

# Visualization of the Chinese Censorship Events Data from Weibo in the Notalert Website

# Joseph Wang, Jedidiah R. Crandall

Department of Computer Science, University of New Mexico

## INTRODUCTION

In this project, the purpose is to visualize the Chinese celebrities censorship event data from [1] Sina Weibo website in the [2] Netaalert website by using the timeline library. The timeline is used to visualize the event data so people can read these events easily. If the user is interested in a specific celebrity or person, the user can click on the celebrity name in the timeline and look through his or her previous events. This Chinese censorship event data was collected over 3 years by us which was censored by the Chinese government and was being taken down by [1] Sina Weibo. The censorship data included a total of 319 celebrities personal basic information and their events, these celebrities are classified into 24 groups base on their job titles. However, group 6, Economist 经济学家, group 13, Monk 僧, and group 19, Real estate tycoon 房地产大亨, doesn't have any event data that exists in the weiboscope-censored.CSV so these three groups were excluded. This paper is organized into seven sections. In section II we present the definitions of the important terms that are mentioned in the abstract and the introduction section. In section III we show the methods and outline that we performed to visualize the censorship data. In section IV we present the result of our CSS webpage and the result of the timeline to display our data. In section V is the suggestion for the researcher and wrong information in the CSV file that needs to be improved. In section VI is the conclusion of this project and the things that we observed from the censorship data.

## CELEBRITY GROUPS IN THE CHINESE CENSORSHIP EVENTS DATA

**There are a total of 24 celebrity groups which are**

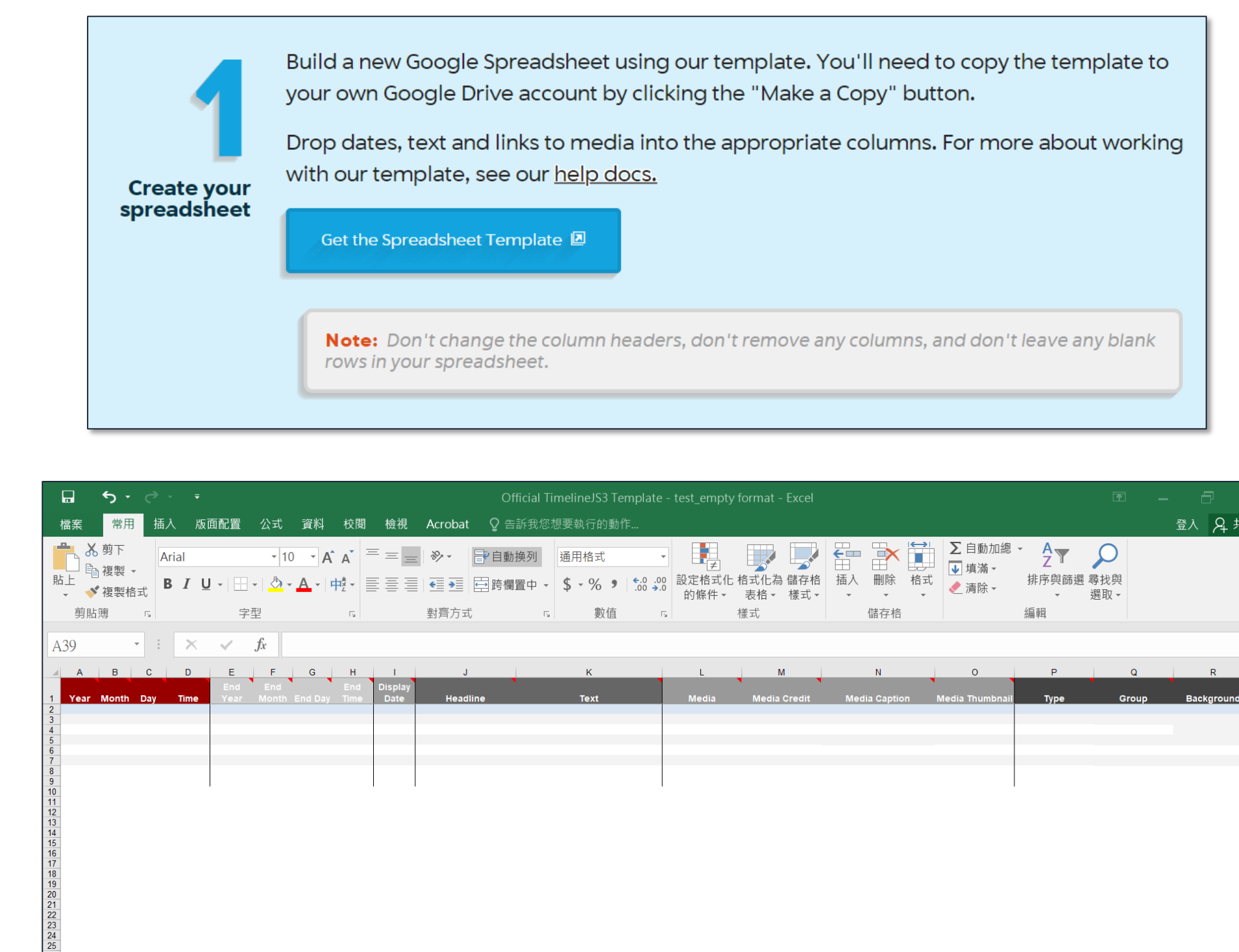
1. Activist 活动家
2. Artist 艺术家
3. Athlete 运动员
4. Businessman/Businesswoman 商人
5. Citizen 公民
6. Economist 经济学家
7. Fictional character 虚构人物
8. Foreign politician 外国政客
9. Human rights lawyer 人权律师
10. Journalist 记者
11. Media 媒体
12. Military general/leader/officer 军事将领/领导/ 官员
13. Monk 僧
14. Pastor 牧师
15. Philosopher 哲学家
16. Police officer 警官
17. Politician/CPC member 政治家/共产党员
18. Professor/Scientist 教授/研究员
19. Real estate tycoon 房地产大亨
20. Soldier 士兵
21. Taoist 道
22. Teacher 老师
23. Unknown 未知
24. Writer/Essayist/Novelist 作家/圣手/小说家

