Real-time Hybrid Dashboard and App for Mpox Outbreak Surveillance

Daniel Quezada
Computer Science Department
California State University,
Fullerton, USA
Daquezada19@csu.fulleron.edu

Sampson Akwafuo
Computer Science Department
California State University,
Fullerton, USA
0000-0001-8255-4127

Aniket Wattamwar
Computer Science Department
California State University,
Fullerton, USA
Aniket.wattamwar@csu.fulleron.edu

Abstract— **With the continuous rise in the global threat of infectious diseases, surveillance systems have become powerful tools for monitoring the development and transmission patterns of fast-changing disease outbreaks, by public health officers. These surveillance systems typically track relevant epidemiological data such as number of cases, fatality rates, hot spots, and in most cases, evaluate the impact of public health intervention strategies in these locations. Data from these surveillance apps can also be used for efficient emergency response preparatory logistics, as well as early containment. In this paper, we presents a modular, open-source, GIS-enabled, web-based dashboard and mobile app that are capable of visualizing and identifying epidemiological patterns for any infectious disease of interest, in real time. These visualizations can be in the form of GIS maps, charts, and other metrics to track certain indicators. Python and Streamlit library are used to manage the frontend of the application. Data from the 2022 Mpox outbreak in the United States were used to evaluate the application. Non-governmental organizations or other community-based groups can leverage on, and adapt this dashboard to monitor the spread of any quick-onset disease outbreak in their regions.**

Keywords— Disease Surveillance, Dashboard, Mobile Apps, Outbreak Tracking, Mpox,

I. INTRODUCTION

Active surveillance of disease outbreaks is imperative for efficient pre-disaster planning, ensuring prompt and targeted intervention following a disaster or epidemic outbreak and impact reduction. It can be aptly defined as a continuous, systematic collection, analysis, and interpretation of health-related data needed for the planning, implementation, and evaluation of public health practice. Generally, there are about three types of data collection approaches to public health surveillance. These include passive surveillance, active surveillance, and syndromic surveillance. Passive surveillance is the most common type of surveillance. It includes reports from laboratories, physicians, or other case notifications that are regularly sent to the local or state health department [1]. Case reports that are based on a standardized case definition for a particular disease are included in this category. A second form of surveillance is the active surveillance. This occurs when the collection of data from the lab, physician, or other healthcare provider is initiated by the health department. Active surveillance is often used during investigations of sudden disease outbreaks or for research studies. Active surveillance has an advantage over passive surveillance because it achieves more wholistic, complete and accurate reporting. However, it is usually more resource intensive (personnel, finance and time) for the public health agencies than other types of surveillance. Syndromic surveillance describes ongoing, systematic collection, analysis, interpretation, and application of real-time indicators for disease tracking. This allows for early detection of possible disease outbreaks before public health authorities would otherwise identify them. This surveillance approach collects samples and disease-related data such as the speed and size of an outbreak; when does it start? how fast does it grow? And how long does it last?

Each of these surveillance type utilizes the four basic components of surveillance techniques, including collection, analysis, dissemination, and response. Outbreak data collection and analysis can be conducted at the local, state, federal, or international level by public agencies as well as by private organizations. Dissemination and response are specific public health activities. Disease surveillance therefore includes analysis, interpretation, and feedback of outcome-specific data. Hence, surveillance may monitor cases of disease reported by clinicians or identified in laboratories, changes in practice, or other related indicators of public health importance.

A. Epidemiological Data Stream

Surveillance of infectious diseases is therefore very critical, mostly in epidemiology and public health in general. in order to monitor the progression of new outbreaks. Modern telecommunications allow for governments and other public health organizations to disseminate epidemiological data to one another in a near instantaneous manner. This can be attributed to the rise of information technology across the world as low and middle income countries (LMICs) begin to adopt similar disease surveillance systems as high-income countries (HICs). These data streams have low enough latency that institutions have been publishing epidemiological data in near real-time to better help researchers and officials understand how an outbreak progresses. However, when sharing this data through the internet, there is no international standard across these public health institutions on how to format their epidemiological data [1]. Many public health institutions will have differing interfaces for how their data is uploaded: examples include comma-separate values (CSV) files, extensible markup language (XML), portable document format (PDF), JavaScript Object Notation (JSON), or through an application programming interface (API). This is a well known problem in epidemiological studies, as there is no set standard on how researchers might extract data from these public health institutions to use in their research activities.. Recent studies Althouse et all have established a framework for how researchers may be able to use existing or novel epidemiological data streams to gather data in a systematic manner [2], [3].

