining and pro...	education

	topic	num_talks
0	technology	727
1	science	567
2	global issues	501
3	culture	486
4	TEDx	450

	description	related
0	Sir Ken Robinson makes an entertaining and pro...	Bring on the learning revolution!
0	Sir Ken Robinson makes an entertaining and pro...	How to escape education's death valley
0	Sir Ken Robinson makes an entertaining and pro...	How to fix a broken school? Lead fearlessly, I...
0	Sir Ken Robinson makes an entertaining and pro...	Education innovation in the slums
0	Sir Ken Robinson makes an entertaining and pro...	A short intro to the Studio School

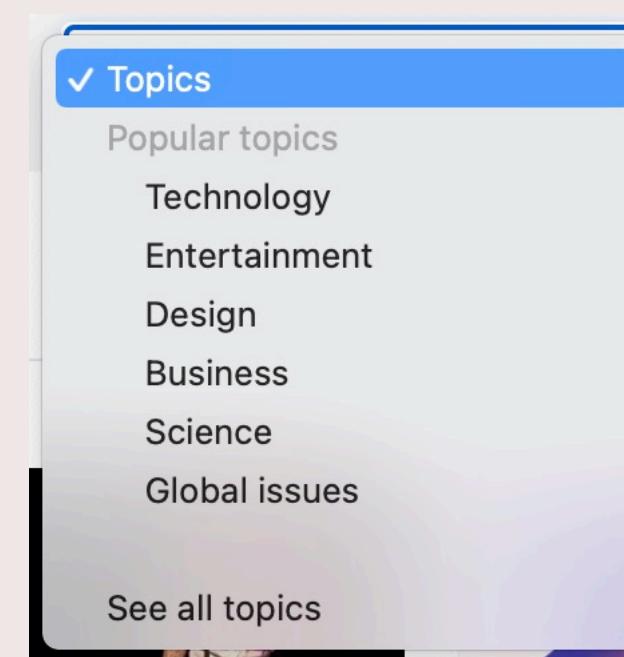
CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Design Evolution

- **Inspiration**

For the topic part, we went to the ted website to see how they represent the topic categories. In fact, generally speaking, websites don't choose to present categories in a visual way; they generally present all categories in a more concise, comprehensive way:



Topics			
A	B	C	D-E
Activism	Algorithm	Anthropocene	Astrobiology
Addiction	Aliens	Anthropology	Astronomy
Africa	Alzheimer's	Archaeology	Atheism
Aging	Ancient world	Architecture	Audacious Project
Agriculture	Animals	Art	Augmented reality
AI	Animation	Asia	Autism spectrum disorder
AIDS	Antarctica	Asteroid	

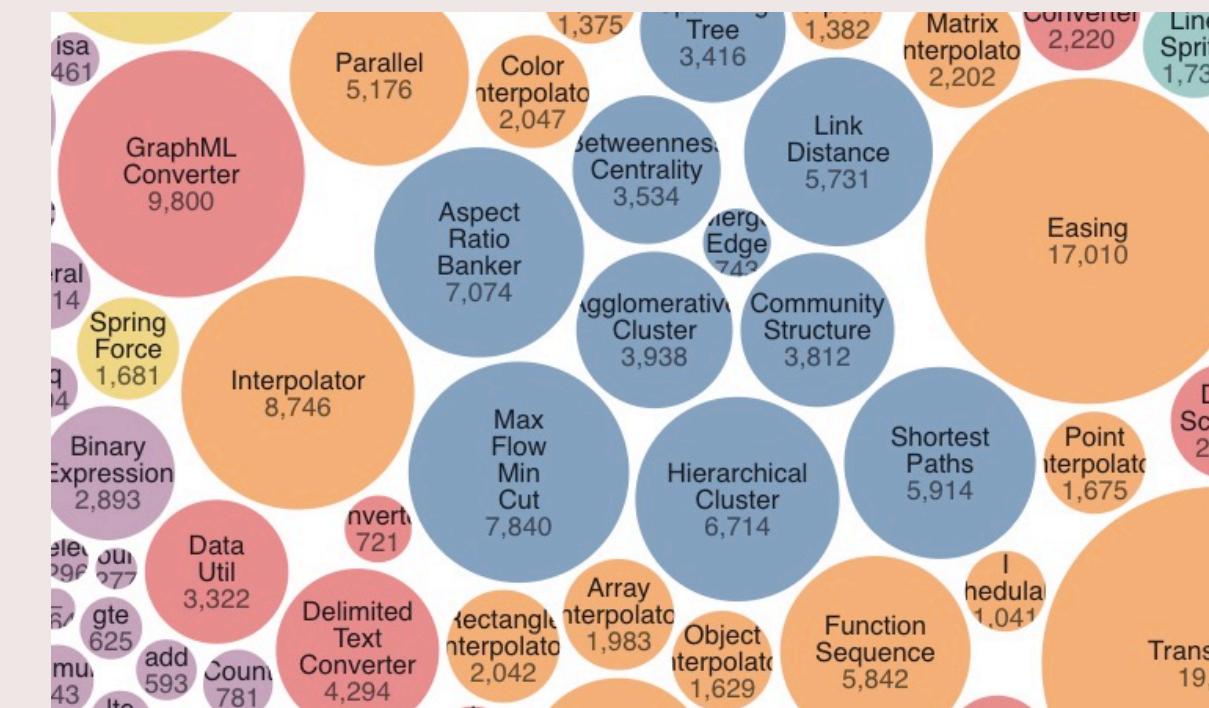
CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Design Evolution

- **Inspiration**

When it comes to representing the categories of data through visualization, the first type of visualization that comes to our mind is the graphs with dots.



So we decided to use bubble diagram to show the classification of topic display.

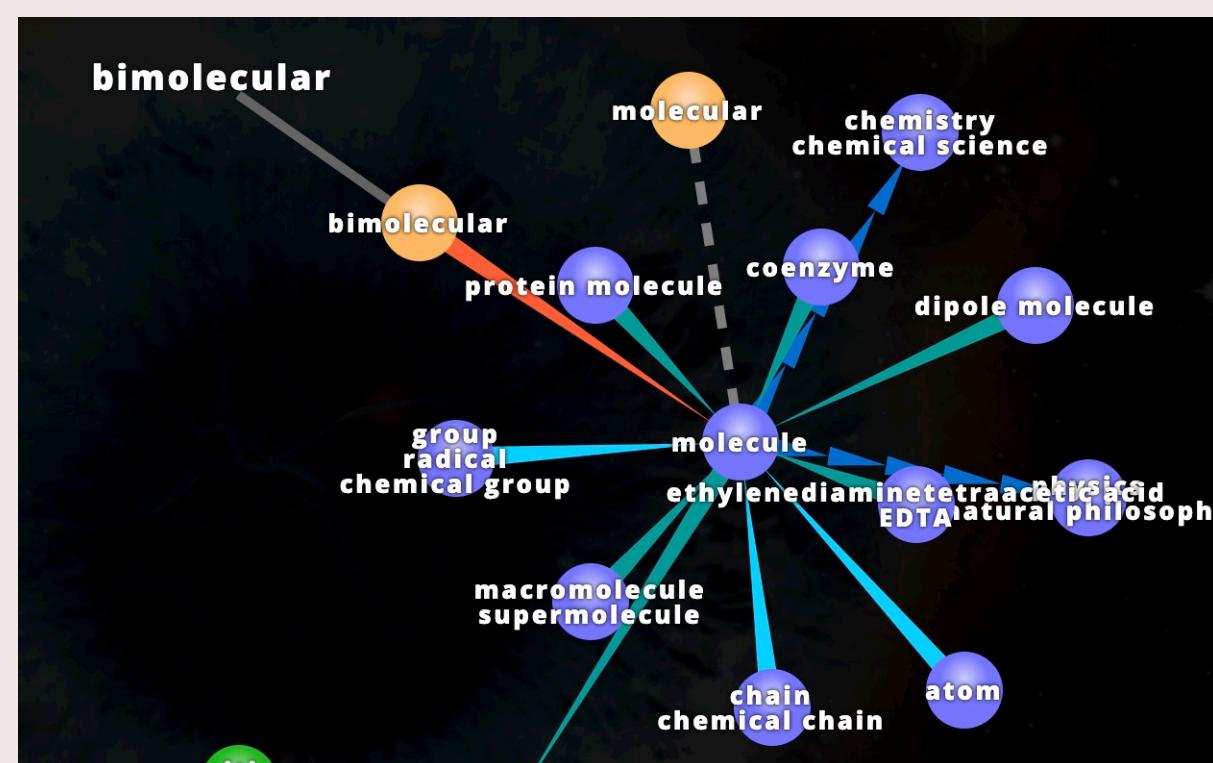
CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Design Evolution