## METHOD

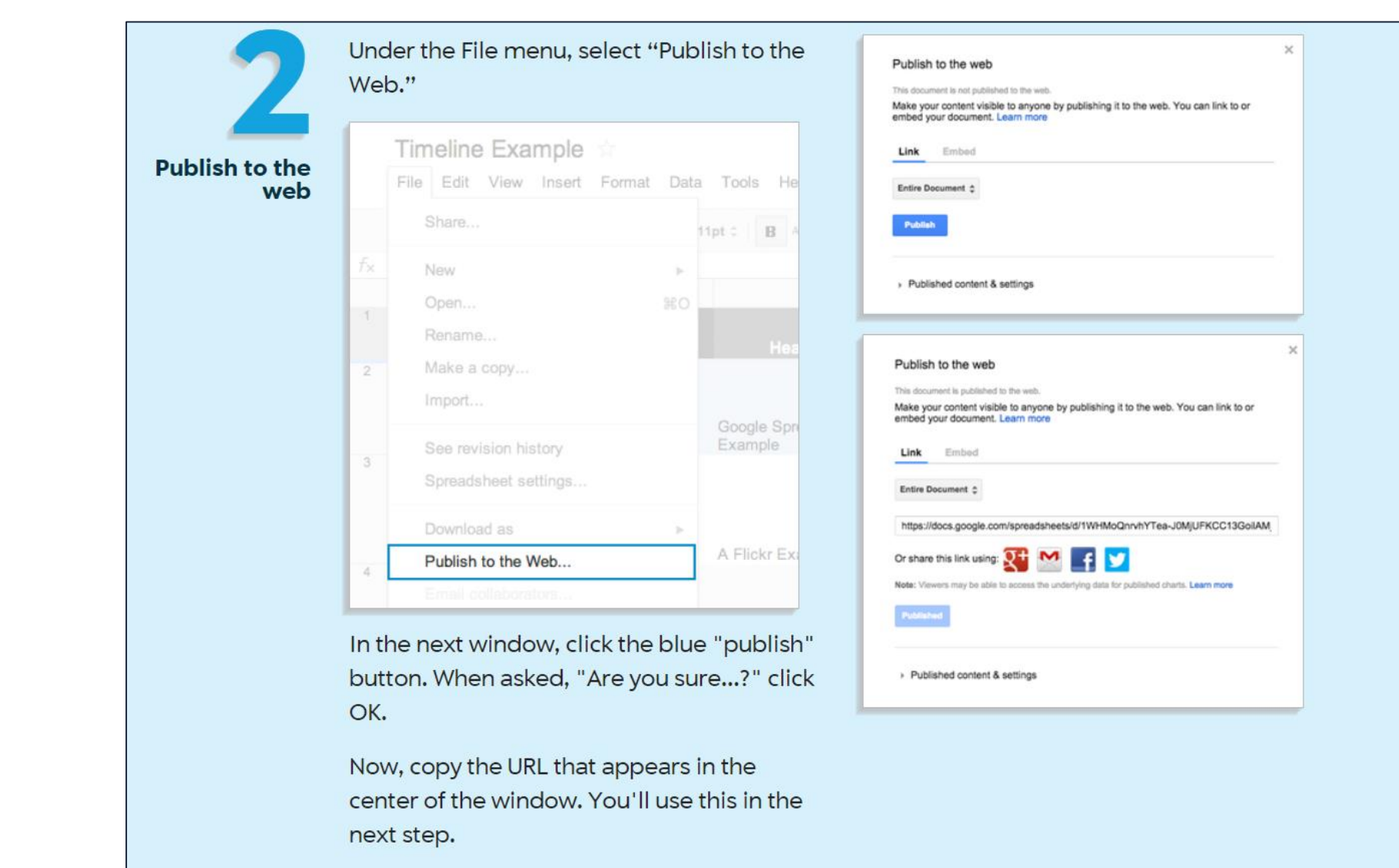
Based on the previous Netaalert website [2] using CSS language we created our own webpage to visualize these Chinese censorship event data (total 4904 events), we used knight lab Timeline JS Library <https://timeline.knightlab.com>.

There are four steps to create the timeline in my CSS webpage.

