Development of a Covid-19 Information Dashboard to Access the Health Care Resources and Requirements Online

Krishna Sathya A C
Amrita Vishwa Vidyapeetham
Bengaluru, India
acsathya333@gmail.com

Parimi Mastan Rao
Amrita Vishwa Vidyapeetham
Bengaluru, India
pr_mastan@blr.amrita.edu

Shekar Babu
Amrita Vishwa Vidyapeetham
Bengaluru, India
sb@amrita.edu

2022 3rd International Conference on Computing, Analytics and Networks (ICAN) | 978-1-6654-9944-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICAN56228.2022.10007167

***Abstract*—With the emergence of coronavirus and the rapidly increasing number of cases, we observed that people were suffering from a lack of information about where to source critical requirements such as oxygen cylinders, beds, ambulance service, and ventilators. In this research paper and project, the authors have designed and developed an interactive information dashboard to access the availability of these resources and requirements at different locations across India from myriad sources. The information dashboard will show information for all the states across India. The dashboard visualizes the number of corona cases, hospital beds, blood bank, etc. The end-users and the COVID-19 support team can imagine all the requirements in the dashboard that is factual and credible within their vicinity from multiple information sources. The user need not search various information sites to search for different conditions. The project collates the data of other requirements from primary sources of information like Twitter and government websites and lists them on the dashboard. The authors used open-source programming and database technologies to display the data per user requirements. The dashboard was helpful to the end-users during COVID-19 times in terms of convenient accessibility and efficiency.**

Keywords—Covid-19, web scraping, information dashboard, data gathering, pandemic

I. INTRODUCTION

Over the past year, India has been facing one of the biggest challenges it has ever faced: ravaging the country by the dreadful pandemic. India has been witnessing a very high peak in COVID-19 cases where the double mutant / B.1.617 variant had a significant role in this massive causation. What followed the unexpected spike of the increasing positive cases was nothing but the 2nd Covid-19 wave, which had set an enormous record of approximately more than 400,000 daily cases, which made the 1st wave pale in comparison [1]. Owing to this vast surge, the Indian community responsible for healthcare was confronted with significant issues due to the lack of necessary healthcare resources required in the battle against the pandemic [2]. Meanwhile, from the Indian citizen's side, the main survival challenge faced was a lack of information. It often gave rise to questions like: Where should I go? How do I know which hospitals have available spots for beds /ICUs? Where can I procure oxygen cylinders? Where should I go to get Remdesivir? And a wide range of similar questions bombards the frazzled human minds.

It was not where no information was available; many individuals and government organizations across India have created numerous websites, applications, and even chatbots filled with data or information related to covid-19 necessities. Nevertheless, when faced with many of these resources, it isn't easy to search for the right help. Moreover, people don't have enough time to afford to search for the right individual resources across multiple designated sites.

Coronavirus has been one of the popular moving subject matters on the web since Jan 2020, and Twitter is no exception and has been constantly on the buzz and examined to date. So, in short, what Twitter does is that it brings constant updates to its platform to aid the population who are frantically looking for assistance amid the Covid-19 pandemic. Since the majority of the Indian population had taken to Twitter to seek out information regarding the availability of required resources like beds, plasma, oxygen concentrators, pills & medicine (Remdesiver, hydroxychloroquine, etc.), food and many others, Twitter had taken steps to launch a COVID-19 pandemic SOS which is nothing but an information resource page that was established to inform the citizens about the availability of the requirements sought out by them. Twitter also made efforts to display a prompt/signal in the user-associated timeline with links that redirect to sources giving out updates or information about Covid19 vaccinations. In addition to these, many other 3rd party tools function separately that aid users to quickly locate resources on a particular platform or site only where you can acquire basic amenities such as Oxygen beds, plasma, oxygen concentrators, and others needed in the battle against the crisis. Another reliable source of information is state government websites.

Here in this project, we have tried to identify and list all the information resources currently popular on the search site both Twitter and the web, thereby making it a one-stop platform displaying consolidated, efficient results. Then we utilize the apt data extraction technique for each case, API for Twitter, and traditional Web data scraping for government sites due to their unique web page layout structure across each government site.

