HOW COULD CENTRALIZED DATA SHARING MITIGATE THE EFFECTS OF COVID-19?

A Research Paper submitted to the Department of Computer Science
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

Michael Chang

By

Sung Joon Park

April 17, 2021

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISOR
Paul McBurney, Department of Computer Science

A Simple COVID-19 Dashboard for Virginia

Michael Chang
Computer Science
University of Virginia
Charlottesville Virginia USA
myc6cp@virginia.edu

Sung Joon Park
Computer Science
University of Virginia
Charlottesville Virginia USA
sp3bk@virginia.edu

ABSTRACT

Since the advent of the pandemic, online dashboards have become the main means for the public to better understand the COVID-19 pandemic. However, after using the COVID-19 dashboard built by the Virginia Department of Health (VDH) and several others, it has become apparent that most COVID-19 dashboards are not built for the everyday user. In fact, it seems as if COVID-19 dashboards are more so built for those with authority such as health and government organizations. One of the problems with existing dashboards is that they seek to display as much data as possible, which may be too complex for most people. These modern dashboards are also at much larger scale, showing data on the state, national, and the global scale with few that attempt to show data on the county level, which may not provide much actionable information for the public.

The main objective of this technical Capstone project will be to develop a user-friendly COVID-19 dashboard for the state of Virginia that delivers data on the state and county levels. The goal of the COVID-19 dashboard is to create a simple interface with localized data that provides vital and actionable information for users in Virginia. To do this, the majority of the project will consist of research and implementation. Research will involve looking at online resources for different kinds of data and frameworks to create the dashboard. Implementation will involve using APIs to obtain data and using frameworks in JavaScript to both engineer and display the visualizations.

Through this pandemic, it has become apparent that data and technology have been vital to suppressing the spread of the virus. Although COVID-19 dashboards play a major role in providing a summary of the pandemic to key decision-makers, they should also be used to inform the general public. Through this simple COVID-19 dashboard, ordinary citizens of Virginia will be able to access data on the pandemic that is more applicable to them in an easily digestible way. Through this dashboard, citizens of Virginia will be empowered to make smarter decisions.

INTRODUCTION

Throughout the COVID-19 pandemic, technology has been shown to play a crucial role in preventing the spread of the virus. In fact, countries that have effectively implemented digital technology into mitigation processes early on have done the best with flattening their "incidence curves" [6]. A digital health solution, known as dashboards, was first introduced to help extract and visualize useful information from a large amount of data. Dashboards were meant to display datasets in ways that were easily understandable, usually with "time-series charts and geographic maps, ranging from region-level statistics to case-level coordinate data [1]. As the pandemic progressed, various dashboards were developed by governments, universities, and other organizations to help aid decision makers. One of the most popular dashboards has been the John Hopkins dashboard, as it was developed early on in the pandemic and now has a large array of features available to use. As stated by Dong, Du, and Gardner (2020), some of the creators of the John Hopkins dashboard, their dashboard has been "crucial to help inform modelling efforts and control measures" and with this dashboard, they have been able to "provide researchers, public health authorities, and the general public with a user-friendly tool to track the outbreak as it unfolds" [2].

However, dashboards have not been as user-friendly as anticipated. As Pietz, McCoy, and Wilck state, many dashboards lack consistency, key metrics, and can be hard to understand [3]. Additionally, as dashboards have increased in complexity, they have become less suited for the general public. As Wernimont (2021) states, "covid-19 dashboards are remarkably ill-suited to the complexities of our pandemic reality" because "historically the dashboard has been designed for a reader or viewer who has some power to act" [5]. Current dashboards have been influenced by dashboards of the past, which were made for figures of authority. This has now led to a lack of COVID-19 dashboards that are targeted towards the ordinary citizen. While decision makers need as much information as possible in order to make the best decision possible, ordinary citizens would be overwhelmed by this amount of information. Another problem that has been found has been that many dashboards do not display data on the county level in the United States. Wernimont (2021) states that "some local and county level data are available, but not always in the user-friendly format of the dashboard, and certainly not consistently across the country" [5].