II. Background and Review

A. Monitoring Disease Outbreaks

Continuous monitoring of disease outbreaks or fluctuations in disease outbreaks' contributory factors remains an importance tool in public health management and policy development. A core approach to monitoring is surveillance. As described in introductory section, surveillance. The objective is to associate disease activity to geographic location. It serves as an early warning system for impending public health emergencies [4], as well as documenting the impact of an intervention. It further tracks progress towards specified goals, such as mitigation or containment. It also monitors and clarifies the epidemiology of health problems, informing public health policy and strategies [4].

B. Human Mpox (Monkeypox) Fever

Human monkeypox (Mpox) is a zoonotic viral infection caused by the monkeypox virus (MPXV). The virus is a double-stranded DNA specie of the *Orthopoxvirus* genus in the *Poxviridae* family, which includes smallpox, cowpox, vaccinia and other viruses. It appears in two distinct genetic clades: Clade I (previously known as the Congo Basin or Central African) clade and Clade II also known as the West African clade [5]. The monkeypox virus was discovered in 1958 in Denmark from sone research monkeys. However, the first reported human case of Mpox was in 70, from a nine-month-old boy in the Democratic Republic of the Congo. Similar to outbreaks of other zoonotic viral infections, recurrent outbreaks of Mpox have occurred in most countries of West and Central Africa [6]. Mpox can spread from person to person or from animals to people. Following eradication of smallpox in 1980 and the end of global smallpox vaccination programme, mpox steadily emerged in central, east and west Africa. Various small mammals such as squirrels and monkeys are suspected to be the natural reservoir of the virus. Detection of viral DNA by polymerase chain reaction (PCR) remains the preferred laboratory method of diagnoses. Antigen and antibody detection methods may not be useful as they do not distinguish between orthopoxviruses [7].

On July 23, 2022, the World Health Organization (WHO) declared the ongoing monkeypox outbreak to be a public health emergency of international concern. This represents the first time in almost twenty years that other regions outside African would experience an Mpox outbreak. Those infected with Mpox may develop rashes and flu-like symptoms (headache, chills, muscle soreness) within three weeks of exposure. The incubation period can range from five to twenty-one days, although typically only last six to thirteen days, and an infected person may transmit the virus up to twenty-eight days after the incubation period [8]. In past outbreaks, Farahat et estimates that case fatality rate of Mpox has been as high as 11%, although it is currently estimated at 3-6%. Currently, there are no FDA approved therapeutics or vaccines for Mpox specifically, but antiviral treatments and vaccines for smallpox have been approved to treat Mpox [9].

C. *Digital Syndromic Dashboards as Disease Surveillance Tools*

Digital dashboards have proven to be a dependable tool in the study and analysis of epidemiological data for tracking disease outbreak progression. They summarize and present information through a variety of graphs, plots, and indicators, thereby connecting public health data to protect the health of communities. Typically, these surveillance system dashboards are tailored to a specific infectious disease [10]. This was seen throughout the COVID-19 pandemic where the governments in HICs had built their own web-based dashboards to monitor the progression of the virus in their respective countries [11]. Although these systems proved to be an effective tool for policymakers to develop intervention strategies against the prevalence of SARS-CoV-2, the same could not be said for LMICs where development of such dashboards are usually hindered by lack of resources [12]. A recent study found that while most countries are working towards developing a surveillance system, there is a concern that web-based dashboard surveillance systems that rely on paper-based records containing the epidemiological data would be too slow to update. In addition to this, research has been done into expanding disease surveillance systems in rural parts of LMICs, specifically in areas where humans have an increased risk of contracting zoonotic diseases [13], [14]. Without proper surveillance in these areas, outbreaks are more likely to occur and will spread rapidly to neighboring communities, provinces, or nations. Digital dashboards have been created as disease surveillance tools in rural communities in China and Somalia, and data from these surveillance tools can provide a much-needed inputs in the development of models for disaster and re-emergent outbreak pre-planning and management [15]–[17]. We present an efficient and cost-effective approach to dashboard development. It involves the use of open source and modular techniques.