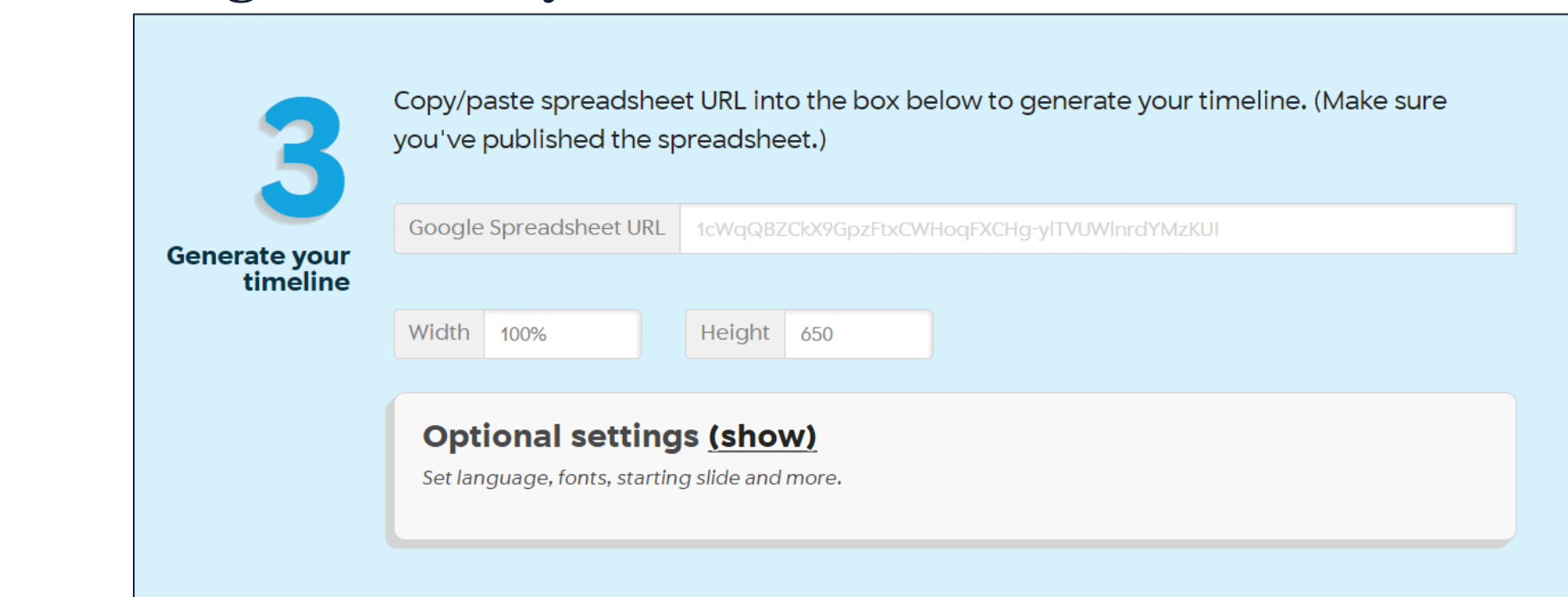
First, create my spreadsheet by using their