

# CTSP 2019 Summer Project

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

Joseph Huan Chih Wang <sup>#1</sup>, Jedidiah R. Crandall <sup>#2</sup>

Computer Science Department, University of New Mexico  
1155 University Blvd SE, Albuquerque, NM, 87106

<sup>1</sup>[jo520@unm.edu](mailto:jo520@unm.edu) Student

<sup>2</sup>[Crandall@cs.unm.edu](mailto:Crandall@cs.unm.edu) Advisor

### ABSTRACT<sup>#1, #2</sup>

In this project, the purpose is to visualize those Chines Celebrities censorship events data from Sina Weibo website in the Netalert website by using the timeline library. These three years Chines censorship events data that we collected was censored by the Chinese government and was being taken down. There are a total of 319 celebrities and their events and the celebrities are classified into 21 groups base on their job and background. The Chines government wanted the Sina Weibo website to take these events down and with the use of the timeline to visualize these event data, people can read these events easily. If the user is interesting in a specific celebrity or person, the user can click on the celebrity name at the timeline and look through his or her previous events. At the end of this project, we can observe that group 17. Politician/CPC member 政治家/共产党员 has the most data quantity, approximately over 2,000 events. The next one is group 11. Media 媒 which included movie stars and singers from China, Hong Kong, and Taiwan. So we can know that 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, accuses their leader as a dictator or said that their leader symbolizes Hitler.

### I. INTRODUCTION<sup>#1, #2</sup>

In this project, the purpose is to visualize those Chines Celebrities censorship events data from Sina Weibo website [1] in the Netalert website [2] by using the timeline library. These three years Chines censorship events data that we collected was censored by the Chinese government and was being taken down. These censorship data included three CSV file

which are identities.CSV, weiboscope-censored.CSV and weiboscope-uncensored.CSV. The identities CSV file has a total of 319 celebrities, their Chinese and English name, their group (job title), their gender and their nationality. The weiboscope-censored.CSV file is those events that the Chines government took it down from the Sina Weibo website [1]. The data inside weiboscope-censored.CSV file included events original picture, events deleted time, events created year, date and time, events text with simplified Chinese character, retweeted status, reposts count, comments count, text identities, image identities. The weiboscope-uncensored.CSV file has the same content title as weiboscope-censored.CSV file except the events deleted time. We won't use the weiboscope-uncensored.CSV file data in the timeline because the Chinese government doesn't take those events down from the Sina Weibo website.

There are a total of 24 celebrity groups in the Chines censorship events data 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

---

#1: Done by Joseph Huan Chih Wang.

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

However, group 6. Economist 经济学家, 13. Monk 僧 and 19. Real estate tycoon 房 地 产 大 亨 doesn't have any events data that exist in the weiboscope-censored.CSV file so I ignored these three groups. The total categories for our timelines are 21 groups. With the use of the timeline to visualize these event data, people can read these events easily. If the user is interesting at a specific celebrity or person, the user can click on the celebrity name at the timeline and look through his or her previous events.

This paper is organized into seven sections. On section II we present the definition of the important terms that mention in the abstract and the introduction section. On section III we show the methods and outline that we did to visualize the censorship data. On section IV we present the result of our CSS webpage and the result of the timeline to display our data. On section V is the suggestion for the researcher and wrong information in the CSV file that needs to be improved. On section VI is the conclusion of this project and the things that we observed.

*Keywords— Sina Weibo, Netalert website, censorship, visualized.*

## II. TERM DESCRIPTION<sup>#1</sup>

Sina Weibo:

A Chinese microblogging (weibo) website. Launched by Sina Corporation it is one of the biggest social media platforms in China.

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 Netalert website is developed by the previous student from the University of Toronto.

Chines Celebrities censorship events data:  
The source is from Sina Weibo and the event was being taken down by the Chinese government.

## III. METHODS AND OUTLINE<sup>#1</sup>

I will base on the previous Netalert website [2] Figure 1 using CSS language (Cascading Style Sheets is a style sheet language used for describing the presentation of a document written in a markup language like HTML) to create my own webpage with the timeline. To visualize these Chinese censorship event data (total 4904 events), I used knight lab Timeline JS library [3].

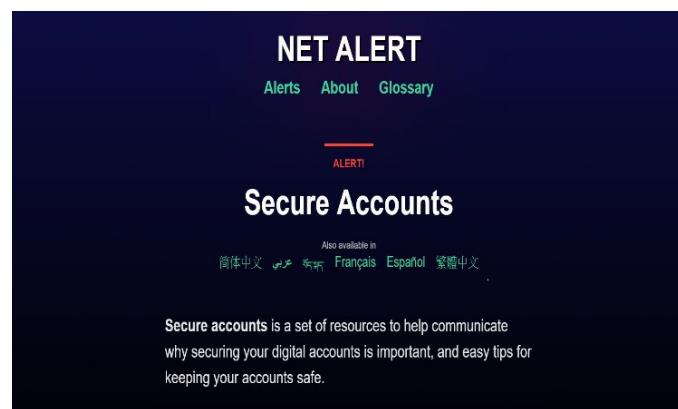


Figure 1.  
Netalert website

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

First, create my spreadsheet by using their [3] template.