- **Inspiration**

For the related videos part, there is an interesting project named VISUWORDS which shows the relationships between words that gives up inspiration.



Initially we wanted to create a complete network diagram showing the connections between all the videos, after the feedback of the proposal, we were concerned that an excessive number of nodes would lead to poor utility, so we decided to show each video separately for its associated video, which would involve a smaller number of nodes and be more functional.

From the previous data analysis, we learned that the number of associated videos for each video is small, and the connection is relatively simple, so we initially decided to use network graph to visualize the connection between videos.

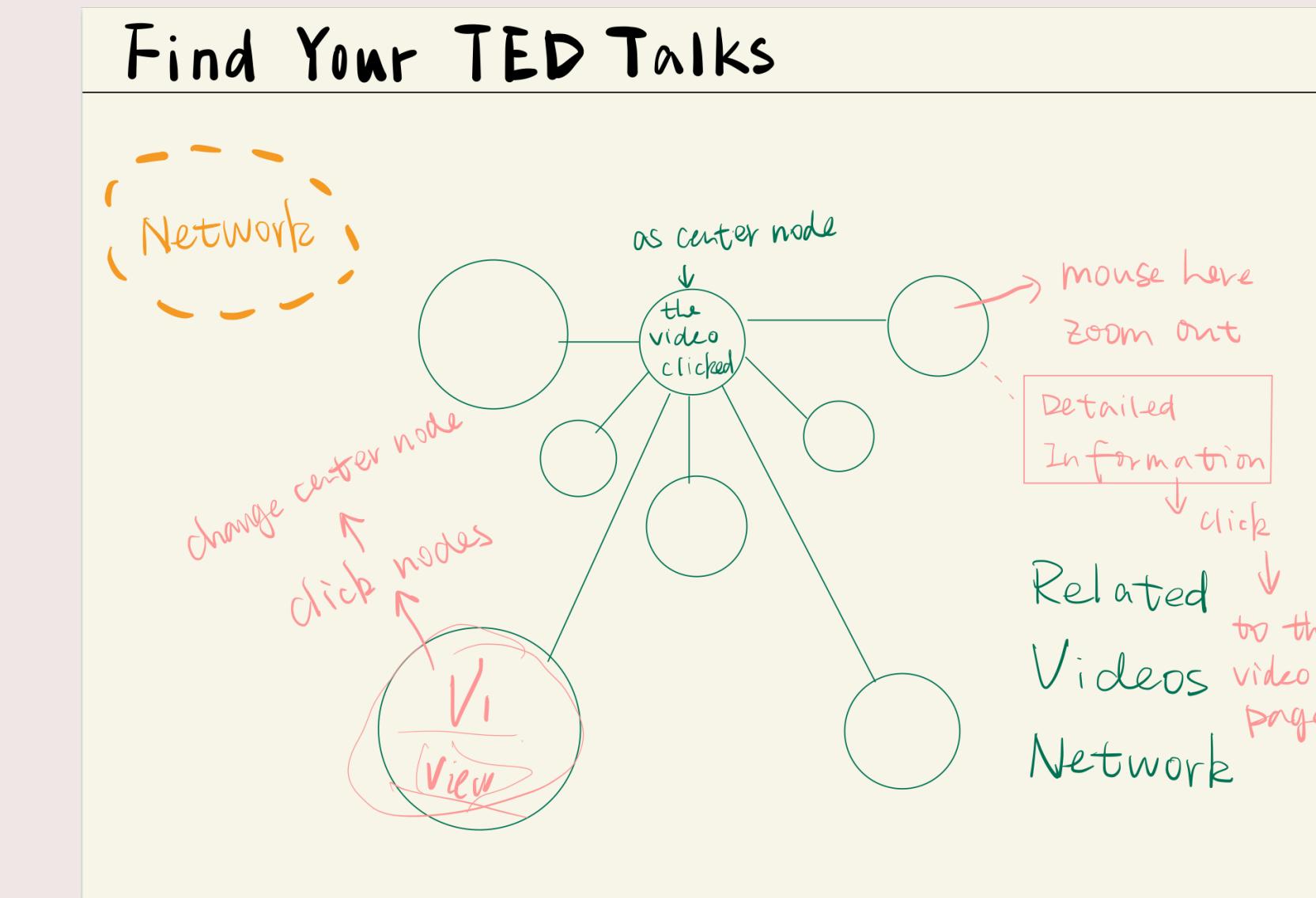
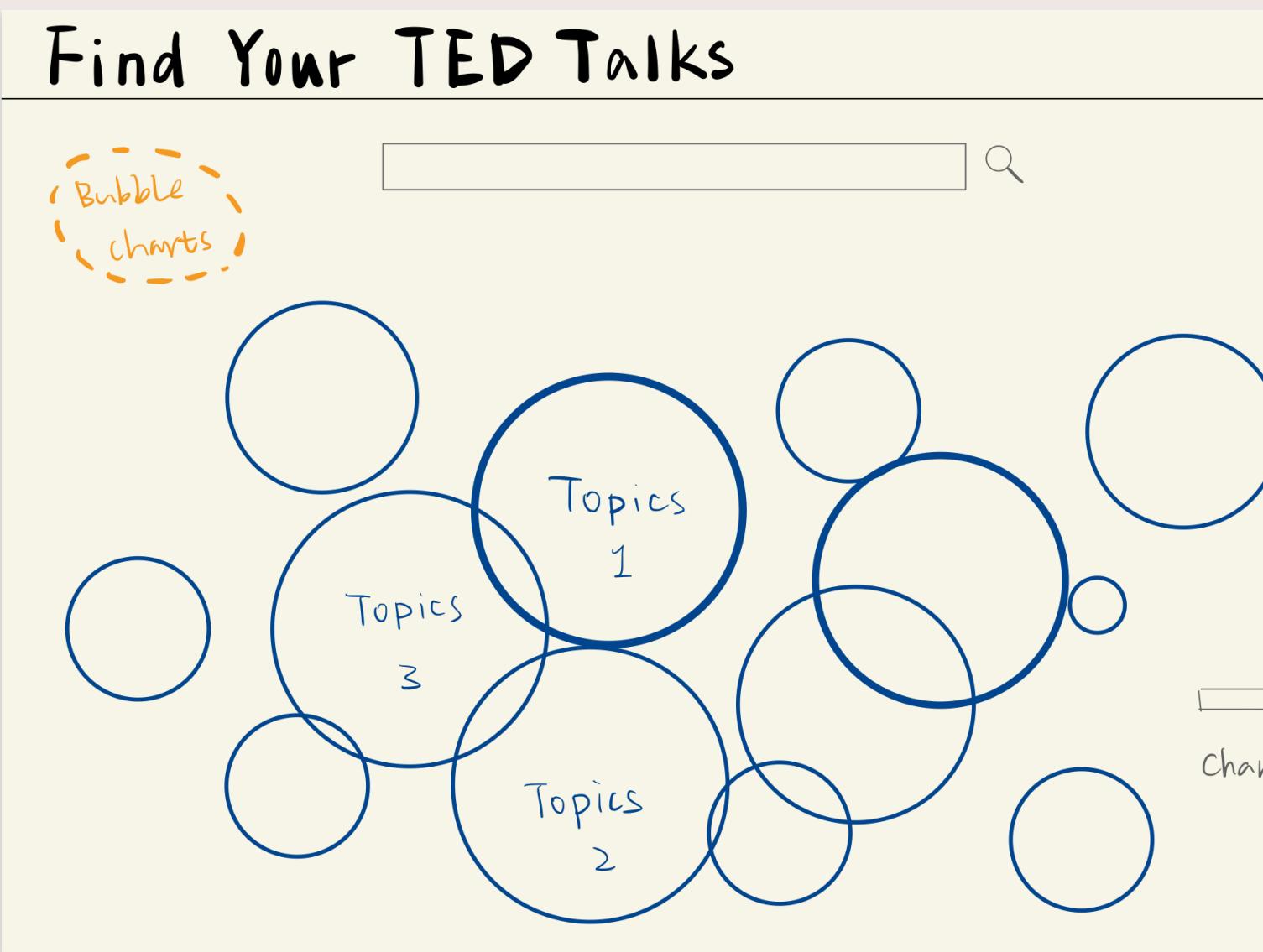
CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Design Evolution