III. Methodology

The comprehensive surveillance systems was created using Streamlit, an open-sourced web app framework capable of handling both the frontend and backend of an application [18]. Streamlit allows for dashboard developers to create and deploy web apps using Python scripts. Streamlit has since proven to be a useful tool for researchers as it has been used in similar studies for epidemiology and geospatial applications [19], [20]. In addition to Streamlit, three other Python libraries were used to create the dashboard application: pandas, json, and plotly. The input and data acquisition processes are as shown in Fig. 1.

A. *Extract, Transform, Load Process*

In order to power the dashboard app, the epidemiological data concerning the target disease was pulled from the CDC's weekly updates of Mpox cases across the United States. Here,

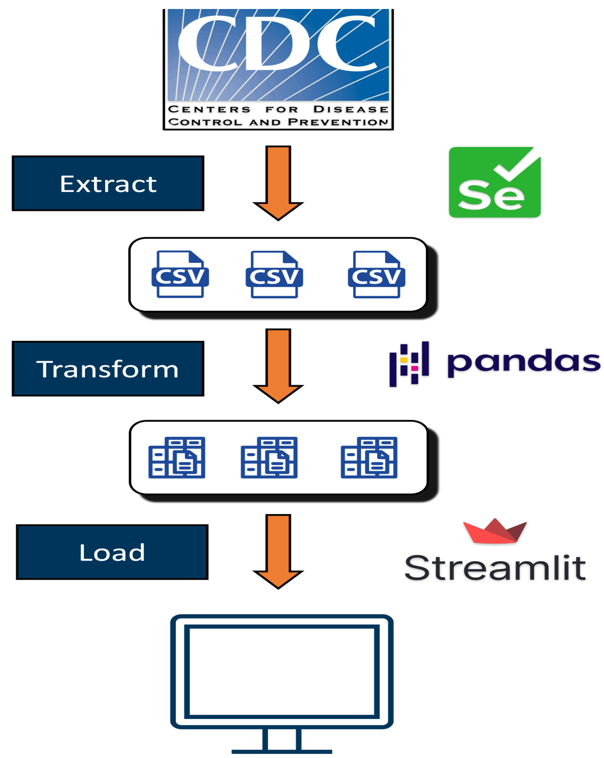

Fig. 1: Flow chart of the system and techniques used in the development

we begin the extract, transform, load (ETL) process described in [21], where a web scraping bot was created using the Selenium library so as to automatically download the csv files when published by the CDC. This will launch a Firefox browser window, load predetermined CDC web-pages hosting Mpox data, search for HTML buttons that would download the csv data, and finally save the csv files to the project folder. Moving onto the second phase of the ETL process, the csv data goes through a 'sanitation' phase where a Python script using data-analysis library, Pandas, will drop

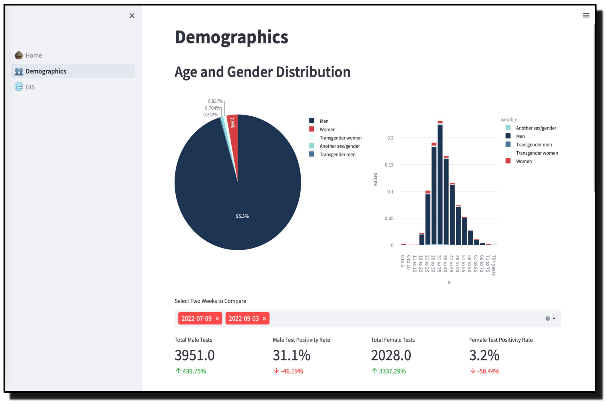

Fig. 2: Menu items and demographics of initial outbreak cases

any unnecessary tables, headers, or rows in addition to reformatting calendar dates to the ISO 8601 standard. Once the csv data is validated and reformatted into Pandas DataFrame objects, the project can finally load these DataFrames into other Python scripts across the dashboard. Plotly is an open-source Python library that offers simple to complex data analytics and visualization. Streamlit offers support for Plotly figures, so most of the epidemiological visualizations in the dashboard were created using Plotly.