template.



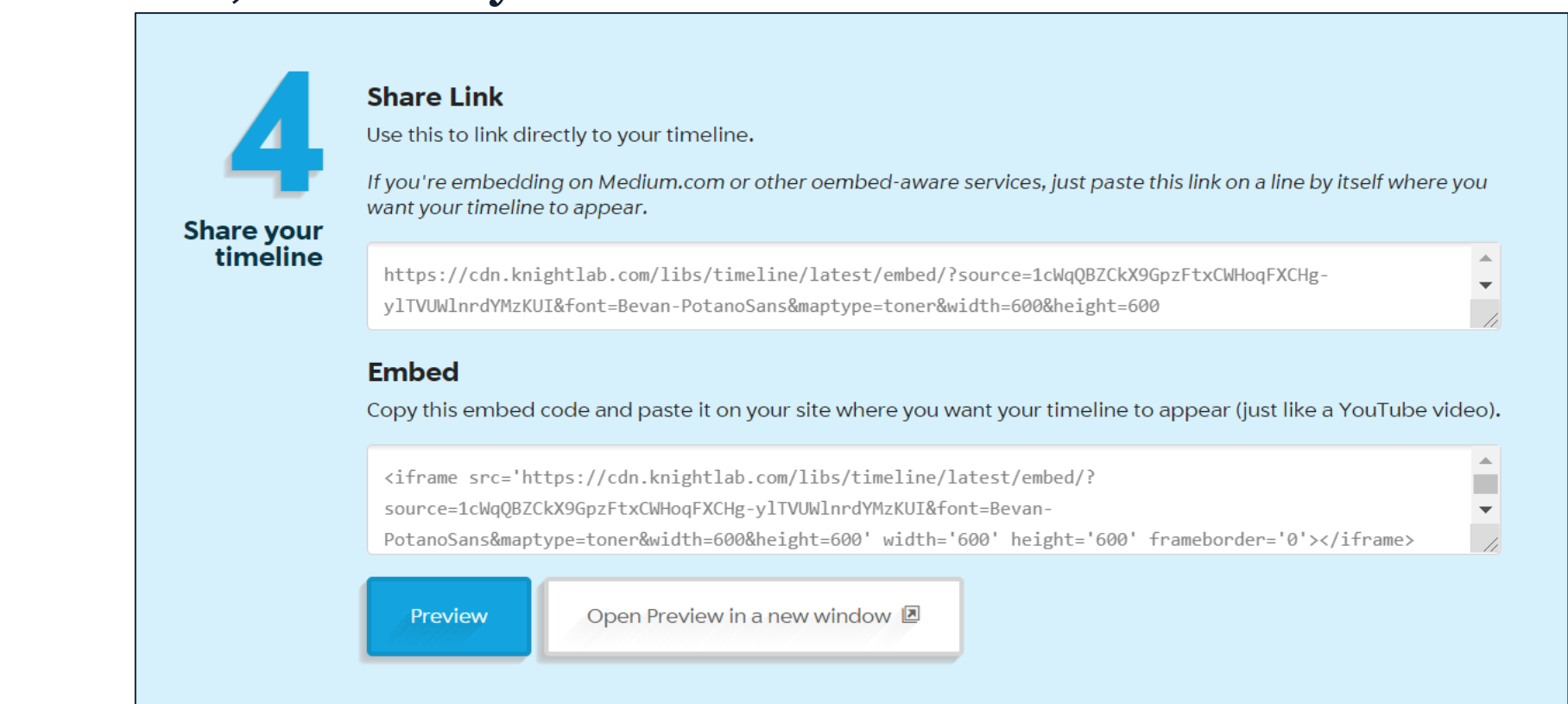
Second, publish to the web.