I need to copy the template to my own Google Drive account by clicking the "Make a Copy" button. Drop dates, text, and links to media into the appropriate columns. For the contents that I used in the weiboscope-censored.CSV file are events original picture, events created year, date and time, events text and text identities.

Second, publish to the web.

Under the File menu, select "Publish to the Web." In the next window, click the blue "publish" button. When asked, "Are you sure...?" click OK. Then, Close the 'Publish to the web' window. Copy the URL for your Timeline from the browser's address bar. Use this in the next step.

Third, generate your timeline.

Copy/paste spreadsheet URL into the box below at the timeline JS website to generate my timeline. (Make sure you've published the spreadsheet.)

Fourth, share my timeline.

Select the share link part, use this to link directly to my timeline.

After finished the first group timeline (Activist 活动家) I used the same method to create another group's timeline. In the end there are going to be 21 timelines on my webpage. Then the Citizen Lab can combine my webpage code including the timeline link into their Netalert website [2].

#### IV. RESULT

Figure 2 display part of my CSS webpage.

1. Activist 活动家  
[Activist 活动家 Timeline](#)

2. Artist 艺术家  
[Artist 艺术家 Timeline](#)

3. Athlete 运动员  
[Athlete 运动员 Timeline](#)

4. Businessman/Businesswoman 商人  
[Businessman/Businesswoman 商人 Timeline](#)

5. Citizen 公民  
[Citizen 公民 Timeline](#)

7. Fictional character 虚构人物  
[Fictional character 虚构人物 Timeline](#)

Figure 2.  
My CSS webpage

When selecting a specific group to look at such as click at the web link 12. Military general/leader/officer 军事将领/领导/官员 we can see the first-page Figure 3 display the group name and all the celebrities English and Chinese name. Figure 6 is the first page of group 17. Politician/CPC member. Figure 7 and Figure 8 is one of the events of Politician/CPC member.



Figure 3.  
The first page of the Military general/leader/officer timeline

After the first page is the celebrity events which is display in chronological order Figure 4.



Figure 4.  
Celebrity CHIANG KAI-SHEK event

If the user wants to see a specific celebrity event at a specific year and date, he or she can drag the timeline to that year and date then click on the below rectangle which has the celebrity name inside it like Figure 5.



Figure 5.

The first red rectangle is the person's name the second is the event year and date.



Figure 6.

The first page of the group Politician/CPC member



Figure 7.

Group Politician/CPC member XI JINPING



Figure 8.

Group Politician/CPC member XI JINPING

The more the celebrity events data that I put inside the timeline the lower the speed of the web page to display the contents (need to wait for a few seconds).

## V. SUGGESTION AND IMPROVEMENT PART OF THE CSV FILE

This part is for the researcher only who provided the identities.CSV file, weiboscope-censored.CSV file, and weiboscope-uncensored.CSV file. All data contents inside the timeline in this project, I use the same data contents as in the CSV file that my advisor provided to me, even though there is some mistake in the CSV file (wrong information) I didn't modify any of them. I have several suggestions for the researcher and also there are things that the researcher need to modify in the CSV file such as wrong information, duplicate celebrity names, etc. For the suggestions, first, now we already know that the timeline JS library [3] can display tradition and simplify Chinese character also when using this timeline library we have to follow their template to display the picture and the news event, etc (like their demo example). So when using the code to produce the weiboscope-censored.csv file, at the event created time part (the created\_at column) Figure 9, I suggest that for the year, date and time put them in separate column in order to match the timeline template Figure 10, this way it will be easier for the person to copy the whole thing and just past it into the timeline template so the person no need to pick out the year, the date and the time then paste them separately into the template, this is time consuming and painful.