B. *Top Level Menu Items*

There are three major navigational items on the first level menu: Home, Demographics, and GIS, as shown in Fig. 2. When the user clicks Home, various data filtering, parameter selection and visualization options are presented. The user can select to query number of vaccines administered, number of cases of demographics of affected individuals. Data acquisition form (Fig. 3), to be used by field officers and data entry personnel can also be obtained from home menu. The demographics and GIS sections are described in the next section. The pie chart in Fig. 2 is a sample output of the demographics section. Each slice represents a different gender with their size being proportional to the number of cases contributed to the total Mpox case count. In relation to gender, a stacked bar chart comparing the proportion of cases with the age range over the different genders was generated as well. In the interest of comparing which ethnicities were affected the most by Mpox, a multi-line chart highlighting the weekly percentage difference was created. Two Plotly line charts and a bar chart represented the total cumulative cases, daily cases reported, and seven-day average over time, respectively. In addition to these Plotly figures, interactive modules were created using built-in Streamlit methods. An overview module displays the confirmed cases as well as Mpox lab tests administered for three time periods: *total, last month, last week*. A comparison module built using the Streamlit multi-select widget allows user to select two different weeks to compare the rate of change between total tests for male and female gender, as well as the positivity rates.

C. *Real-time reports from Filed Offciers and Agents*

In addition to data and related inputs from established databases, field officers and local monitoring and evaluation (M&E) officers can enter data directly to the dashboard. The mobile app is customized to work with low internet

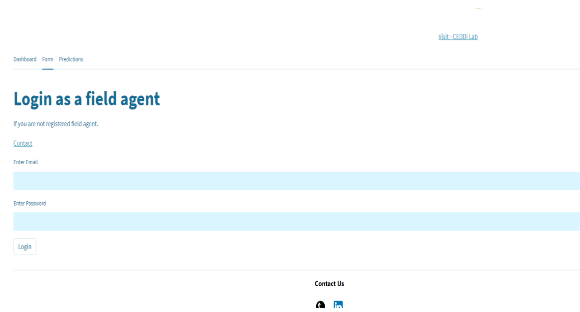

Fig. 3: Data and Report Acquisition Form

bandwidth. The form tab shown in Fig. 3 allows these agents to report about emerging outbreaks, new or updated about ongoing outbreaks. The google sheet is linked to the Power BI dashboard, ensuring real-time updating of the charts and plots on the dashboard. The link database is updated as well.

IV. Results and Discussions

Given the global ongoing outbreak of human monkeypox (mpox), this study aims to build a friendly web-based surveillance system dashboard that will monitor the progression of mpox cases in the United States. The system can be easily re-configured to track data for any other emerging disease outbreak in a region of interest. Using the epidemiological data published by the CDC in conjunction with the visualization tools that the dashboard offered, the following chats were generated:

- The cumulative cases of the outbreak over a selected period of time. [see Fig. 4]
- The daily confirmed cases over time. This includes all age groups and gender categories. [see Fig. 5]
- Tracking reveals sudden and sustained increase in the number of cases, necessitating the public health emergency declaration [see Fig. 6]

- The top three ethnicities made up 39%, 33%, and 22% of reported cases. [see Fig. 7]

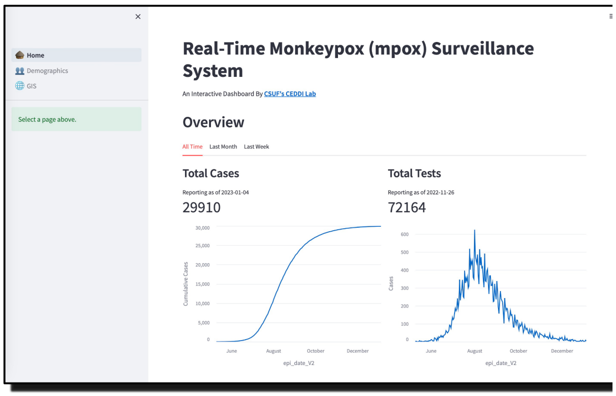

Fig. 4: Sample Output: Number of confirmed cases of the outbreak over a selected period of time