Specifically for Virginia, the official dashboard made for the state has been the COVID-19 dashboard made by the Virginia Department of Health [VDH] (https://www.vdh.virginia.gov/coronavirus/covid-19-in-virginia/). However, in an analysis on government dashboards, with the VDH dashboard being one of them, Pietz et al. (2020) state that these dashboards can be too confusing for ordinary users, as the dashboards would overwhelm the user with information [3]. Given that this is the official dashboard for Virginia, it makes sense that

the dashboard would try to display and visualize every piece of data that is available in Virginia. However, this can be confusing for most users, who would most likely be coming to this dashboard to find a few COVID-19 statistics for their locality. Although the VDH dashboard does have county data, the data is not in a place that is easily accessible, as it takes navigating through multiple pages to find it.

The objective of this technical project is to create a user friendly COVID-19 dashboard for Virginia that will solve the issues of information overload, usability, and the lack of county data. The main goal of this dashboard is to make it easy to understand, while still providing the general public of Virginia with key metrics on COVID-19. To do this, we focused heavily on the aspect of usability. Specifically, we considered the core components of usability which consist of ease of learning, efficiency of use, memorability, and subjective satisfaction [4]. These aspects are what influenced our design choices for the frontend of the web application. To solve the issue of information overload, we selected COVID-19 statistics that we found to be the most common across many dashboards. To solve the problem of a lack of county data, we also dedicated a large portion of the dashboard to county data for Virginia, showing the information in an easily understandable tabular format while also giving the user the option of selecting one specific county to look at. Our target audience for this dashboard is any citizen interested in COVID-19 in Virginia.

RELATED WORK

Currently, there are many dashboards available that provide data on the COVID-19 pandemic. As mentioned earlier, the John Hopkins dashboard has been a really powerful tool used by many during this pandemic. As stated by Dong, Du, and Gardner (2020), this dashboard is an international dashboard, reporting cases "at the province level in China", "at the city level in the US, Australia, and Canada", and "at the country level otherwise" [2]. This dashboard gets its data from DXY, which is "an online platform run by members of the Chinese medical community" that "aggregates local media and governments reports to provide cumulative totals of COVID-19 cases in near real time at the province level in China and at the country level otherwise" [2]. For city-level cases in the USA, Australia, and Canada, the dashboard relies on reports from the US CDC, the government of Canada, the Australian Government Department of Health, and various state or territory health authorities [2]. Additionally, before the dashboard is manually updated, the case numbers are confirmed with both regional and local health departments and also city-level and state-level health authorities [2]. When looking at the actual website, some of the main features are an interactive map showing where COVID-19 is most prevalent in the world, tables containing various statistics, and a graph showing daily cases over a period of time. Although this dashboard is very powerful and reliable, it is not suited for ordinary citizens looking to find information about their local county or state. Since it is an international dashboard, the scope of the data is very wide, but this limits the depth and granularity of data that can be achieved. For this reason, a user would have a hard time navigating to their state or city to find data that is relevant to them, as this dashboard is not meant for this functionality. Additionally, due to the complex features on this website and the large amount of data it deals with, it is also slow to load and cumbersome to navigate for older computers. When considering the core components of usability, the Johns Hopkins dashboard fails in the four aspects of usability (ease of learning, efficiency of use, memorability, and subjective satisfaction) for a user looking to find COVID-19 data in their locality, as it is too difficult for an ordinary citizen to use to find their county and state data.

Additionally, the Virginia Department of Health (VDH) has also created a dashboard for the state of Virginia (https://www.vdh.virginia.gov/coronavirus/covid-19-in-virginia/). According to their website (https://www.vdh.virginia.gov/coronavirus/about-the-data/), the VDH reports "detailed, near real-time health data" related to the pandemic for Virginia and the data is sourced from VDH surveillance data systems. The VDH has multiple dashboards, which consist of a "COVID-19 Daily Dashboard", a "Virginia's Key COVID-19 Measures Dashboard", a "Pandemic Metrics Dashboard", and two others that offer other insights. The dashboard made for the general public of Virginia is the "COVID-19 Daily Dashboard". On the "Summary" tab, it shows some statistics about cases, hospitalizations, deaths, outbreaks, and testing. However, this dashboard is not very intuitive and can be quite confusing to use. For example, when trying to find the number of new cases for Albemarle county, one would go to the "Cases" tab, only to find that this only shows an interactive map graphic of Virginia. Most users do not know where their county is geographically located in the midst of Virginia, so this is not very convenient. Additionally, data was not available in a tabular format under this tab. The tabular format is actually located under the "Locality" tab, but this table does not have text that is searchable through the browser since the entire dashboard is simply a Tableau application inserted into the webpage, making it difficult to find the specific county a user would be looking for. Overall, the VDH dashboard feels clunky and hard to navigate. When considering the core components of usability, the VDH dashboard fails in ease of learning, efficiency of use, memorability, and subjective satisfaction. This is because it is not intuitive, requires users to navigate through many pages to find their desired information, and does not make use of useful browser features such as simple text searching because of the fact that the dashboard is a Tableau application and not a native web application.