- **Sketch**

These are some sketches of our preliminary design.



CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Implementation

- **Tryout**
 1. In terms of implementation techniques, we first tried a combination of Vue and d3. We encountered the following problems. We wrote the Vue framework content and the d3 diagram separately, but found that the pure d3 diagram was written in a way that was not compatible with the Vue framework during the concatenation phase. We can't solve all the problems simply by just using these two technologies, so we started to consider simply using Vue's d3 package or abandoning the use of Vue framework.
 2. We started writing websites in the traditional way, but we were not very familiar with it so we took a lot of wrong turns. For example, Vue's list function is well encapsulated, but using html and javascript to write a dynamic list from scratch made us have a lot of display problems.

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

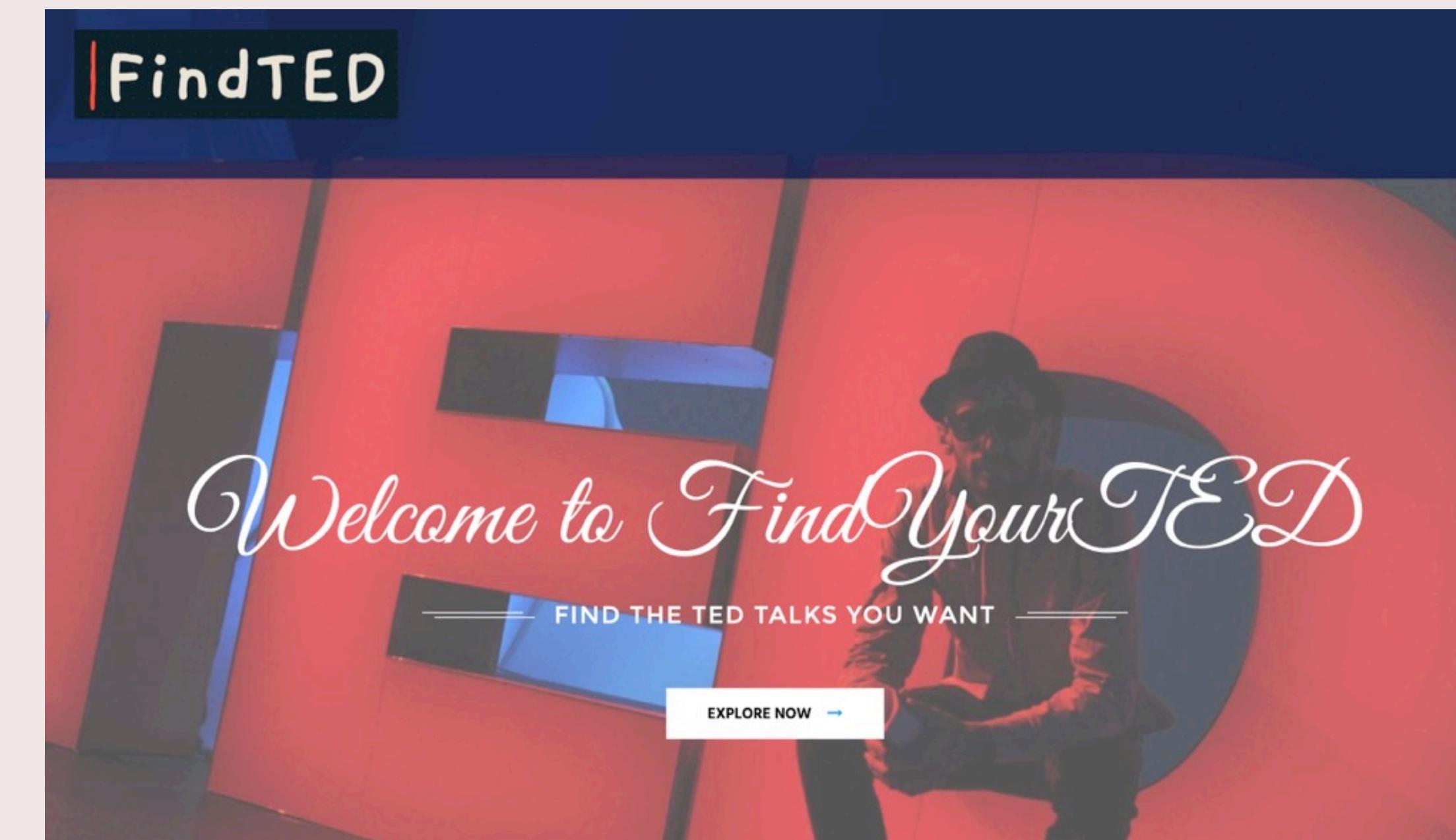
Implementation

- Layout

This is our main page.

