

Weekly homework #2

Data Journalism

Tomás Dodds Rojas

Bente de Bruijn (12191566), Haizea Azpiazu (12288942) Alexander Zeilstra (12306681),

Lisa Betlem (12040002), Jessica Blom (12358681)

1027 words

Topic/issue

It has been a hot topic of discussion for months now, and increasingly more countries are encountering the consequences of the outbreak. Of course, we are talking about the now world-famous coronavirus. The virus broke out in a Chinese region, called Wuhan, in December 2019. Its cause, however, is still unclear. Initially, the fish market in Wuhan was seen as the place of origin. Today however, we are not so sure. According to the World Health Organization, symptoms of the virus include “respiratory symptoms, fever, cough, shortness of breath and breathing difficulties” (World Health Organization, n.d.).

At this moment, the worldwide count of infections is at 89.073, spread over 51 different countries, with 3.048 casualties. Also, in the Netherlands, there are now 18 people infected by the virus. Because of the fact that it spreads so quickly on such a large scale, the media have been covering it nonstop for the past few months. It has even been deemed ‘infodemic’ by the World Health Organisation. Due to the overabundance of information it seems to get harder for people to determine what is real and what is not about the virus.

In April of 2009 a similar global outbreak occurred, namely the flu pandemic H1N1 or swine flu. Starting from an infection in Mexico, the illness got widespread globally through human contact like it would with a seasonal flu. About 11 to 21% of the global population at the occurrence of the outbreak got infected, however the fatality of the illness was around 0.01 to 0.08%. Even though this rate was fairly low, it was still seen as an epidemic as it spread globally, and the World Health Organization (WHO) therefore announced it as a “public health emergency of international concern”. The outbreak was then declared a pandemic as the counting of the cases got difficult to maintain. The symptoms of a normal case of H1N1 usually lasted around four to six days. This low fatality rate resulted in critics claiming that the WHO had exaggerated the danger of the pandemic and were part of the cause for the spreading of fear for the pandemic. This was post-pandemic seen as unnecessary as the risk for children and adults for severe complications was slim. Also, the death rate was supposedly lower than originally expected. It was further thought that the WHO made

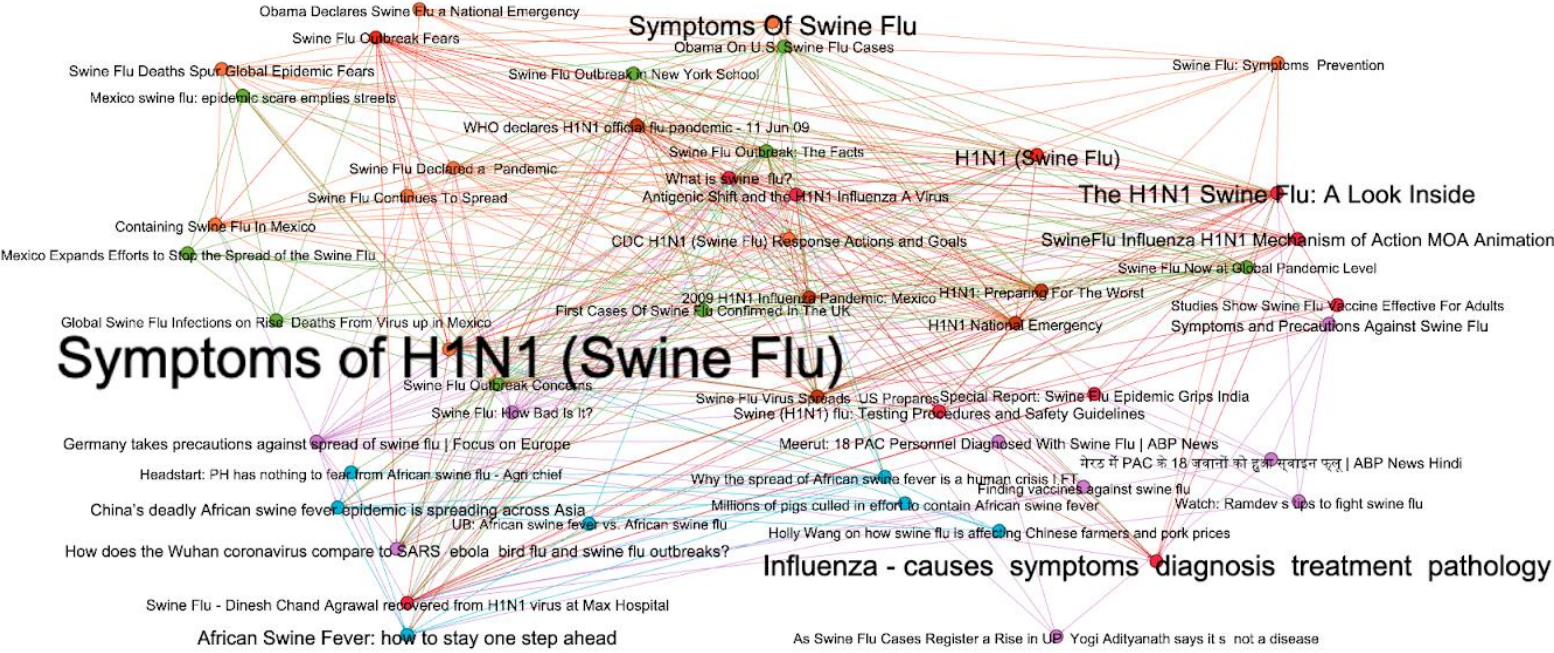
statements to the public seem more alarming and dooming than the situation actually was, with the supposed goal to sell more vaccines as the virus turned out not to be as dangerous as thought at the beginning of the pandemic.