II. RELATED WORK

In the Indian context, many other information dashboards have been developed by the central and also by many state Governments amidst the pandemic (Some

notable ones include: MyGov.in website released in 2021and also WHO Coronavirus (COVID-19) Dashboard deployed in 2021) [3], [4], [5]. Meanwhile, several academic institutions across the world and in India have taken initiatives to develop such dashboards for a social cause (Some of those notable institutes include: John Hopkins with their case tracking dashboard, Harvard T.H. Chan School of Public Health, 2020, and also IIT Delhi with their Resource dashboard 'PRACRITI 2.0' deployed in 2021 and several others). Some of those Dashboards were even created by specific volunteers (One such initiative was Covid tracking. In, 2021). Aside from the dimensions mentioned above, several other mediums have successfully provided insights into the virus's spread and impact. Those mediums include repositories in GitHub, Tableau visualizations, and ArcGIS data-driven pages.

Initiatives via crowdsourced means like WhatsApp and Telegram groups that actively took part in maintaining the constant updates concerning tracking and resource locating websites also played a vital role. One notable website for case tracking (covid19india.org) carefully organized the case count information made available in several state government bulletins, press releases, and official social media handles. It also makes the data open to public use by providing an API. So these are all considered notable efforts in information and communication technology that contribute toward the social good.

A. Function of Online Social Networks as a Source of Information

Right from the advent of social networks, that social network ecosystem has become a viral and engaging channel for communication and socialization [6]. One could view various information types being swapped among participants on social media platforms, from personal contexts through messaging, feelings through expressions, images, and videos, to official and institutional contexts. So, no wonder social networks have become one of the highly go-to sources of real-time information. As per the 2020 statistics, among the highly prominent social networks concerning user volume, Facebook (FB) is 1st with a total of 2449 million, followed by YouTube, ranking second with about 2000 million, then Instagram is also in the same boat as YouTube with a whopping 2000 million while the Twitter platform has only about 340 million with a considerable volume gap than the before mentioned platform. Meanwhile, other country-specific social network platforms like Qzone(QQ), Sina Weibo(a Chinese microblogging site), and Baidu Teiba/ Baidu post bar communication channel all together managed to accumulate a substantial number of followers, which is about 1000 million [7].

B. Data Extraction Methods

In today's world, extracting data from the wide range of sources available on the web is necessary concerning most of the processes involved in a business, Research and Development (R&D), studies, and many others. So this data extraction process encompasses retrieving relevant data content / even metadata, which might be helpful across a diverse set of purposes set out to achieve. Two unique techniques and methods for this include traditional web scraping and API as a tool. It is common knowledge that both web-scraping and APIs are techniques that involve automation and are viewed to be a few of the most practical ways in which data can be harvested [8]. These methods facilitate data collection from several web pages and data repositories quickly and accurately. Once extraction is done, the extracted data is stored securely in a database for future analysis and usage. Compared to these automated techniques removing information manually does not hold a candle to accuracy and time.

C. Applications and Shortcomings of Web Scraping

Web scraping is a systemic fashion in which extraction and combining of content under consideration from the web takes place. During this procedure, a user agent, a web scraping bot, or a scraping script simulates the browsing interaction between the web servers and the human clients concerning a traditional web-based traversal. The robot/script can gain access to multiple sites as per the user's intent, and then next, the contents are parsed to locate and extract the data contents and align/structure those contents as the user intends to. These all occur in a step-by-step fashion [9]. Web scrapers are very apt and helpful in quickly extracting and processing vast amounts of data from a particular site. Thus, for any information that is portrayed on a web browser and as long as the specific website has provided them permission to crawl by gaining access to the robots.txt file, then the website information can be explicitly accessed through a robot/script to retrieve the data contents and dump those contents in a database and then utilized for future use or analysis [9]

Web scraping is applied mainly on web pages that make use of structured mark-up languages like (Hyper-Text Markup Language) HTML [or] (eXtensible HyperText Markup Language) XHTML. So here, in this situation, the scraping process involves: locating and parsing the hypertext HTML tags and extracting the plain text information encapsulated within those tags. An HTTP protocol is utilized for the web scraper script to set up a communication line with the intended target website. Sometimes, regular expression matching coupled with extra logic is also needed when extracting the data of interest [9]. One of the well-known limitations of web scraping is network speed because it directly influences how much time the website takes to load and how exactly its contents will be structured in case of a low rate. Since conventional web scraping is based on website-specific programming, it naturally does not comply with any changes that may occur on the HTML source page [9], [10]. So when this happens, the web scraping script should be continuously updated and ensure that it is adaptable to the new format or structural layout of the website information. This maintenance of the web scraping process is vital and is mainly viewed as a tedious, time-consuming programming activity. In addition, web scraping developers must also make it a priority to look into the legal & policy aspects of the data they intend to extract from a website to avoid copyright infringement issues [9]. Moreover, one important thing to note is that there is no significant distinction between a user visiting a