A	B	C	
	deleted	created_at	text
1 original_pic			
2 http://ww1.sinaimg.cn/large/664db5fbgw1erwpm9cizj20it08y758.jpg	42:16.6	2015/5/12 03	这样的情况，鲁迅可以起
3 http://ww4.sinaimg.cn/large/d3e4af42jw1esk4d37d7ej20im0h2403.jpg	10:20.1	2015/5/28 18:04	仅隔五年，这种差别你们
4 http://ww3.sinaimg.cn/large/e23b3959jw1eocxdfxjxj4j20b30diwf9.jpg	30:54.5	2015/1/17 23:18	在当今乃至未来的历史长河
5 http://ww4.sinaimg.cn/large/48c688ecgw1eouhf7dwxj20b407fmxg.jpg	30:21.2	2015/2/2 03:49	【矛盾】教育部长袁贵仁之
6 http://ww3.sinaimg.cn/large/005Q0hdzgw1eoajq0p3qj30b08cq3i.jpg	32:00.5	2015/1/15 21:50	今天（2015年1月15日）：
7 http://ww2.sinaimg.cn/large/61/e6c0l2g1bw1e3jmbj20kp0hmwy.jpg	41:14.4	2015/12/16 14:00	#世界互联网大会#【习近平
8 http://ww4.sinaimg.cn/large/66501ebgw1ewqcw5xmnv20ku1279t.jpg	30:15.1	2015/9/4 13:29	【抗日是挣钱 媚日是姻缘
9 http://ww3.sinaimg.cn/large/da7e3f17jw1esllzfzifc20ap060gm3.jpg	35:10.0	2015/5/30 00:55	{怒}试问李克强总理？沈阳
10 http://ww1.sinaimg.cn/large/85&fad8tw1esk98mpd1j2og0b2wgl.jpg	45:22.1	2015/5/23 22:39	天安门城楼上的毛泽东像

Figure 9.  
weiboscope-censored.csv file

	A	B	C	D	E	F	G	H	I	J
1	Year	Mon	Day	Time	End Y	End M	End D	End T	Display Date	Headline
3513	2015	5	6	18:44					Chen Jiahuo 陈佳罗	
3514	2015	4	6	18:24					Chen Jiahuo 陈佳罗	
3515	2015	4	7	23:39					Chen Jiahuo 陈佳罗	
3516	2015	4	2	22:05					Chen Jiahuo 陈佳罗	
3517	2015	4	10	09:25					Chen Jiahuo 陈佳罗	
3518	2015	4	6	15:52					Chen Jiahuo 陈佳罗	
3519	2015	4	6	14:38					Chen Jiahuo 陈佳罗	
3520	2015	4	7	20:55					Chen Jiahuo 陈佳罗	
3521	2015	4	8	19:57					Chen Jiahuo 陈佳罗	
3522	2015	4	2	19:44					Chen Jiahuo 陈佳罗	
3523	2015	5	29	22:46					Chen Jiahuo 陈佳罗	
3524	2015	4	2	19:13					Chen Jiahuo 陈佳罗	
3525	2015	4	20	17:38					Chen Jiahuo 陈佳罗	
3526	2015	4	2	17:10					Chen Jiahuo 陈佳罗	

Figure 10.  
The timeline template

Second, for the timeline I only used these contents inside the weiboscope-censored.csv file, such as original\_pic, created\_at, text, text\_identities (in order to know this event belongs to which celebrity) for the deleted time Figure 11 (yellow part) I didn't use it because I'm not going to add approximately 5,000 data also the purpose is display the celebrity event that the Chinese government took down so I just used the event created time.

	A	B	C	
1	original_pic	deleted	created_at	text
2	http://ww1.sinaimg.cn/large/664db5fbgw1erwp9m9zj3j20it08y758.jpg	42:16.6	2015/5/8 12:03	这样的情况
3	http://ww4.sinaimg.cn/large/d3e9af42jw1esks4d37d7e20im0h2403.jpg	10:20.1	2015/5/28 18:04	仅隔五年，这
4	http://ww3.sinaimg.cn/large/e23b3959jw1eoexdfx4j20b30liw9q.jpg	30:54.5	2015/11/17 23:18	在当今乃至未
5	http://ww4.sinaimg.cn/large/48c688cgw1fjw7du20b407fmxg.jpg	30:21.2	2015/2/2 03:49	【矛盾】教科
6	http://ww3.sinaimg.cn/large/005QHdgzw1eoajq3j50bi08cq3i.jpg	32:00.5	2015/1/15 21:50	今天（2015年
7	http://ww2.sinaimg.cn/large/f16e6012gw1ezl1g33nbq20kbi0hmyw.jpg	41:14.4	2015/2/16 14:00	#世界互联网
8	http://ww4.sinaimg.cn/large/66501cebgw1ewcqv5xmvmj20ku11279.jpg	30:15.1	2015/9/4 13:29	【抗日是铮铮

Figure 11.

## The deleted time column

Third, for the identities.csv file, I used all of the content such as Chinese name, English name, Group, Nationality. However, I suggest producing another new column that includes both the Chinese name and English name Figure 12 (yellow part) so I can just only copy one column and paste it into the template that way I don't need to copy the Chinese name first then the English name. It will be faster.

	A	B	C	D	E	F
1	Chinese name	English name	Group	Gender	Nationality	English and Chinese
2	白岩松	Bai Yansong	Journalist	Male	China	Bai Yansong 白岩松
3	吴爱英	Wu Aiying	Politician/CPC member	Female	China	Wu Aiying 吴爱英
4	沈冰	Shen Bing	Unknown			Shen Bing 沈冰
5	舒淇	Shu Qi	Media			Shu Qi 舒淇
6	金秀贤	Jin Jiuxian	Unknown	Female	Taiwan	Jin Jiuxian 金秀贤
7	谢霆锋	Nicholas Tsui	Media	Male	Hong Kong	Nicholas Tsui 谢霆锋
8	贺卫方	He Weifang	Professor/Scientist	Male	China	He Weifang 贺卫方
9	谷雨	Gu Yu	Unknown			Gu Yu 谷雨
10	李白	Li Bai	Artist	Male	China	Li Bai 李白

Figure 12.