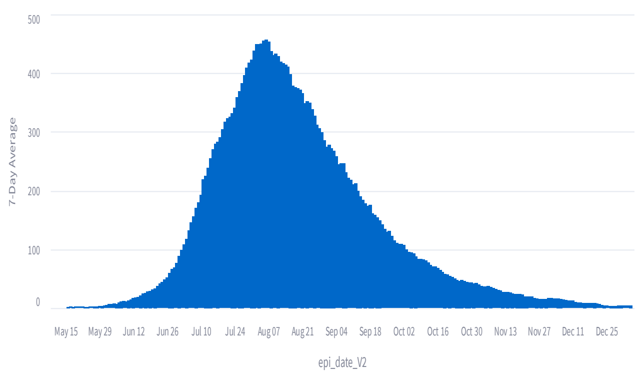

Table 2 highlights the four largest age groups from every gender group and their respective joint proportions:

A. Geographic Information System Capabilities

A geographic information system (GIS) was considered essential for the creation of this dashboard surveillance system. To avoid introducing more libraries, tile-map choropleth maps were created using Plotly figures. GeoJson data containing the outline of every US state was downloaded from the US Census Bureau's website [20]. Two different GIS maps were created that highlight the cases recorded and the vaccines administered in each US states. Building a web-based dashboard allows for the dissemination of public health data to a wide range of key stakeholders, including public health officials, healthcare workers, researchers, and the public. By using free and open-source software, this dashboard project provides LMICs or NGOs an example of how a web-based disease surveillance system can be implemented using limited resources. Using this dashboard

Table 1: Proportions of Selected Age Groups from Reported Gender Categories

Joint Proportions	Age Groups			
	21-25	***26-30***	***31-35***	***36-40***
Men	9%	18%	22%	16%
Women	0.54%	0.55%	0.47%	0.35%
Trans Women	0.12%	0.15%	0.24%	0.10%
Trans Men	6.12e-4%	5.44e-4%	3.74e-4%	3.06e-4%
Another	0.10%	0.15%	0.17%	0.11%

system to analyze the current 2022 Mpox outbreak in the United States, we can deduce the following:

- Men were overwhelmingly affected by Mpox, especially young to middle adults
- Around 65% of all confirmed Mpox cases in the United States came from men ranging in age from 20-40 years old
- The total number of confirmed Mpox cases in the US was 29,910 (as of January 4th 2023)
- The largest number of daily confirmed cases was recorded as 624 cases on July 31st 2022

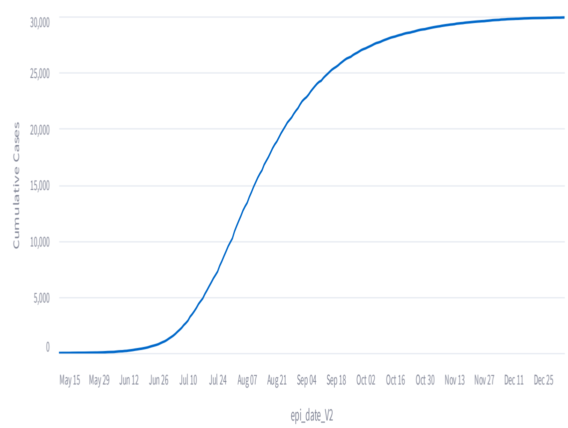

Fig. 6: sudden and sustained increase in the number of cases, necessitated the public health emergency declaration

- The largest seven-day average was recorded as 457 cases on August 5th 2022
- The ethnicities most impacted by the Mpox outbreak were: Black/African Americans, White, and Latino/Hispanics