The function of this page is to show the theme of our project.

Click on the button that says EXPLORE NOW to enter the function page.



CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Implementation

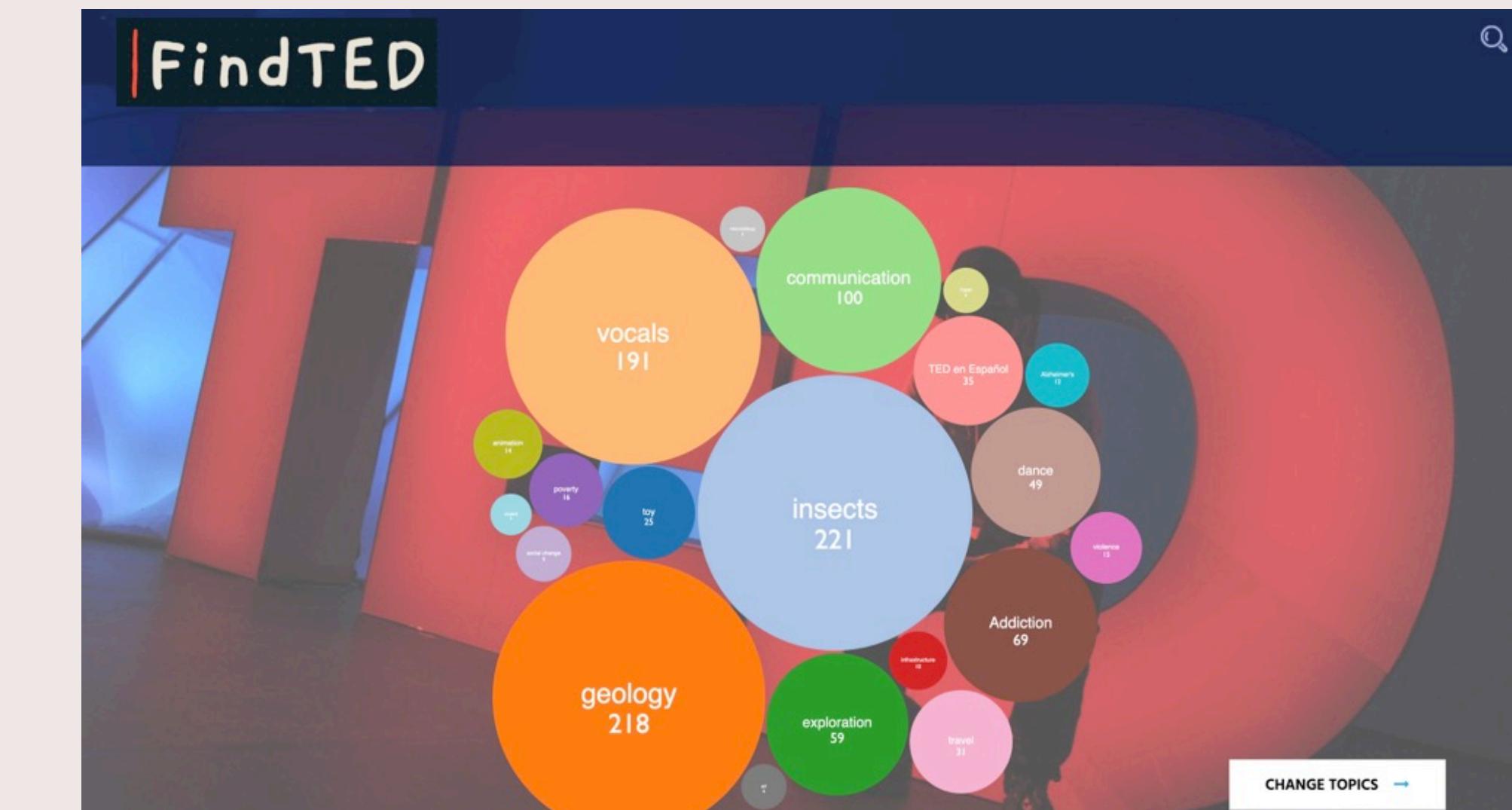
- **Layout**

This is the topic page.

The main body of the screen is a number of bubbles, and the bubble shows the name of the ted talk topic and the number of its videos.

The top right corner is the search function for topic.

The button in the lower right corner allows users to change the topics displayed.