#### The Chinese and English name column

Fourth, when producing the CSV file follow the timeline template format Figure 13 also combine the first, the second and the third suggestion that I mention above. This will be very helpful for the person to classify the group events in the timeline also it will save a lot of time.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Year	Month	Day	Time	End Year	Month	End Day	End Time	Display Date	Headline	Text	Media	Media Credit	Media Caption	Media Thumbnail	Type	Group	Background
1	2015	3	9	21:46						China Celebrity	范范(林依晨), Lu Yang, 龙洋, 刘洋, 林心如, 林志玲, 林心凌	<a href="#">http://tiny.cc/meyarw</a> <a href="#">http://tiny.cc/meyarw</a> <a href="#">http://tiny.cc/meyarw</a> <a href="#">http://tiny.cc/meyarw</a> <a href="#">http://tiny.cc/meyarw</a> <a href="#">http://tiny.cc/meyarw</a>	Zach Weinersmith	42	范范(林依晨)和龙洋(刘洋)同台，林志玲(林心凌)也来了！“范范”(林依晨)和“龙洋”(刘洋)同台，林志玲(林心凌)也来了！		Photo	http://tiny.cc/meyarw
2	2015	4	28	17:03						范范(林依晨)	【林依晨】公开了！我们终于等到了！范范(林依晨)同台，她好美，她好美，她好美！	Zach Weinersmith	42	范范(林依晨)公开了！我们终于等到了！范范(林依晨)同台，她好美，她好美，她好美！		Photo	http://tiny.cc/meyarw	
3	2015	4	5	23:05						范范(林依晨)	【林依晨】公开了！我们终于等到了！范范(林依晨)同台，她好美，她好美，她好美！	Zach Weinersmith	42	范范(林依晨)公开了！我们终于等到了！范范(林依晨)同台，她好美，她好美，她好美！		Photo	http://tiny.cc/meyarw	
4	2015	4	1	20:13						范范(林依晨)	林依晨的粉丝们，快为自己喜爱的范范(林依晨)投票吧！	Zach Weinersmith	42	林依晨的粉丝们，快为自己喜爱的范范(林依晨)投票吧！		Photo	http://tiny.cc/meyarw	
5	2015	5	30	16:43						范范(林依晨)	粉丝们，范范(林依晨)的微博粉丝数突破200万+，感谢大家！	Zach Weinersmith	42	粉丝们，范范(林依晨)的微博粉丝数突破200万+，感谢大家！		Photo	http://tiny.cc/meyarw	
6	2015	6	2	09:39						范范(林依晨)	粉丝们，范范(林依晨)的微博粉丝数突破200万+，感谢大家！	Zach Weinersmith	42	粉丝们，范范(林依晨)的微博粉丝数突破200万+，感谢大家！		Photo	http://tiny.cc/meyarw	
7	2015	5	8	18:48						范范(林依晨)	感谢你们！感谢你们！感谢你们！	Zach Weinersmith	42	感谢你们！感谢你们！感谢你们！		Photo	http://tiny.cc/meyarw	
8	2015	6	9	15:58						范范(林依晨)	刘姐继续唱！范范(林依晨)王宝强(王宝强)同台，她好美，她好美，她好美！	Zach Weinersmith	42	刘姐继续唱！范范(林依晨)王宝强(王宝强)同台，她好美，她好美，她好美！		Photo	http://tiny.cc/meyarw	
9	2015	5	23	00:13						范范(林依晨)	刚才是谁唱的歌？范范(林依晨)王宝强(王宝强)同台，她好美，她好美，她好美！	Zach Weinersmith	42	刚才是谁唱的歌？范范(林依晨)王宝强(王宝强)同台，她好美，她好美，她好美！		Photo	http://tiny.cc/meyarw	
10	2015	6	10	05:16						范范(林依晨)	共同期待的时刻，快乐的时光！范范(林依晨)王宝强(王宝强)同台，她好美，她好美，她好美！	Zach Weinersmith	42	共同期待的时刻，快乐的时光！范范(林依晨)王宝强(王宝强)同台，她好美，她好美，她好美！		Photo	http://tiny.cc/meyarw	
11	2015	6	28	09:41						范范(林依晨)	【范】（范范）的身材！范范(林依晨)王宝强(王宝强)同台，她好美，她好美，她好美！	Zach Weinersmith	42	【范】（范范）的身材！范范(林依晨)王宝强(王宝强)同台，她好美，她好美，她好美！		Photo	http://tiny.cc/meyarw	
12	2015	4	7	15:47						范范(林依晨)	感谢你们！感谢你们！感谢你们！	Zach Weinersmith	42	感谢你们！感谢你们！感谢你们！		Photo	http://tiny.cc/meyarw	
13	2015	4	7	09:19						范范(林依晨)	今日份的范范(林依晨)来了！范范(林依晨)王宝强(王宝强)同台，她好美，她好美，她好美！	Zach Weinersmith	42	今日份的范范(林依晨)来了！范范(林依晨)王宝强(王宝强)同台，她好美，她好美，她好美！		Photo	http://tiny.cc/meyarw	
14	2015	4	7	13:46						范范(林依晨)	范范(林依晨)和王宝强(王宝强)一起跳舞了！	Zach Weinersmith	42	范范(林依晨)和王宝强(王宝强)一起跳舞了！		Photo	http://tiny.cc/meyarw	
15	2015	5	22	12:46						范范(林依晨)	范范(林依晨)和王宝强(王宝强)一起跳舞了！	Zach Weinersmith	42	范范(林依晨)和王宝强(王宝强)一起跳舞了！		Photo	http://tiny.cc/meyarw	
16	2015	5	22	13:36						范范(林依晨)	【范】（范范）的身材！范范(林依晨)王宝强(王宝强)同台，她好美，她好美，她好美！	Zach Weinersmith	42	【范】（范范）的身材！范范(林依晨)王宝强(王宝强)同台，她好美，她好美，她好美！		Photo	http://tiny.cc/meyarw	