Viewing this retrospectively, once major public health

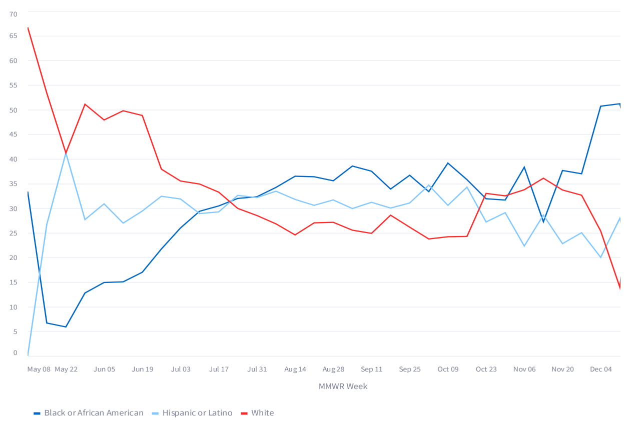

Fig.7: The top three most affected ethnic groups

Fig. 8: Use of GIS features to study hospots for cases or vaccines acceptance ration

interventions were undertaken by US public health officials, there was a decline in cases. Four days before the United States hits it daily case peak, the FDA approves JYNNEOS, a smallpox vaccine, for Mpox usage [21]. Four days later from daily case peak, US Department of Health and Human Services (HHS) secretary, declares Mpox a public health emergency in the United States [22]. After the public health emergency declaration by the HHS secretary, the development of medical counter measures, such as vaccines and antiviral treatments, were fast tracked and distributed to various states. This is reflected in the data, as cases began to drop after mpox was declared a public health emergency. As of December 2nd, 2022, HHS does not expect to renew the public health emergency declaration for mpox when it expires on January 31st, 2023 [23].

V. Conclusion

Web-based dashboards that provide real-time visualizations of epidemiological data have been proven to identify trends, hotspots, and other patterns that can inform response and prevention efforts. In this project, we have demonstrated that a mobile friendly dashboard surveillance system can be implemented using free and open-source software. This dashboard allows for easy sharing and the dissemination of data with the key stakeholders in public health, further increasing awareness and understanding of disease outbreaks.

A. Future Work

We intend to extend the features of the of the surveillance system by adding a section on machine learning predictive models. The machine learning component is still under development. Using data generated by these dashboards, users can select and view predictive models about the study disease, for the location of interest.

References

[1] C. Rojanaworarit, "Principles of public health surveillance: a revisit to fundamental concepts," *J. Public Heal. Dev.*, vol. 13, no. 1, pp. 69–86, 2015.

[2] B. M. Althouse *et al.*, "Enhancing disease surveillance with novel data streams: challenges and opportunities," *EPJ Data Sci.*, vol. 4, no. 1, pp. 1–8, 2015, doi: 10.1140/epjds/s13688-015-0054-0.

[3] G. Fairchild *et al.*, "Epidemiological data challenges: Planning for a more robust future through data standards," *Front. Public Heal.*, vol. 6, no. NOV, pp. 1–9, 2018, doi: 10.3389/fpubh.2018.00336.

[4] S. L. Groseclose and D. L. Buckeridge, "Public Health Surveillance Systems: Recent Advances in Their Use and Evaluation," *Annu. Rev. Public Health*, vol. 38, pp. 57–79, 2017, doi: 10.1146/annurev-publhealth-031816-044348.

[5] Y. Huang, L. Mu, and W. Wang, “Monkeypox: epidemiology, pathogenesis, treatment and prevention,” *Signal Transduction and Targeted Therapy*, vol. 7, no. 1. Springer Nature, Dec. 01, 2022, doi: 10.1038/s41392-022-01215-4.

[6] S. Akwafuo, X. Guo, and A. Mikler, “Epidemiological modelling of vaccination and reduced funeral rites interventions on the Reproduction number , R0 of Ebola virus disease in West Africa .,” *J. Infect. Dis. Med. Microbiol.*, vol. 2, no. 3, pp. 7–11, 2018.

[7] E. M. Bunge, B. Hoet, L. Chen, and F. Lienert, “The changing epidemiology of human monkeypox—A potential threat? A systematic review,” *PLoS Negl. Trop. Dis.*, vol. 16, no. 2, Feb. 2022, doi: 10.1371/journal.pntd.0010141.