website or a web scraping developer scraping one because in both instances, the user is treated as a guest, thereby clearly outlining that they are not the actual owner of the data/information content being extracted.

D. Applications and Shortcomings of API Tool

Inherently an API, a module coming under object-oriented programming language, facilitates software developers to help construct a specific software application via a library associated with a reference program [11]. The API is utilized by an operating system or any software application in which the client user/another device usually makes relevant requests and, in turn, expects responses from the application side. APIs accommodate communication between various software applications and the permission of their respective services. It mainly encompasses the specification/parameters of data structures, protocols, object classes, and runtimes, ensuring that the customer communicates about the API's resources [12]. Extending existing courses to accommodate the addition of new features/ functionalities or constructing new systems can also be done by developers. So when a client API is called via an endpoint, the component listens when a client-side request is made to the server-side through a Hypertext Transfer Protocol, expecting the response for the demands being made. Aside from the most popular ones when viewing social networks and web information sources, most of them do not offer an API for gaining access to the available content for various reasons [12], such as the desired data in a contest being small or uncommon (b) the source being not capable enough to facilitate the appropriate infrastructure / technical capability required for building an API (c) when the data intended to be scraped either valuable or safeguarded and not meant to be laid out in the open and spread widely (d) even upon the existence of an API, there bound to be certain limitations not only on the volume of requests but also on the rate limits. Moreover, the resultant types and data format that the API provides might not be enough to achieve the intended purpose. In addition to this, there exist certain limitations concerning APIs that involve information is not exercised to its full capability / not enough or even fittingly demonstrated. The access is further limited due to the existence of data protection laws associated with privacy and also by the majority of the Terms of Service [13]

III. Proposed Work (Building An Automatic Resource Data Gathering Pipeline)

Unlike the previous cases where data seems to be fragmented in terms of the specific region-related dashboards and specific social media platform-related dashboards, we aimed to design a one-stop COVID-19 information dashboard that sources credible information across multiple points (like Twitter & multiple government websites) which helps the user to gain verified details of the health resource they want within their vicinity.

A. Proof of Concept Architecture and Implementation

Fig. 1 portrays the implementation outline of this project. It depicts how the information flows from the backend to the front end. A detailed explanation of the functionalities in each module is listed in the sections below.

B. Identify Top 'N' Current Resource Keyword

The first step is to identify which data you want to extract. So one should settle on the top 11 highly searched current resource keywords on social media sites like Twitter and also by finalizing on a resource based on factors like rising demand numbers & scarcity of specific resources like saying plasma and some rare blood groups, which mostly seemed hard to find during this pandemic. The list of top 11 requirement keywords we had chosen includes Oxygen cylinders, concentrators, general ward beds, ICU beds, Oxygen Beds, ventilators, Ambulance, Blood, Amphotericin, Remdesivir, and Tocilizumab.

C. Gather Data Sources

Collect a list of websites or platforms from where the availability information of the healthcare resources data for the chosen top 11 keywords can be sourced. Among all social media platforms, the authors had settled on Twitter owing to the highly active tweet ecosystem in terms of announcing /discussing the availability of resources. The authors also collected the list of all possible government websites categorized by each state to expand our possibilities to search. The authors also aimed to increase the odds of obtaining more credible information for specific resources that might not be often available on Twitter.

D. Gather Data Sources

Collect a list of websites or platforms from where the availability information of the healthcare resources data for the chosen top 11 keywords can be sourced. Among all social media platforms, the authors had settled on Twitter owing to the highly active tweet ecosystem in terms of announcing /discussing the availability of resources. The authors also collected the list of all possible government websites categorized by each state to expand our possibilities to search. The authors also aimed to increase the odds of obtaining more credible information for specific resources that might not be often available on Twitter.