CS573 - Final Project - Find Your TED

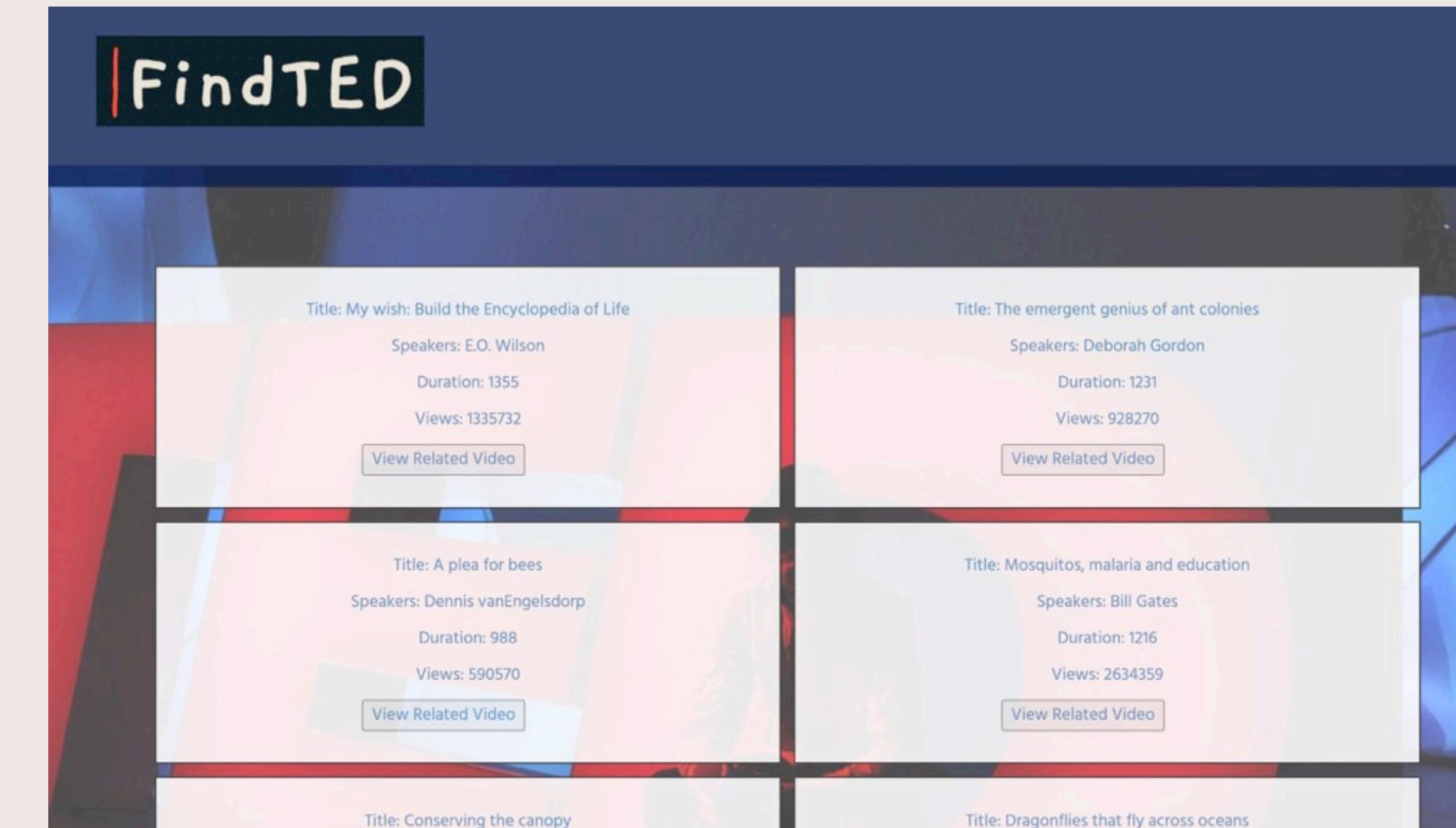
Mingjie Zeng & Yufei Lin

Implementation

- **Layout**

This is a simple list page, each item shows the title, speaker, duration and number of views of the video in this topic.

Clicking the button inside the item will display the network page for the video and its related videos.



CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Implementation

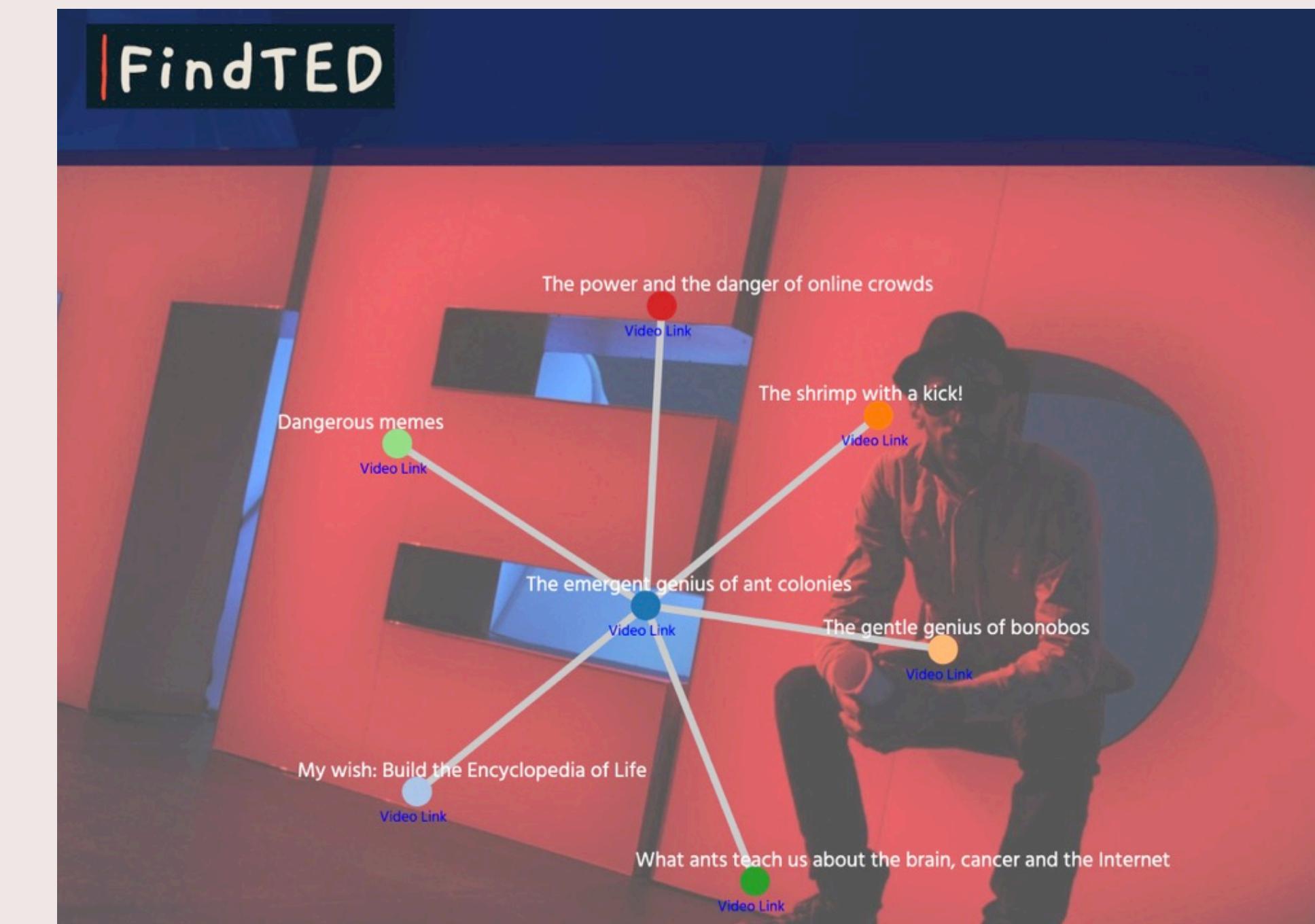
- **Layout**

This is the related videos page.

The middle node represents the video clicked by the user, and the children nodes represent the videos related to the video clicked by the user.

When you click the children nodes, the center node will change and its related videos will change too.

Clicking the video link can direct users to the TED video page.