[8] R. A. Farahat *et al.*, “Human monkeypox disease (MPX),” *Infez. Med.*, vol. 30, no. 3, pp. 372–391, 2022, doi: 10.53854/liim-3003-6.

[9] B. K. Titanji, B. Tegomoh, S. Nematollahi, M. Konomos, and P. A. Kulkarni, “Monkeypox: A Contemporary Review for Healthcare Professionals,” *Open Forum Infect. Dis.*, vol. 9, no. 7, pp. 1–13, 2022, doi: 10.1093/ofid/ofac310.

[10] H. Burkom, W. Loschen, R. Wojcik, and R. Holtry, “Electronic surveillance system for the early notification of community-based epidemics (ESSENCE): Overview, components, and public health applications,” *JMIR Public Heal. Surveill.*, vol. 7, no. 6, pp. 1–13, 2021, doi: 10.2196/26303.

[11] CDC, “COVID Data Tracker: Home. #datatracker-home,” *Center for Disease Control and Prevention (CDC)*, 2021. https://covid.cdc.gov/covid-data-tracker (accessed Dec. 04, 2022).

[12] A. Vahedi, H. Moghaddasi, and F. Asadi, “Applications, features and key indicators for the development of Covid-19 dashboards: A systematic review study,” *Informatics Med. Unlocked*, vol. 33, no. January, 2020.

[13] K. E. L. Worsley-Tonks *et al.*, “Strengthening global health security by improving disease surveillance in remote rural areas of low-income and middle-income countries,” *Lancet Glob. Heal.*, vol. 10, no. 4, pp. e579–e584, 2022, doi: 10.1016/S2214-109X(22)00031-6.

[14] K. Jayatilleke, “Challenges in Implementing Surveillance Tools of High-Income Countries (HICs) in Low Middle Income Countries (LMICs),” *Curr. Treat. Options Infect. Dis.*, vol. 12, no. 3, pp. 191–201, 2020, doi: 10.1007/s40506-020-00229-2.

[15] W. Yan, L. Palm, X. Lu, and S. Nie, “ISS-An Electronic Syndromic Surveillance System for Infectious Disease in Rural China,” *PLoS One*, vol. 8, no. 4, 2013, doi: 10.1371/journal.pone.0062749.

[16] R. Kamadjeu and C. Gathenji, “Designing and implementing an electronic dashboard for disease outbreaks response-Case study of the 2013-2014 Somalia Polio outbreak response dashboard,” *Pan Afr. Med. J.*, vol. 27, no. Supp 3, pp. 1–5, 2017, doi: 10.11604/PAMJ.SUPP.2017.27.3.11062.

[17] S. E. Akwafuo, A. R. Mikler, and C. Ihinegbu, “Data-driven depot pre-positioning model and location-routing algorithm for management of disasters and disease outbreaks,” *Proc. ACM Symp. Appl. Comput.*, pp. 681–684, 2022, doi: 10.1145/3477314.3507174.

[18] S. Documentation, “Streamlit: A faster way to build and share data apps,” 2013. https://streamlit.io/.

[19] A. Saxena and S. . Robila, “Analysis of the New York City’s Vehicle Crash Open Data,” in *IEEE International Conference on Big Data (Big Data),* 2021, pp. 6017–6019, doi: doi: 10.1109/BigData52589.2021.9672012.

[20] S. Srinivasan, R. Rujula Singh, R. R. Biradar, and S. A. Revathi, “COVID-19 monitoring system using social distancing and face mask detection on surveillance video datasets,” *Int. Conf. Emerg. Smart Comput. Informatics, ESCI 2021*, pp. 449–455, 2021, doi: 10.1109/ESCI50559.2021.9396783.

[21] K. M. Gokhale, J. S. Chandan, and K. Toulis, “Data extraction for epidemiological research (DExtER): a novel tool for automated clinical epidemiology studies,” *Eur. J. Epidemiol.*, vol. 36, no. 2, pp. 165–178, 2021, doi: 10.1007/s10654-020-00677-6.