E. Data Extraction From Twitter

Twitter: Data can be extracted from Twitter using either Twitter API or web scraping to extract the data and perform further analysis. Both extraction techniques have advantages and disadvantages. Regarding time performance, web scraping is faster than Twitter API and is more flexible in obtaining data; however, web scraping is susceptible to website changes. Since the authors can't afford to change the web scraping code whenever Twitter makes changes concerning website layout, authors are willing to trade out the time difference worth a few seconds. So the authors opted for the Twitter API approach, which is more straightforward to use where it is possible to obtain a variety of information only by knowing the tweet ID or user ID.

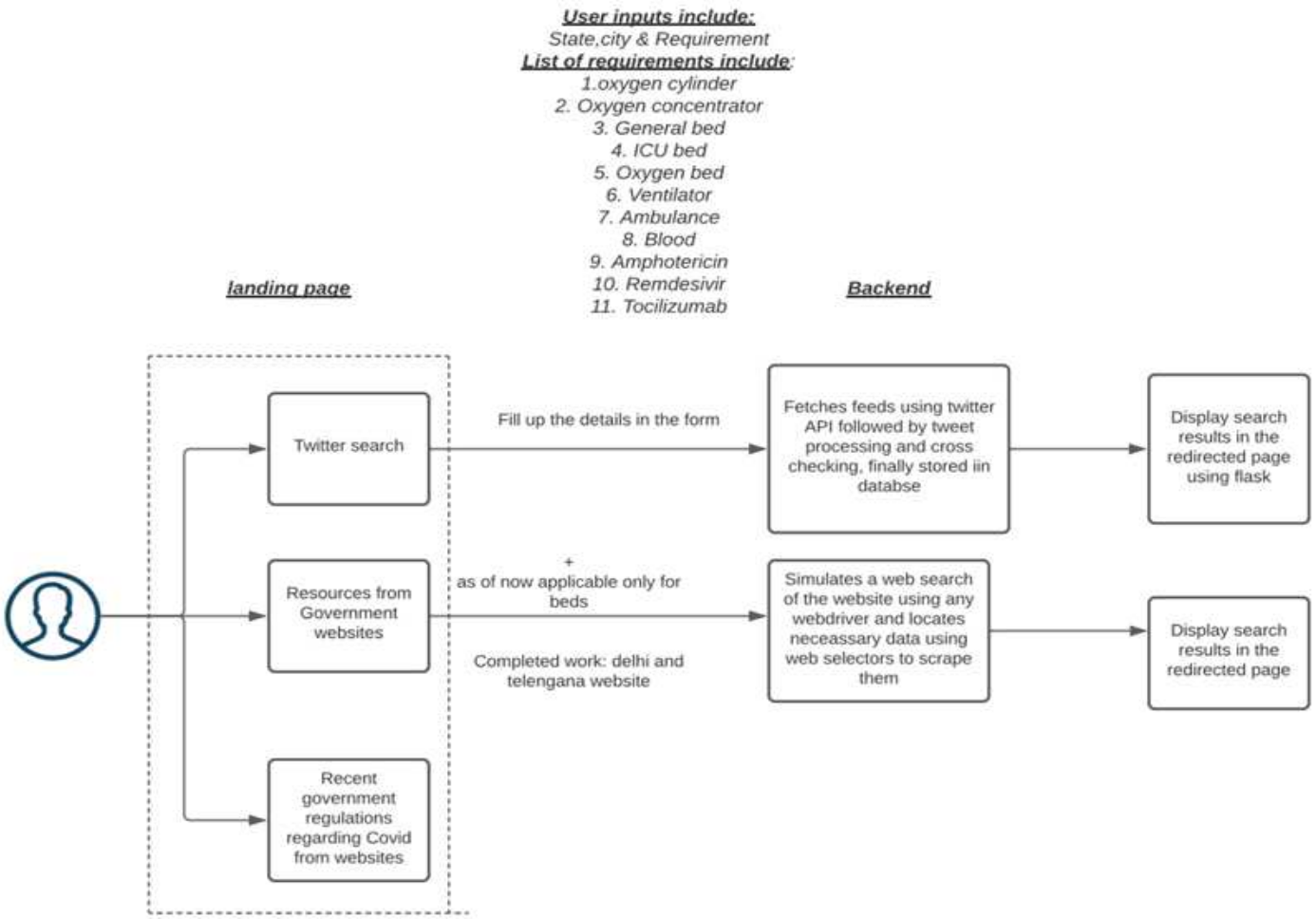

Fig. 1. Concept architecture

Generating a twitter search query: Twitter has an advanced search option through which we can refine or narrow down our search results using any combination of 4 broad fields: words, people, places, and dates. Some well-known advanced search query options include Tweets excluding any negative keywords, Tweets containing exact phrases, Tweets with a specific hashtag, Tweets from a particular account, Tweets sent from a geographic location (e.g., a specific city, state, country), and many more advanced options. The Twitter search query we composed can be decomposed into three variables:

- Wanted - This wanted variable includes all the must-have keywords like 'verified' or 'available.'
- Unwanted - This variable includes the list of keywords that should not be present in our list of tweets like not verified, unverified, help, please, Does anyone know, etc. These negative keywords were primarily excluded because they include tweets that are not reliable, or they include tweets that are just requesting help and not indicating any resource availability.
- Category - This variable includes the list of vital keywords in a user's search query. The index consists of Requirements: What resource is the user searching for? (Oxygen cylinder, ICU bed, etc.); State: The state where the user wants to search for the availability (Tamil Nadu, Karnataka, etc.); City: (Chennai, Delhi, etc.).

Finally, the authors concatenate these three variables to generate a wholesome refined twitter search query. So this twitter query generator can create questions for all combinations of requirements, state, and city as per the user's input in the HTML form, thereby keeping the wanted and unwanted variable alone static. One such example of a Twitter query when a user searches for an oxygen concentrator in Chennai, Tamil Nadu, is listed in Fig. 2.

```
(verified OR available) (Telangana OR Hyderabad) ventilator -"know" -"Does anyone" -"where" -"Seeking"-"want"-"Requires"-"wan