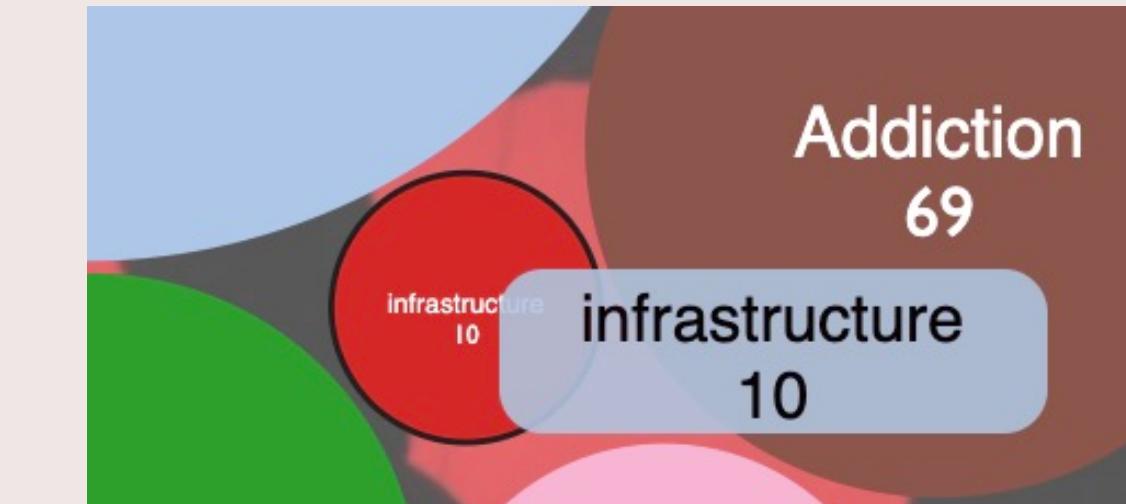
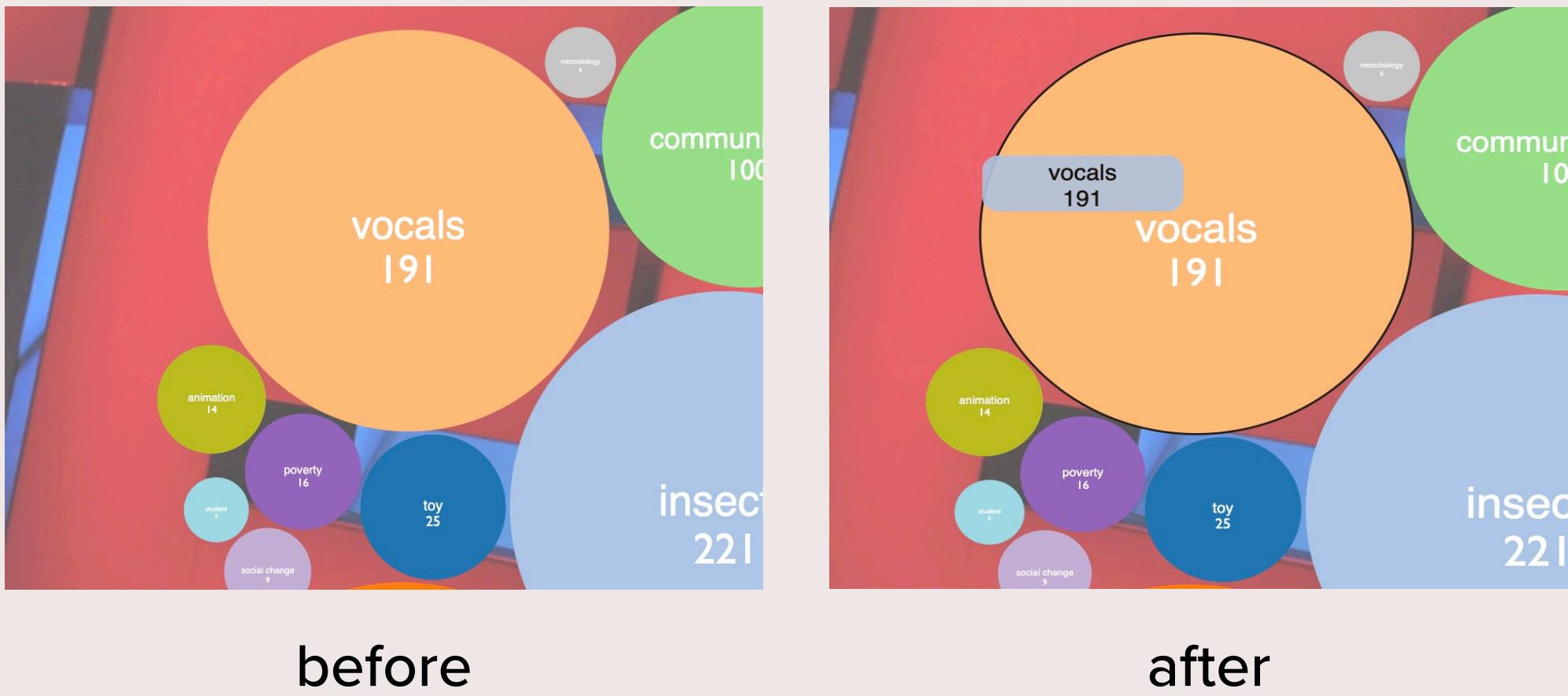
CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Implementation

- **Interaction**

When users hover over the bubble, the bubble will appear with a black border and a hover box to display the specific text content on the bubble. This helps users to see the contents of the bubble, especially when the bubble and its text is displayed in a very small scale.



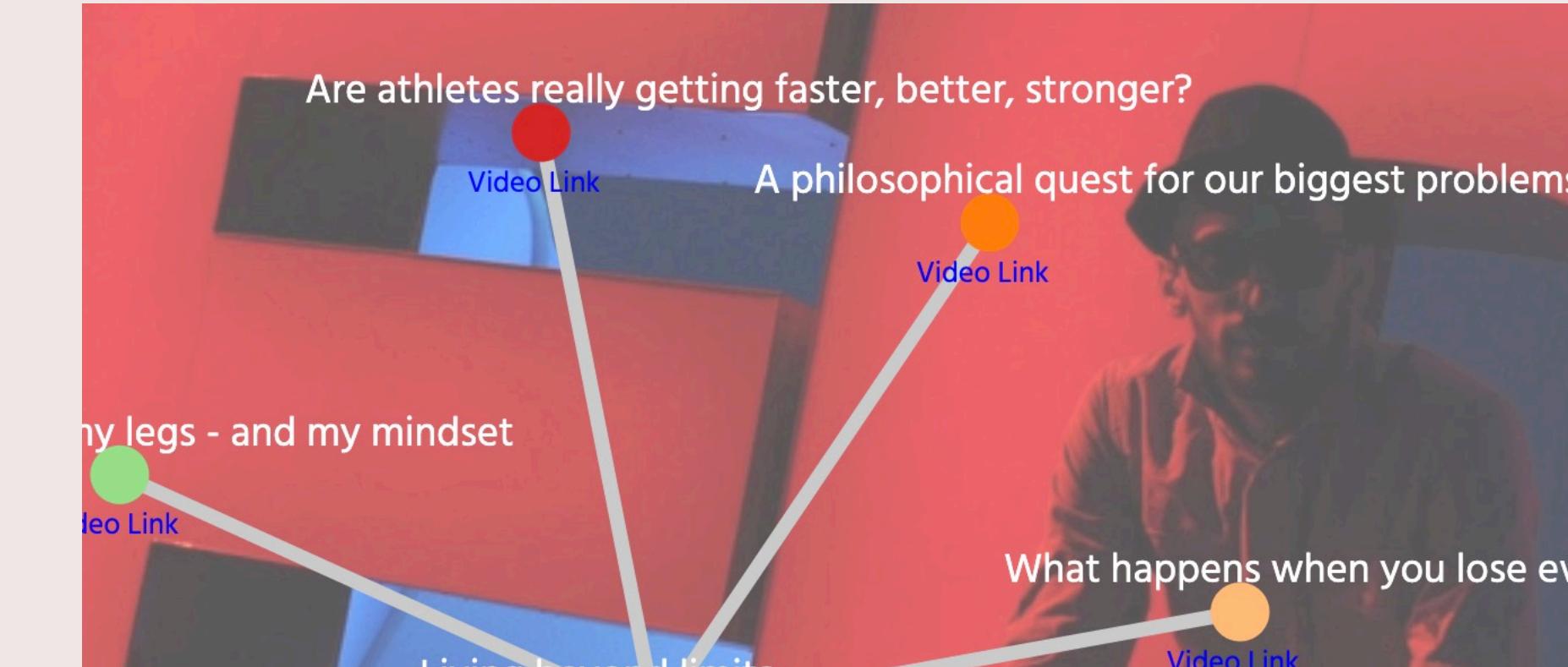
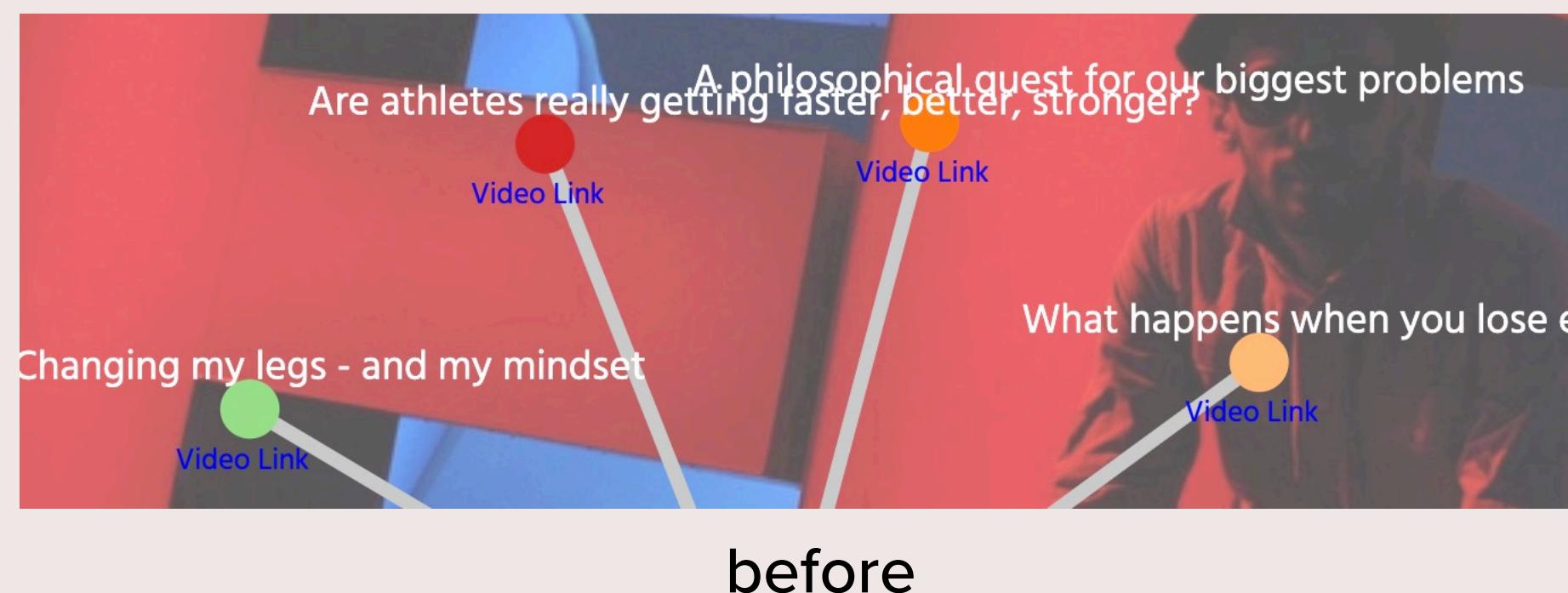
CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Implementation