Justification + Approach

Previous outbreaks such as the swine flu have caused global panic like the recent coronavirus. However, the fear surrounding the coronavirus seems to be intensified by social media. Because of the overabundance of information about the coronavirus it would potentially allow for more misinformation and disinformation to be spread, which could add to the global panic. The way information is circulating has changed at a quick rate within the 21st century. This also means that the way false information is spread has changed because of the rise of the internet and social media. “The spreading of misinformation on online social media has become a widespread phenomenon to an extent that the World Economic Forum listed massive digital misinformation as one of the main threats for the modern society” (Bessi et al, 2016). So, one of the main threats to society, in the case of the coronavirus an infodemic, added with a pandemic seems like a recipe for disaster. To see if there is more of a spread of misinformation surrounding the coronavirus compared to a previous viral outbreak, in our case the swine flu, we take a look at the largest video-sharing social networking site: YouTube. Does the coronavirus really stand out from previous outbreaks in relation to misinformation? How much has the spread of misinformation possibly changed in the last 11 years since the swine flu outbreak?

We have extracted data from YouTube by using the YouTube Data Tools software. We will be specifically taking a look at the video networks surrounding both the coronavirus and the swine flu. We want to compare the coronavirus to the swine flu because they seem to be similar global outbreaks and can both be classified as pandemics. We used both viruses as search queries and put a time frame on both. Meaning that we chose to only look at the videos that were made on the viruses on YouTube within the first two months of the outbreaks. We chose this time frame because of how recent the coronavirus outbreak is, thus to do the whole duration of both viruses would be impossible. The video network will be based on the relevancy of the videos to the search query. Then we put both data sets in Gephi, an open-source visualization software that creates graphs and networks, to create two graphs we can compare. We will be looking at the titles of the videos in the networks to determine which could be misinformation or disinformation. To do this we should look at certain variables

within the title to decide whether it contains false information. We established that titles that contain false information fall under misinformation or disinformation. Also, what can be considered ‘clickbait’ falls under those same categories. “‘Clickbait’ refers to ‘content whose main purpose is to attract attention and encourage visitors to click on a link to a particular web page’ [‘clickbait,’ n.d.] and has been implicated in the rapid spread of rumor and misinformation online.” (Chen et al, 2015). Clickbait can contain for example false promises to which, in this case, the videos or the truth does not live up to.



Gephi network of video-results on YouTube under the search query “Swine Flu” in the time period between 01-04-2009 and 31-06-2009 (crawl depth = 0) (color division by modularity) (label size based on number of views)