ted"-"Require"-"Required"-"Please"-"Plz"-"Urgent"-"need"-"Any verified lead"-"not available"-"filter:retweets"
```

Fig.2. Generated twitter query as per user's input requirements.

Simulating the Twitter search: On constructing the Twitter query, one can simulate the search request using the Twitter API credentials (See Fig 3) and specifying the attributes required to obtain the search results in our intended format.

```
#The standard API only allows you to retrieve tweets up to 7 days ago and is limited to scraping 18,000 tweets per a
tweets = tw.Cursor(api.search,tweet_mode='extended',q = search_word, lang ='en').items(100)
tdetails=[[tweet.id,tweet.user.screen_name,tweet.created_at,tweet.full_text,tweet.user.location]for tweet in tweets]
```

Fig. 3. Twitter query call function

Tweet pre-processing: The next step involves filtering and cleaning the unprocessed tweet content. Initially, upon the Twitter API search using the generated twitter search query, the authors can obtain four fundamental fields: Tweet ID, username, tweet date, tweet content, and finally, the location of who posted the tweet. The purpose of cleaning is not only to get extended attributes like verification status, contact info, and tweet URL from the tweet content variable but also to recheck or verify if the must-have categories are present correctly in the tweet, like the state city, requirement, etc. Another primary purpose of cleaning is removing redundant tweets, whether retweets or the same tweet content posted by different people. A sample of initial dataset fields before preprocessing and after pre-processing is given in Fig. 4 and Fig. 5.

F. Data Extraction from Government Websites:

Unlike the previous data extraction process in Twitter here, the authors used traditional web scraping because each government website has a different structure and layout, so no static API exists for efficiently extracting the data. So this challenge calls for writing separate web scraping code snippets tailored exclusively for each government website which takes a lot of time. So, the authors extracted data from a few such government websites. Below listed Fig. 6 and Fig. 7 are a few example snippets from Telangana and Delhi government websites.

G. Cloud Storage and Process Automation

The entire web scraping process across government websites and Twitter is automated so that they are scheduled to run every few minutes in a multiple thread fashion. The scraped and processed results are stored in the Mongo DB Cloud database. The MongoDB cloud functions dump their results into the allotted database, keeping the day-to-day historical data. Then, the data is fetched, organized, and displayed as per the user's requirement in the front end, which is on the dashboard.