Figure 13.  
The whole timeline template

Several places in the identities.csv file need to be an improvement or modify.

1.

There are total 319 celebrity name in the identities.csv file, however, from the row 235 to 319 the Chinese name is blank Figure 14 also I can't find any events in the weiboscope-censored.csv file for these celebrities but I did find these celebrities in the weiboscope-uncensored.csv file. This might be the reason why in the weiboscope-censored.csv file these people don't have their Chinese name or maybe the researcher doesn't know their Chinese name, so in case the research doesn't know their Chinese name I put in the celebrities simplify Chinese name that I know Figure 15.

	A	B	C	D	E	F	G
1	Chinese name	English name	Group	Gender	Nationality		
240	Fu Yuanhui	Athlete		Female	China		
241	Song Wei	Athlete		Male	China		
242	Andy Lau	Media		Male	Hong Kong		
243	Ma Huateng	Businessman/Businesswoman		Male	China		
244	Zeng Qinghong	Politician/CPC member		Male	China		
245	Rui Chenggang	Media		Male	China		
246	Lu Wei	Politician/CPC member		Male	China		
247	Ren Zhiqiang	Real estate tycoon		Male	China		
248	Puttin	Foreign politician		Male	Russia		
249	Lei Yang	Activist		Male	China		
250	Tsai Ing-wen	Foreign politician		Female	Taiwan		
251	Ling Jihua	Politician/CPC member		Male	China		
252	Jack Ma	Businessman/Businesswoman		Male	China		
253	Zhang Zhaozhong	Military general/leader/officer		Male	China		
254	Park Shin Hye	Media		Female	South Korea		
255	Sima Nan	Media		Male	China		
256	Zhang Weiwei	Professor/Scientist		Male	China		
257	Zhang Yixing	Media		Male	China		
258	Du Fu	Artist		Male	China		
259	Sun Yang	Athlete		Male	China		
260	Zhao Liying	Media		Female	China		

Figure 14.