- **Interaction**

The user can click on a node to drag it around to change its position. This feature is mainly to solve the problem that the titles will be blocked from each other due to the dynamic generation of the network diagram and the possibility of long video titles. If the occlusion occurs, users can drag the nodes to get a better experience.



CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

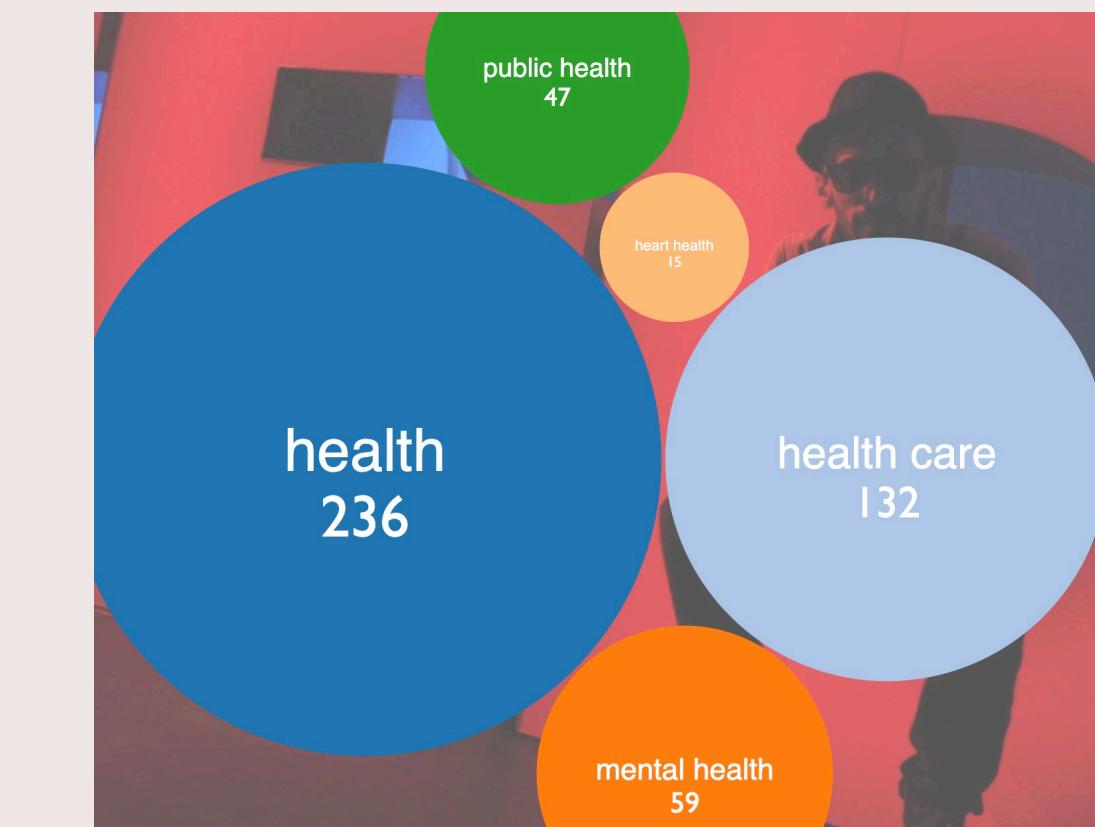
Implementation

- **Improvement**

At first, we used the exact search for topic, and the result page only had a bubble, which was very inefficient and unattractive. Then we turned the search into a fuzzy search, where users can find some related topics and the results are displayed in a more harmonious way.