Lastly, Wissel et al. (2020), realizing the importance of tracking outbreaks at the city level, created their own interactive dashboard called "COVID-19 Watcher", which showcases a few metrics on the municipal level, such as testing capacity, cases, and deaths [7]. This dashboard utilizes a methodology "to aggregate county-level COVID-19 data into metropolitan areas" ("Introduction, para. 3). The data is also passed through a quality control check which "ensures that updated data files are the anticipated size and format" ("Code", para. 1). However, the problem with this dashboard is that it only displays one graphic at a time. It requires the user to select

the county they are in and what data they want to see on the graphic. Additionally, the user can only select between two statistics, "Cases" or "Deaths". The graphic itself is also not interactive and the time period on the graph is limited to two months with no option to extend this. The lack of choices and requirements for user input on this dashboard make it really cumbersome to use, as a user is very limited in what they can see and do on this dashboard. The dashboard was minimally designed but because of this, many aspects of usability were sacrificed. When considering the core components of usability, the COVID-19 Watcher dashboard does very well in ease of learning and memorability, as its minimal design makes it very simple to use. However, efficiency of use and subjective satisfaction were sacrificed because of the minimal design. For efficiency of use, this dashboard failed because having a user select a new set of options to see the other statistic is very inefficient. For subjective satisfaction, this dashboard failed because the user is limited to two statistics, they can only view one at a time, and the time period shown on the graphic is too short and has no option to modify it.

SYSTEM DESIGN

1 Software Stack

For the frontend of the application, we decided to use React JS, mostly because our team had prior experience with working with this framework. For the backend of our application, we decided to use Express JS. Since our backend would not be overly complex and would simply be responsible for getting data from the API's we have chosen, Express JS was a great choice as it was simple, minimalistic, and fast.

1.1 Frontend

The client side was designed so that it was minimalistic and simple to view and use. Our website can be divided into two parts: the first half deals with the data on the state of Virginia overall and the second half deals with data related to the counties in Virginia. Both parts of our dashboard contain a mixture of text and visualizations. Implementation wise, the frontend would use the backend endpoints to query for certain data. The backend then receives those requests from the frontend, fetches the relevant data from the API, and returns it back to the frontend. The frontend would then retrieve, process, and display the data. For visualizations, the team researched different data visualizations libraries for JavaScript and chose to use Recharts JS for its simple user interface and ease of use.

For the part of our website that dealt with data related to the state of Virginia as a whole, we had to consider what statistics and visualizations we should show that would be the most beneficial to the average user in Virginia. To do this, we used our research and experience with using other dashboards. By comparing various dashboards and seeing which COVID-19 statistics were most common throughout them, we made our decision to display certain statistics and visualizations over others. For statistics, we use plaintext to show data related to cases, deaths, hospital capacity, and vaccines. For visualizations, we used line charts for total cases and deaths, daily new cases and deaths, and the death rate. We also used pie charts to visualize hospital and ICU bed capacity.

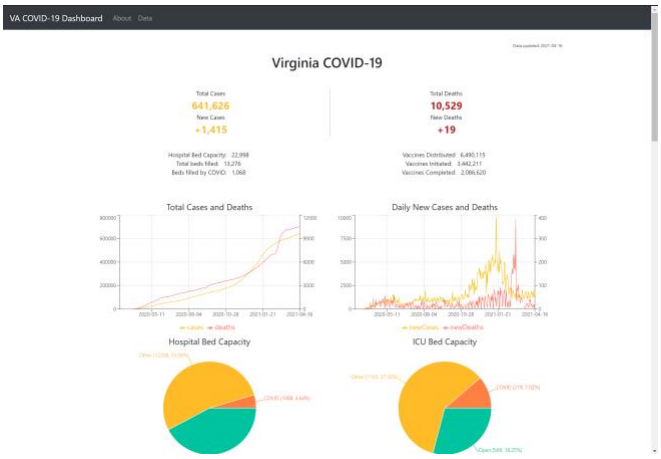

Figure 1: Overall State Data View on the Virginia COVID-19 Dashboard.