From row 235 to 319 the Chinese celebrity name is blank

傅园慧	Fu Yuanhui	赵丽颖	Zhao Liying	林心如	Lin Xiumu	安倍晋三	Abe Shinz
刘松伟	Song Wei	万里	Wan Li	林心如	Lin Yuran	张炜	Zhang Wei
刘德华	Andy Lau	宁泽涛	Ning Zetao	林丹	Lin Dan	荷西·穆尼奥	Mourinho
马化腾	Ma Huateng			希拉莉克林顿	Hillary Clinton	Chen Wei	
曾庆红	Zeng Qinghong	马英九	Ma Ying-jeou	张柏芝	Cecilia Cheung	刘涛	Liu Tao
芮成钢	Rui Chenggang	刘强	Jiao Qiang	李冰冰	Li Bingbing	王石	Wang Shi
鲁伟	Lu Wei	伊隆·马斯克	Elon Musk	林志玲	Lin Chi-ling	刘威葳	Liu Weiqi
任志强	Ren Zhiqiang	李健	Jia Jian	许魏洲	Xu Weizhou	王菲	Faye Wong
普京	Putin	姚晨	Yao Chen	潘基文	Pan Ki-moon	罗伯·穆加比	Robert Mugabe
雷洋案	Lei Yang case	郑爽	Zheng Shuang	宋仲基	Song Zhongji	柳传志	Liu Chuanzhi
蔡英文	Tsai Ing-wen	贾跃亭	Jia Yueting			王源	Wang Yuan
今计划	Ling Jihua	吴世勋	Wu Shixun	胡歌	Hu Ge	苏打绿	Soda Green
马云	Jack Ma		Jia Shao	潘永信	Shi Yongxin	马龙	Ma Long
张召忠	Zhang Zhaozhong	孙楠	Sun Nan	林彪	Lin Biao	施晋哲·布拉特	Blatter
孙信惠	Park Shin Hye	贾晓	Jiao Wang	张爱玲	Zhang Ailing	尼基塔·赫鲁晓夫	Khrushchev
司马南	Sima Nan	郭富城	Aaron Kwok	李敏镐	Lee Min Ho	郑智	Zheng Zhi
张维为	Zhang Weiwei	董明珠	Dong Mingzhu	孙悟空	Sun Wukong	金秀贤	Kim Soo-hyun
张艺兴	Zhang Yixing	王嘉尔	Wang Jiale	孟盛华	Meng Shen Aoli	张馨予	Zhang Xinxiu
杜甫	Du Fu	朴槿惠	Park Geun-hye	梁振英	Liang Zhenying	李子軾	Mao Yush
小杨	Sun Yang	洪秀柱	Hóng Xiùzhù	张根硕	Zhang Keun Suk	李纪恒	Li Jiheng

Figure 15.

The celebrity Chinese simplified name that I know

## 2.

Inside both weiboscope-censored.csv and identities.csv file there is an activist name “Lei Yang”, however, some place inside the weiboscope-censored.csv name this activist “Lei Yang case 雷洋案” this is the same people, the “case” in Chinese is “案” it means case such as case 1, case 2 or Watergate scandal case.

## 3.

For the China leader “Xi Jinping 习近平” he has several nicknames in China such as “习大大” that mean master Xi or boss Xi.

## 4.

For the media row 40 Jay Chou 周杰伦 he nationality is Taiwan, not China.

## 5.

Row 24 Figure 16 “范冰冰李晨” this is two people, not one person, “范冰冰” is one person her English name is Fan Bingbing and “李晨” is another person his English name is Li Chen he is Fan Bingbing boyfriend.

	A	B	C	D	E
16	王磊	Wang Lei	Media	Male	China
17	林祖恋	Lin Zulian	Citizen	Male	China
18	李乐斌	Li Lebin	Police officer	Male	China
19	李志	Li Zhi	Philosopher	Male	China
20	释迦牟尼	Sakyamuni	Philosopher	Male	India
21	薄熙来	Bo Xilai	Politician/CPC member	Male	China
22	王小洪	Wang Xiaohong	Police officer	Male	China
23	韩寒	Han Han	Athlete	Male	China
24	范冰冰李晨	Fan Bingbing	Media	Female	China
25	吴小晖	Wu Xiaohui	Businessman/Businesswoman	Male	China
26	宋祖英	Song Zuying	Media	Female	China
27	毛泽东	Mao Zedong	Politician/CPC member	Male	China
28	楼继伟	Lou Jiwei	Politician/CPC member	Male	China
29	沈跃跃	Shen Yueyue	Politician/CPC member	Female	China
30	洪森	Hun Sen	Foreign politician	Male	Cambodia
31	王岐山	Wang Qishan	Politician/CPC member	Male	China
32	王全璋	Wang Quanqi	Unknown		

Figure 16.

范冰冰李晨 is two people

## 6.

Jin Xiuxian 金秀贤 Figure 17 he is a Korean star so the group is Media, Gender is male, Nationality is South Korean.

	A	B	C	D	E
1	Chinese name	English name	Group	Gender	Nationality
2	白岩松	Bai Yansong	Journalist	Male	China
3	吴爱英	Wu Aiying	Politician/CPC member	Female	China
4	沈冰	Shen Bing	Unknown		
5	舒淇	Shu Qi	Media	Female	Taiwan
6	金秀贤	Jin Xiuxian	Unknown		
7	谢霆锋	Nicholas Tse	Media	Male	Hong Kong
8	贺卫方	He Weifang	Professor/Scientist	Male	China
9	谷雨	Gu Yu	Unknown		
10	李白	Li Bai	Artist	Male	China
11	夏霖	Xia Lin	Human rights lawyer	Male	China
12	卡斯特罗	Castro	Foreign politician	Male	Cuba
13	朱镕基	Zhu Rongji	Politician/CPC member	Male	China
14	李晨	Li Chen	Media	Male	China

Figure 17.

Row 6 Jin Xiuxian 金秀贤

7.

For China previous leader “Mao Zedong 毛泽东” he also has nicknames such as “毛” his last name, “毛主席” leader Mao so these are the same people.

8.

For the group Businessman/Businesswoman “Ma Yun 马云” also call Jack Ma so these two name is the same people.