before



after

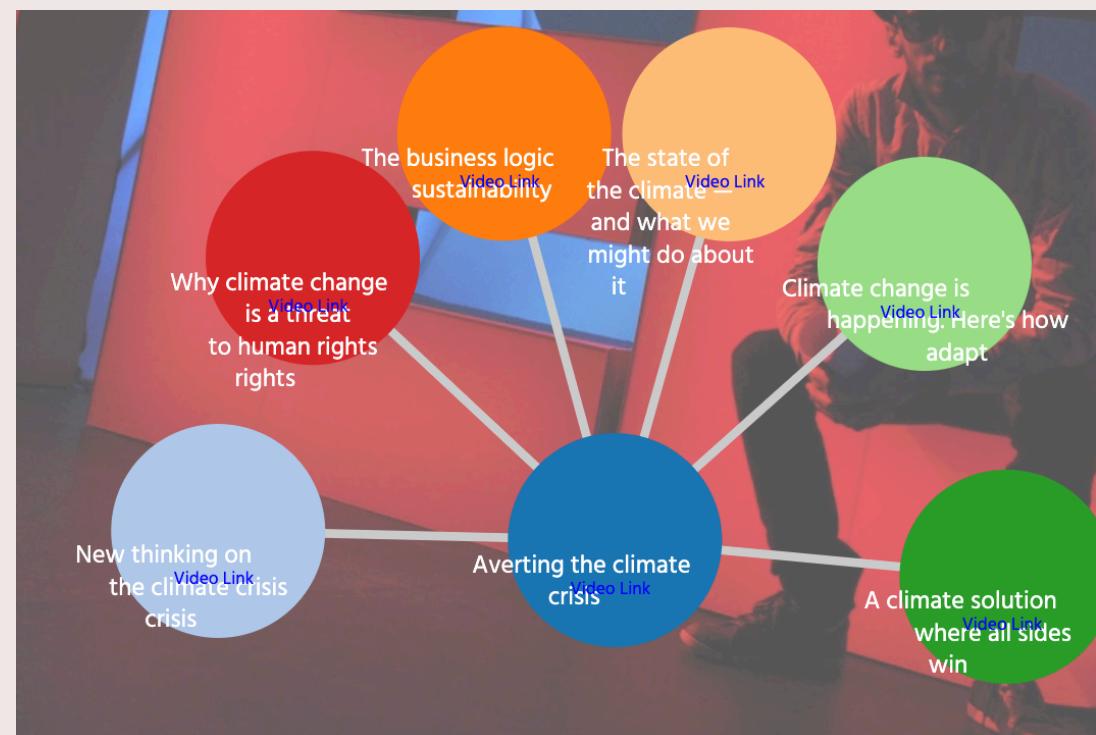
CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

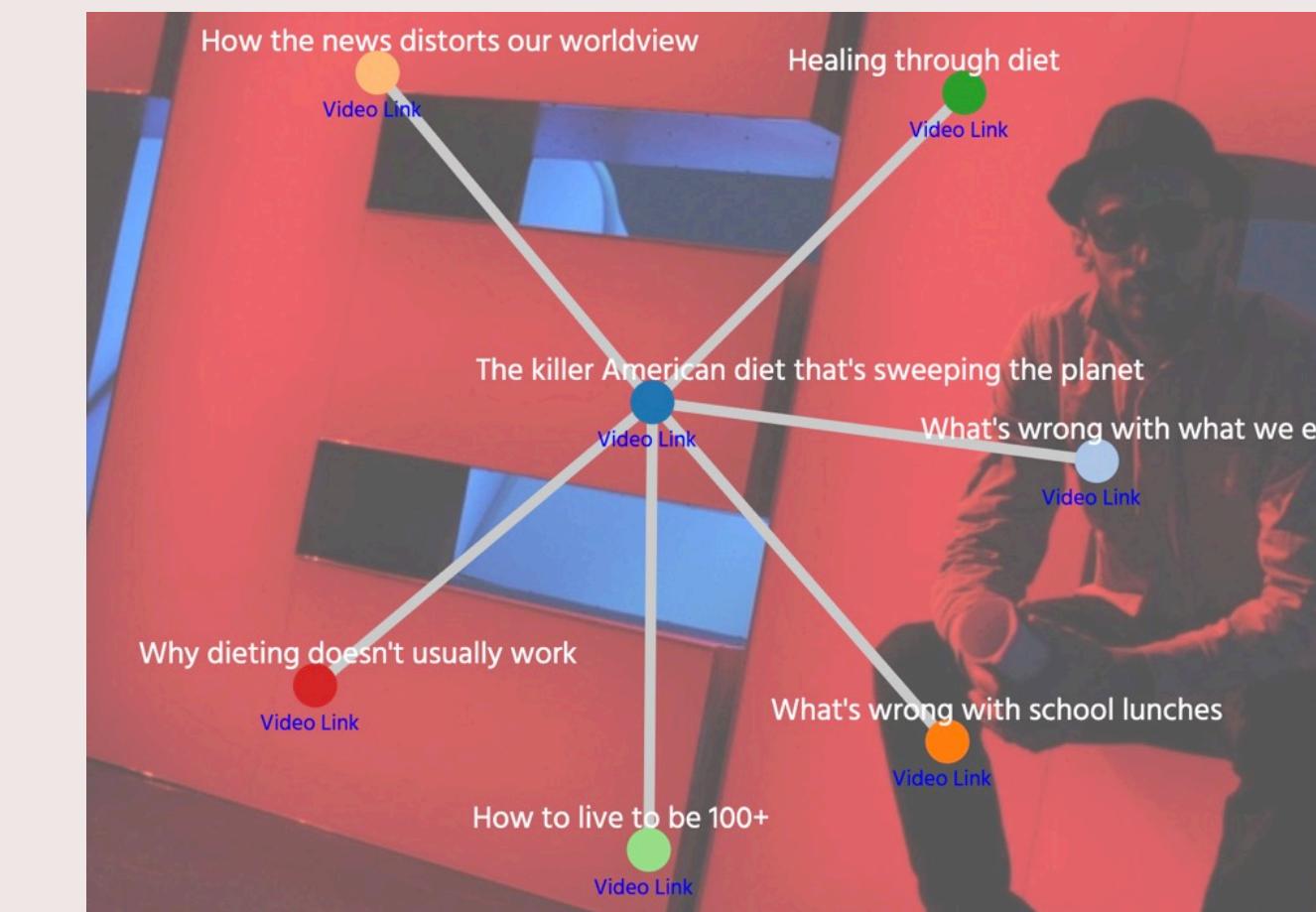
Implementation

- **Improvement**

At first, we used the common network graph style, which is large nodes, but it was not harmonious to display. So we tried to reduce the existence of the nodes and highlight the title content more.



before



after

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Evaluation

- **Learn from Data**

What interests us most about this project is the connection between the data. Data such as video titles, for example, are difficult to represent their relationship in other ways, but data visualization directly presents the connections between them to users. This is one of the amazing things about data visualization.

This project is more focused on how to better present the data and make it more useful than exploring the pattern or distribution of certain data. The same data can be visualized in different ways for different purposes, and finding the right presentation is a key point of data visualization application.

In addition, in the application field, not all data can be visualized directly, in order to achieve the data format requirement of some kind of visualization, such as network graph, it is necessary to reconstruct the existing data, which is also challenging for applications with a lot of data.

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

Evaluation

- **Future Work**

Due to our lack of experience, this project is still a very preliminary work and there is still much space for improvement.

The most obvious deficiency is the lack of design, we hope we can find a more elegant layout to make this website more attractive. In addition, we only covered some basic information data of the videos in this project, we are also thinking whether we can dig deeper into the data, such as determining the hot topics by counting the proportional change of video topics over time, and then, for example, linking more videos together through a network of connections between speakers.

Also, according to what we envisioned at the beginning, how to determine the connection between videos through data rather than with the help of external information, and how to really build a video relationship network graph involving all ted videos, we will also continue to think about it and hope to improve this find your ted project again in the future.

CS573 - Final Project - Find Your TED

Mingjie Zeng & Yufei Lin

The End
Thanks for Reading!

May 2, 2022