For the portion of our website dedicated to county level data, we first had a table that displayed all of the current information for every county in Virginia. One problem that we found with the VDH dashboard was that their table that displayed county data was not text searchable in the browser, but since our table is part of our native web application, this issue is fixed. For county level data, we show the data in order of decreasing significance, with the most important data starting from the left. From left to right, we display the statistics: cases, deaths, cases increase, death increase, county population, case density, and then infection rate. Directly below the county data table, we have a search bar that allows a user to select a specific county to view. The user has the option of either selecting a county from the drop down or typing in the county they want. This will then display a view similar to the one we showed for the overall state data, with a mix of text and data visualizations. One challenge we had was figuring out how to best display the data for specific counties along with the visualizations. Initially, we thought about creating a separate webpage that would show data for specific counties, but after finding a npm library for a search bar we were able to keep the county data on the same webpage as the state data. The search bar gave us the ability to show the respective county's visualizations based on the user's selected county.

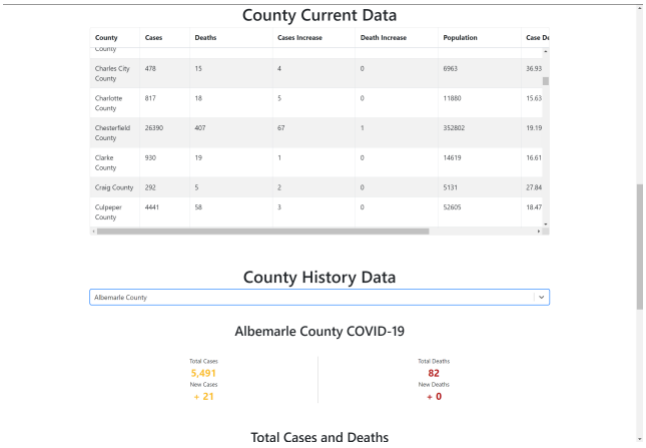

Figure 2: Table Containing All County Data and County Search Bar.

Our design choices were heavily influenced by the core aspects of usability mentioned earlier, which are ease of learning, efficiency of use, memorability, and subjective satisfaction. For example, the reason all of the data on the dashboard is on a single webpage is to minimize user navigation to improve usability. Since this dashboard is meant for the ordinary citizen of Virginia, we wanted to make sure that every aspect of this dashboard was simple to use and that there was minimal user interaction. For this reason, the goal of our frontend was to make it such that all a user would have to do to see data that would be related to them would simply be to visit the dashboard.

1.2 Backend

As mentioned earlier, Express JS was used for its simplicity, minimalism, and quickness. Express was simply used to create the endpoints for both the state and county data that the frontend could retrieve from. It would then fetch the data from the API that we query and return it as JSON back to the frontend.

1.3 Data Source

All of our data was provided by COVID Act Now (covidactnow.org). This API provides free data with rich features for both state and county. The data is from official sources such as the U.S. Department of Health and Human Service, the Centers for Disease Control and Prevention, The New York Times, and official state and county dashboards. The data fields are well documented and the data is updated daily at around noon Eastern Time. Before we knew about this API, we went directly to the New York Times GitHub repository for our county data, as they have compiled CSVs for county data, which are updated daily. However, after finding out about the COVID Act Now API, we directly used the data from there, as their county data was based off the New York Times repository already. We were also previously using data from The COVID Tracking Project (https://covidtracking.com/) but mid-March, the website indicated that they would stop adding in new data, which led us to search for a new API.

PROCEDURE

Given that our target audience is the general public of Virginia, it should be relatively easy for them to use our dashboard. We really wanted to make the dashboard easy to use for any citizen of Virginia, even those that are less familiar with technology. As such, any user of the system would simply have to visit the dashboard to see data relevant to Virginia. Since all of the data is on one page, there is nothing that a user has to do unless they want to see more detailed information about their specific county. That is the only part that requires user interaction. For this, the user has to select the specific county they want to see data for. The user would do this by clicking the drop down menu and then they would either select the county from the predefined list or type it into the search bar.

RESULTS

Our project solves a number of issues that people would normally have with some of the more popular dashboards. The first is simplicity and quick design. As mentioned earlier, one COVID dashboard we looked at was from John Hopkins University (https://coronavirus.jhu.edu/map.html) where we noticed that the data visualization seemed cluttered and the webpage took too long to load. Cluttered data makes it hard for users to understand and properly analyze the data to take appropriate actions. Our dashboard is also simple in that it displays state and county data all in one page while other dashboards require clicking to different pages for different levels of data, which means more latency for users. We believe that providing clean and organized data is paramount to providing a quick and easy to use dashboard.