9.

There are Clinton and Hillary Clinton in the identities.csv file, however, Clinton gender is female he should be male Figure 18.

肖钢	Xiao Gang	Businessman/Businesswoman	Male	China
陈有西	Chen Youxi	Human rights lawyer	Male	China
王立军	Wang Lijun	Politician/CPC member	Male	China
蔡依林	Jolin Tsai	Media	Female	Taiwan
克林顿	Clinton	Foreign politician	Female	USA
陈检罗	Chen Zongluo	Unknown		
陈坤	Chen Kun	Media	Male	China
黄子韬	Huang Zikai	Media	Male	China
奥巴马	Obama	Foreign politician	Male	USA
李嘉诚	Li Ka-shing	Businessman/Businesswoman	Male	Hong Kong
王健林	Wang Jianlin	Businessman/Businesswoman	Male	China
陈光诚	Chen Guangcheng	Human rights lawyer	Male	China
汪峰	Wang Feng	Media	Male	China

Figure 18.

Clinton gender

10.

For the Chinese media “Tang Wei 汤唯” in row 33 the researcher make a mistake with row 192 “ Tang Wei 唐僧” Figure 19, I think the English name is not correct too because “唐僧” is a monk and the event is about a story that “唐僧” has three apprentices. They protect him ( 唐僧 ) to the journey of the west to receive a Buddhist scriptures Figure 20. Through their journey there is lots of challenge waiting for them. Such as monsters, evil spirit etc. These evil creatures want to kidnap the monk (their master 唐僧 ) and his apprentices need to save him. Three of his apprentices all have special skills the monkey is the strongest he can do 72 kinds of transform or tricks and the other apprentices is a pig and a Water demon they have skills and personality. For detail, you need to see the story google “Journey to the West 西遊記 ”. So the “唐僧” is a fictional character.

30	洪森	Hun Sen	Foreign politician	Male	Cambodia
31	王岐山	Wang Qishan	Politician/CPC member	Male	China
32	王全璋	Wang Quanqi	Unknown		
33	汤唯	Tang Wei	Media	Female	China
192	唐僧	Tang Wei	Media	Female	China
193	周强	Zhou Qiang	Politician/CPC member	Male	China
194	特朗普	Trump	Foreign politician	Male	USA
195	齐白石	Qi Baishi	Artist	Male	China
196	小王	Xiao Wang	Unknown		

Figure 19.

Row 33 and row 192 are different people



Figure 20.

The Story Journey to the West 西遊記

11.

For the row 83 “Zhao Wei 赵威” and row 158 “Zhao Wei 赵薇 ” Figure 21 they are the same people, she is a China movie star. “赵威” and “赵薇” pronounce the same sound but this one “赵薇” is the correct Chinese name .

79	李嘉诚	Li Ka-shing	Businessman/Businesswoman	Male	Hong Kong
80	王健林	Wang Jianlin	Businessman/Businesswoman	Male	China
81	陈光诚	Chen Guangcheng	Human rights lawyer	Male	China
82	汪峰	Wang Feng	Media	Male	China
83	赵威	Zhao Wei	Media	Female	China
158	赵薇	Zhao Wei	Media	Female	China
159	王林	Wang Lin	Unknown		
160	曾国藩	Zeng Guofan	Military general/leader/officer	Male	China
161	孟建柱	Meng Jianzhu	Politician/CPC member	Male	China
162	安倍晋三	Shinzo Abe	Foreign politician	Male	Japan

Figure 21.

Row 83 and row 158 are the same people

12.

For row 189 “Putin 普京” and row 248 ” Puttin 普京” Figure 22 they are the same people he is Russian present.

187	崔永元	Cui Yongyuan	Media	Male	China
188	陈佳罗	Chen Jialuo	Fictional character		
189	普京	Putin	Foreign politician	Male	Russia
248	普京	Puttin	Foreign politician	Male	Russia
249	Lei Yang case 雷洋案	Lei Yang	Activist	Male	China
250	蔡英文	Tsai Ing-wen	Foreign politician	Female	Taiwan
251	令计划	Ling Jihua	Politician/CPC member	Male	China
252	马云	Jack Ma	Businessman/Businesswoman	Male	China
253	张召忠	Zhang Zhaozhong	Military general/leader/officer	Male	China

Figure 22.

Row 189 and row 248 are the same people

13.

Row 141 “Nikita Khrushchev 赫鲁晓夫” and row 314 “Khrushchev 尼基塔·赫鲁晓夫” Figure 23 they are the same people.