Third, generate my timeline.



Fourth, share my timeline.



## RESULTS



Figure 1.  
The first page of the group Politician/CPC member



Figure 2.  
Celebrity CHIANG KAI-SHEK event



Figure 3.  
China activist AI WEIWEI

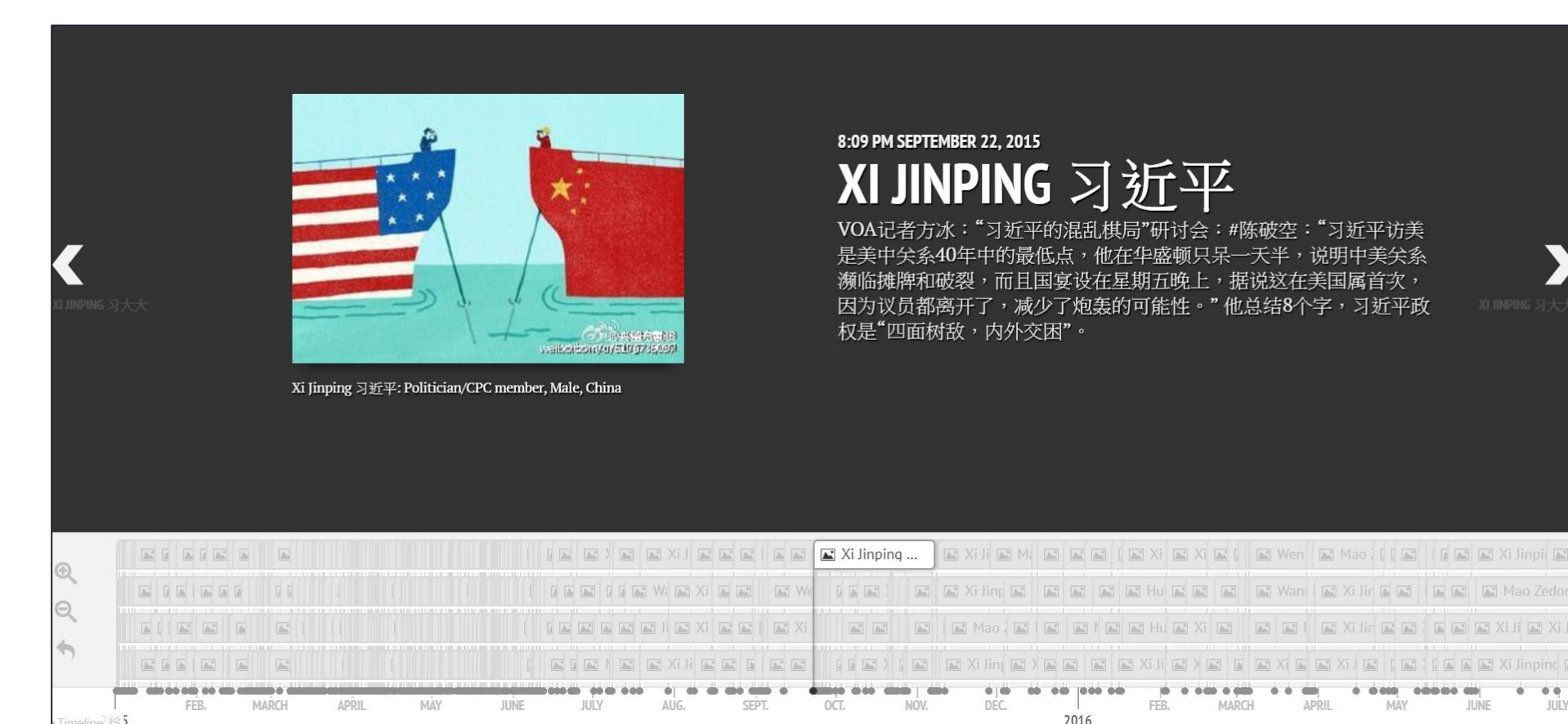


Figure 4. & Figure 5.  
Group Politician/CPC member XI JINPING

### TERM DESCRIPTION

**[1] Sina Weibo:**

A Chinese microblogging (weibo) website. Launched by Sina Corporation it is one of the biggest social media platforms in China.  
<https://www.weibo.com/login.php>

**[2] Netalert website:**

Net Alert is a collaborative project between Open Effect, the Citizen Lab at the Munk School of Global Affairs & Public Policy, University of Toronto, and Professor Jedidiah Crandall at the University of New Mexico. The Netaalert website is developed by the previous student from the University of Toronto.  
<https://netaalert.me/>

### [3] Chines celebrities censorship events data:

The data source is from Sina Weibo and these events were being censored by the Chinese government and being taken down by Sina Weibo. The censorship data included three CSV files which are identities.CSV, weiboscope-censored.CSV and weiboscope-uncensored.CSV. The weiboscope-censored.CSV file is the events that were censored by the Chinese government. We won't use the weiboscope-uncensored.CSV file data in the timeline because the Chinese government doesn't censor these events.

## CONCLUSIONS

In this project, we used the knight lab Timeline JS library in my CSS webpage to display the censorship data that we collected during these three years. We created 21 timelines for each 21 celebrity groups. In every group's timeline, we had to put our data inside the template then publish these links to the web and combine these links into my CSS code. The Citizen lab can combine my CSS webpage code into their Netalet website. There are approximately 5,000 news events in this project. The more data that we put into the timelines, the slower the website operates.

We observed that group 17, Politician/CPC member 政治家/共产党员, has the most data which is approximately over 2,000 events. The next one is group 11, Media 媒, which includes movie stars and singers from China, Hong Kong, and Taiwan. So the politician topics are more sensitive than other topics especially when the article is about people criticizing, mocking or satirizing their government or leader, making jokes of their leader or politician, accusing their leader as a dictator or saying that their leader symbolizes Hitler.

## **ACKNOWLEDGEMENTS**

I would like to thank professor Jedidiah R. Crandall and professor Soraya Abad-M for assisting me with this project, I would also like to thank the CTSP program for funding this project.