Furthermore, because our project has a much smaller scope, it could be more relevant to users in Virginia, as it provides data from the county to state level. Rather than simply providing graphs on the number of cases and deaths, we provide more actionable data for users such as hospital bed usage and capacity, vaccination numbers, population, case density and infection rate. The majority of the dashboards we commonly see provide global scale data, which may not provide much actionable information for ordinary citizens. However, if the data becomes more localized and personalized with more information than just cases and deaths, users can better understand the pandemic situation in their local area and be more likely to take proper action.

CONCLUSIONS

We designed a system to meet the need of simple, easy, and actionable COVID-19 data for users. One issue we noticed was that the majority of dashboards were too cluttered or loaded with too much latency for a typical user. To solve this, our project aims to provide minimal and simple data to users with low latency. Users are able to load data quickly and see the data they need to see. Visualizations are clean and well organized to ease the use of the dashboard. Another issue was providing more personalized data. Many dashboards like to provide firsthand global scale data, but this makes it hard for users to understand the pandemic situation in their own localities, making it harder for them to take actionable steps. However, our dashboard provides rich information from

county level to the state, more than the typical cases and deaths, which could help users to take appropriate action from where they are living. All in all, we believed that providing a good and minimalistic user interface with the proper functionality was paramount to creating a dashboard that users could find useful and enjoyable to use. We believe that our project could serve as a good example of a dashboard that is targeted towards citizens of a specific state.

FUTURE WORK

One area of future work could be on further user interface research, where our team could survey users on their thoughts of our dashboard, asking for any feedback on what they like or do not like about it. Then, appropriate action would be taken by our team to fix those issues. In the future, we could also make the dashboard customizable by user preferences. The users would be able to see which visualizations they would like to see and would have the option of saving their layout for future use.

Another area of improvement could be inclusion of more states and counties. Expanding our dashboard out to more states and their respective counties would make our dashboard usable by anyone in the United States. However, our team wants to be careful to not provide too much information or make the usage of the dashboard more difficult, as our highest goal is to have a dashboard that is simple to use for an ordinary citizen. Potentially, a feature could be added where the dashboard detects user location and automatically sets the state and county to the user location. We could also give the user the option to change the state they want to look out as well. But overall, our dashboard will always show localized information before showing any global information, as we would want to preserve the narrow scope of our dashboard.

The last area of research could be enabling more mobile-friendly features. Whether this means adjusting the dashboard to mobile sizes or creating a mobile application, granting mobile access could allow easier access for a greater body of users. We believe that it is important that platform differences should not inhibit the accessibility of our dashboard.

REFERENCES

[1] Budd, J., Miller, B., Manning, E., Lampos, V., Zhuang, M., Edelstein, M., … McKendry, R. (2020). Digital technologies in the public-health response to COVID-19. *Nature Medicine, 26*, 1183-1192. doi:10.1038/s41591-020-1011-4

[2] Dong, E., Du, Hongru., Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time, *The Lancet Infectious Diseases, 20,* 533-534. doi: 10.1016/S1473-3099(20)30120-1

[3] Pietz, J., McCoy, S., & Wilck, J. H. (2020). Chasing John Snow: data analytics in the COVID-19 era, *European Journal of Information Systems*. doi:10.1080/0960085X.2020.1793698.

[4] U.S. Department of Health and Human Services. (2006). *The research-based web design & usability guidelines, enlarged/expanded edition.* Washington, DC: Author.

[5] Wernimont, J. (2021). Covid-19 dashboards are vital, yet flawed, sources of public information. *The Washington Post*. Retrieved from https://www.washingtonpost.com/

[6] Whitelaw, S., Mamas, M. A., Topol, E., & Van Spall., H. G. C. (2020). Applications of digital technology in COVID-19 pandemic planning and response, *The Lancet Digital Health, 2,* 435-440. doi:10.1016/S2589-7500(20)30142-4

[7] Wissel, B. D., Van Camp, P. J., Kouril, M., Weis, C., Glauser, T. A., White, P. S., … Dexheimer, J. W. (2020). An interactive online dashboard for tracking COVID-19 in U.S. counties, cities, and states in real time. *Journal of the American Medical Informatics Association, 27,* 1121-1125. doi:10.1093/jamia/ocaa071