	ID	username	Tweet posted date	Tweet content	location
0	1393292852587757568	d_prajith	2021-05-14 19:51:16	@Sharanyachander @AnandNodal @ashokhavarshini ...	nellore
1	1393104313438728195	ddileepkumar23	2021-05-14 07:22:05	@PegsRmkr ✅ **Verified**\nCHENNAI\nOxygen Beds Ava...	Andhra Pradesh, India
2	1392940937345703937	dank_azadi	2021-05-13 20:32:53	@kirthanaaaaa This:\n Oxygen Beds Available \n...	
3	1392919000196202503	OwnsWhat	2021-05-13 19:05:43	@ThinkTinkAcad @TellDM #Chennai\nOxygen Beds A...	
4	1392905932158377985	CoatsVasa	2021-05-13 18:13:47	@Preetikaduggal @BabaJotwani @findcovidleads @...	Banjara Hills, Hyderabad
5	1392905724523486753	CoatsVasa	2021-05-13 18:12:58	@Shwetha_Baskar #CHENNAI #TeamIndiaVsCovid \nO...	Banjara Hills, Hyderabad
6	1392905672505716739	CoatsVasa	2021-05-13 18:12:45	@yblk @104_GoTN @Railganesan @skpkaruna @Capta...	Banjara Hills, Hyderabad
7	1392856377765158914	WaliaDivyadeep	2021-05-13 14:56:53	@ReallySwara @srinivasiyc CHENNAI\nOxygen Beds...	
8	1392842982462689284	abhishekaryan17	2021-05-13 14:03:39	✅ **Verified** CHENNAI Oxygen Beds Available (SPO2...	Patna/Muzaffarpur/ Delhi
9	1392836910704259072	COVResourcesIn	2021-05-13 13:39:31	✅ **Verified**\nCHENNAI\nOxygen Beds Available (SP...	India
10	1392832675111600133	NaveenAnakala	2021-05-13 13:22:41	✅ **Verified**\nCHENNAI\nOxygen Beds Available \n(...	Hyderabad, India

Fig. 4. Before pre-processing format

	Tweet ID	username	Tweet posted date	Tweet content	Requirement	Verified	City	State	Phone number	URL
0	1394208176195575809	GirlUpWajood	2021-05-17 08:28:26	#COVIDLEADS Home #ICU Set up Delhi NCR WhatsA...	Oxygen concentrator	no	Delhi	Delhi	9654535285, 8826513407, 9810526521	https://twitter.com/user/status/1394208176195...
1	1394198983736631298	shivanjalingo	2021-05-17 07:51:55	#Oxygen Concentrator Available #Delhi Contact...	Oxygen concentrator	yes	Delhi	Delhi	9560976288	https://twitter.com/user/status/1394198983736...
2	1394198297733001223	shivanjalingo	2021-05-17 07:49:11	#Oxygen Concentrator Available #Delhi Contact...	Oxygen concentrator	yes	Delhi	Delhi	9958065400	https://twitter.com/user/status/1394198297733...
9	1393981465998675972	nandhaVishvaraj	2021-05-16 17:27:35	✅ **Verified** by many user check Kar lena DELHI ...	Oxygen concentrator	no	Delhi	Delhi	9999028878	https://twitter.com/user/status/1393981465998...
10	1393976436159963136	sunal_sharma	2021-05-16 17:07:35	#Oxygen #Concentrator available Capacity 5 an...	Oxygen concentrator	yes	Delhi	Delhi	8285952556, 9599289940	https://twitter.com/user/status/1393976436159...
12	1393972689363083267	SaiVinay072	2021-05-16 16:52:42	✅ **Verified** NEW DELHI Oxygen Concentrator Av...	Oxygen concentrator	yes	Delhi	Delhi	9999119655	https://twitter.com/user/status/1393972689363...