135	崔成浩	Cui Chenghao	Unknown		
136	王思聪	Wang Sicong	Businessman/Businesswoman	Male	China
137	杨幂	Yang Mi	Media	Female	China
138	刘海	Liu Hai	Unknown		
139	李小璐	Li Xiaowei	Professor/Scientist	Male	China
140	柴静	Chai Jing	Journalist	Female	China
141	赫鲁晓夫	Nikita Khrushchev	Foreign politician	Male	Soviet Union
314	尼基塔·赫鲁晓夫	Khrushchev	Foreign politician	Male	Soviet Union
315	郑智	Zheng Zhi	Athlete	Male	China
316	金秀贤	Kim Soo-hyun	Media	Male	South Korea
317	张馨予	Zhang Xinyu	Media	Female	China
318	茅于轼	Mao Yushi	Economist	Male	China
319	李纪恒	Li Jiheng	Politician/CPC member	Male	China

Figure 23.

Row 141 and row 314 are the same people

14.

Row 203 “Liu Dehua 刘德华” and row 242 “Andy Lau 刘德华” Figure 24 they are the same people, this one “Liu Dehua 刘德华” is directly translated from Chinese. Another one “Andy Lau 刘德华” is his first name and last name.

198	陈建刚	Chen Jiangang	Unknown		
199	江青	Jiang Qing	Politician/CPC member	Female	China
200	王增志	Wang Zengzhi	Unknown		
201	乔石	Qiao Shi	Politician/CPC member	Male	China
202	陈云	Chen Yun	Politician/CPC member	Male	China
203	刘德华	Liu Dehua	Media	Male	Hong Kong
242	刘德华	Andy Lau	Media	Male	Hong Kong
243	马化腾	Ma Huateng	Businessman/Businesswoman	Male	China
244	曾庆红	Zeng Qinghong	Politician/CPC member	Male	China
245	芮成钢	Rui Chenggang	Media	Male	China
246	鲁炜	Lu Wei	Politician/CPC member	Male	China
247	任志强	Ren Zhiqiang	Real estate tycoon	Male	China
248	普京	Puttin	Foreign politician	Male	Russia

Figure 24.

Row 203 and row 242 are the same people

15.

Duplicate data Figure 25 at Row 98 “Wang Yizhi 王羲之” and row 238 ” Wang Yizhi 王羲之”.

93	李小琳	Li Xiaolin	Businessman/Businesswoman	Female	China
94	杨崇勇	Yang Chongyang	Politician/CPC member	Male	China
95	李春城	Li Chuncheng	Politician/CPC member	Male	China
96	史迪奇	Stitch	Fictional character		
97	吴淦	Wu Hao	Professor/Scientist	Female	China
98	王羲之	Wang Yizhi	Writer/Essayist/Novelist	Male	China
238	王羲之	Wang Xizhi	Writer/Essayist/Novelist	Male	China
239	克拉伦斯惠	Clarence Hui	Media	Male	Hong Kong
240	傅园慧	Fu Yuanhui	Athlete	Female	China
241	刘松伟	Song Wei	Athlete	Male	China

Figure 25.

Duplicate data

16.

I’m not quite sure about this one but I think they are the same people. Row 185 “Liu Qiangdong 刘强东” and row 265 “Liu Qiang 刘强” Figure 26.

182	邓小平	Deng Xiaoping	Politician/CPC member	Male	China
183	邓超	Deng Chao	Media	Male	China
184	鲁迅	Lu Xun	Writer/Essayist/Novelist	Male	China
185	刘强东	Liu Qiangdong	Businessman/Businesswoman	Male	China
265	刘强	Liu Qiang	Unknown		
266	伊隆·马斯克	Musk	Businessman/Businesswoman	Male	USA
267	李健	Li Jian	Media	Male	China
268	姚晨	Yao Chen	Media	Female	China
269	郑爽	Zheng Shuang	Media	Female	China
270	贾跃亭	Jia Yueting	Businessman/Businesswoman	Male	China

Figure 26.

## VII. CONCLUSION<sup>#1</sup>

In this project, I used knight lab Timeline JS library [3] in my CSS webpage to display the censorship data that we collect during these three years. I created 21 timelines for each 21 celebrity groups. To follow the template format in the timeline library I have to pick out the data contents that we want from the identities.CSV file and the weiboscope-censored.CSV file then I have to paste them into the timeline template. After finishing all the 21 groups timeline I have to put the timeline publish links into my CSS code to display in my webpage. Then the Citizen lab can combine my CSS webpage into their Netalert website. There is a total of 319 celebrity, 21 groups and approximately 5,000 news events in this project. The more the data (censorship news events) that we put inside the timeline the lower the speed that the webpage to display.

After finished all the 4904 celebrities events (319 celebrity members), I found out that the group 17. Politician/CPC member 政治家/共产党员 has the most data quantity, approximately over 2,000 events. The next one is the group 11. Media 媒 which Included movie stars and singers from China, Hong Kong, and Taiwan. So from here, I can observe that 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, accuses their leader as a dictator or said that their leader symbolizes Hitler.

#### REFERENCES

- [1] <https://www.weibo.com/login.php>
- [2] <https://netalert.me/>
- [3] <https://timeline.knightlab.com/>