Fig. 5. Post pre-processing and verification

	S.No.	Hospital Type	Hospital Name	District	Last Updated	Total	Occupied	Vacant	Contact
0	1	Private	AADHYA CHILDRENS HOSPITAL	Warangal Rural	18-05-2021, 06:29:25 PM	2	0	2	Dr.Nehru, 9989840732
1	2	Private	AAROGYA HOSPITAL	Siddipet	18-05-2021, 06:18:54 PM	5	0	5	Dr.Vijaya Shekar, 9000013945
2	3	Private	AARON HOSPITAL-NARSINGI	Rangareddy	17-05-2021, 11:55:00 AM	10	7	3	N/A, N/A
3	4	Private	AASHRITHA HOSPITAL,MALLAPUR	Medchal	22-05-2021, 10:52:48 AM	5	0	5	Medchal, N/A, 9502742949
4	5	Private	ABHAYA MALTISPECIALITY HOSPITAL, PRAKASHAM BAZ...	Nalgonda	23-05-2021, 07:03:29 AM	6	1	5	Private, Nalgonda, N/A, 9000265699
...	...	...	...	...	...	...	...	...	...
1105	1106	Private	YASHODA HOSPITALS - SOMAJIGUDA	Hyderabad	21-05-2021, 08:09:35 AM	168	168	0	DR SRIDHAR PODISHETTY, 9989975559/9390006070
1106	1107	Govt.	YELLANDU CHC	Badradri	23-05-2021, 08:57:06 AM	30	9	21	DR. G.VARUN, 9573735259
1107	1108	Private	YELLAREDDY SUPER SPECIALITY, YELLAREDDY, NIZAM...	Kamareddy	23-05-2021, 08:42:42 AM	3	0	3	Kamareddy, N/A, 9908932322
1108	1109	Private	ZOI, ATTAPUR	Rangareddy	22-05-2021, 09:13:32 AM	20	15	5	Rangareddy, N/A, 9059207482
1109	1110	Private	ZOI HOSPITAL - SOMAJIGUDA	Hyderabad	22-05-2021, 06:38:02 PM	34	30	4	N/A, 9849029580

			REGULAR BEDS			OXYGEN BEDS			ICU BEDS (Ventilator/ CPAP)			TOTAL BEDS		
	S.No.	HOSPITAL	TOTAL	OCCUPIED	VACANT	TOTAL	OCCUPIED	VACANT	TOTAL	OCCUPIED	VACANT	TOTAL	OCCUPIED	VACANT
a	1	GOVERNMENT HOSPITAL	5526	1302	4224	7670	5235	2435	2143	1523	620	15339	8060	7279
b	2	PRIVATE HOSPITAL	16178	3641	12537	13848	7733	6115	9478	6174	3304	39504	17548	21956

Fig. 6. Telangana website bed availability status

	Hospital	Last Updated	Total	Vacant	Oxygen left for	Contact
0	Delhi Govt GTB Hospital	6:01 PM, May 24	900	508\n\nContact	1 day	8595948014 8595948019 8595948020 8595948026 85...
1	Delhi Govt Lok Nayak Hospital	6:09 PM, May 24	900	464\n\nContact	12 hours	8046820570
2	Delhi Govt Rajeev Gandhi Super Speciality Hosp...	7:56 PM, May 24	325	224\n\nContact	6 hours	01122890773, 01122890789, 01122890611 , 859552...
3	Delhi Govt Deen Dayal Upadhyay hospital	6:27 PM, May 24	250	91\n\nContact	16 hours	1125401075
4	Delhi Govt Dr Baba Saheb Ambedkar Hospital	6:29 PM, May 24	200	56\n\nContact	20 hours	7303885606
...	...	...	...	...	...	...
174	Pvt Pt.RRM Hospital	5/24/2021 10:53	2	2\n\nContact	3 days 18 hours	8751002222
175	Pvt Tulsi Multispeciality Hospital & Critical ...	7:13 PM, May 24	2	2\n\nContact	23 hours	8826824672, 8447720118, 9990495495
176	Pvt Angel Agasta Multispeciality Hospital	2:02 PM, May 24	1	1\n\nContact	3 hours	01141222222, 8130390341
177	Pvt Parmeshwari Devi Hospital	5/24/2021 10:09	1	0\n\nContact	NaN	9582585676, 01127550441, 01127550442

Fig. 7. Delhi website ICU bed availability status

H. Front-End Development Website

With the advent of COVID-19, premade tutorials and templates for GIS software and tools on quickly creating a dashboard about COVID-19 came into existence[12]. However, the authors built the dashboard from scratch. The author used the Flask framework to create an interactive and

dynamic web application and later integrated it with the MongoDB database. Concerning the website layout, the authors utilized the well-known Bootstrap framework. In addition, the authors incorporated certain JS events (For example, a refresh SVG shape to activate the refresh button). Fig. 8 shows our web interface.

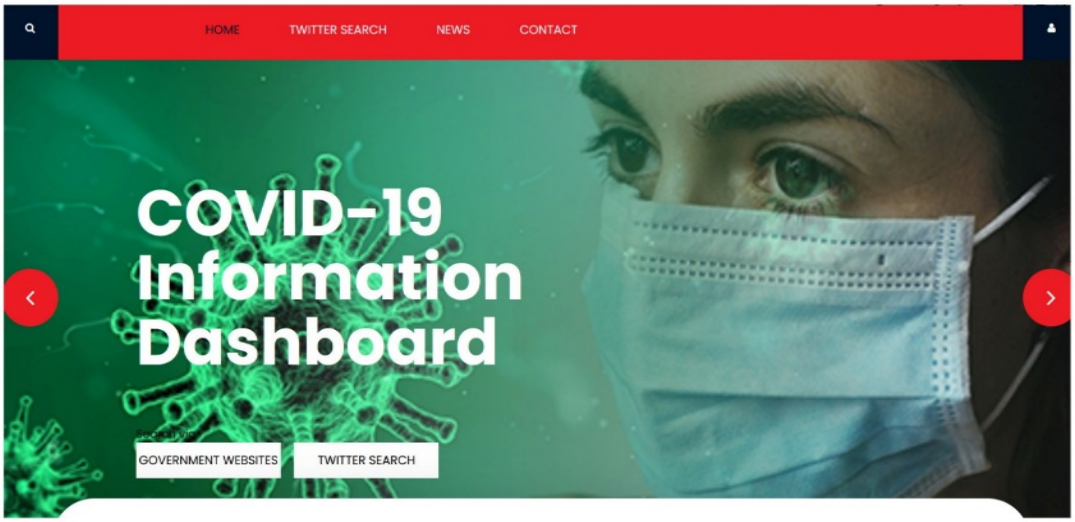

Fig. 8. Web Interface

IV. Conclusion

Designing and maintaining a consolidated resource data convening system responsibly in a demanding time constraint setup is challenging, mainly in times of substantial changes and the far-flung unrestful state as in a worldwide pandemic. In short, this paper portrays the process behind the development and concept architecture of a convenient use one-stop COVID-19 information dashboard developed for the public in general. The authors were quick enough to adapt to the ever-changing pandemic situations, suitable interface layout, and user demands from their perspective and maintain the automated, scalable and convenient platform. From the development process the authors depicted so far, a few main challenges include: Automating the data gathering process is no easy task as collecting resources from over 28 states across India. It requires substantial programming effort for such a large scale.

Moreover, another challenge was that certain official government source websites kept changing their layout, destroying the scraping code done earlier. This calls for vigilant lookouts and constant updates. Another main challenge is the choice of technology to substitute in the stack if authors look at the project with a short vision and choose the wrong technology to work with, only to realize later in the far end stage of the project that the technology poses certain constraints that may cost us in terms of cost and time.

References

[1] L. John and S. Bhauma, "Indian's struggle to track new COVID variants could worsen crisis," Bloomberg, 05 05 2021. [Online]. Available: https://www.bloomberg.com/news/articles/2021-05-05/vaccineswork-work-on-this-india-variant-experts-fret-about-the-next. [Accessed 06 05 2021].

[2] P. Vikas and N. Shadab, "Covid-19 in India - why thw second coronavirus wave is devastating," BBC, [Online]. Available: https://www.bbc.com/news/world-asia-india56811315.amp. [Accessed 30 04 2021].

[3] Government of India, MyGov India, "National Informatics Centre, Ministry of Electronics and Information Technology," [Online]. Available: https://self4society.mygov.in/. [Accessed 15 05 2021].

[4] Government of Wesrt Bengal, "Integrated Covid Management System (WB-ICMS)," [Online]. Available: https://excise.wb.gov.in/chms/Portal_Default.aspx. [Accessed 30 04 2021].

[5] India Covid SOS, "COVID-19 Resources Charbot," [Online]. Available: https://www.covidsos.live/app/index.html. [Accessed 30 04 2021].

[6] M. Alrubaian, M. Al-Qurishi, A. Alamri, M. Hassan and A. Alamri, "A real-time tweets credibility assessing system," in *International Conference on Advances in Social Networks Analysis and Mining*, 2016.

[7] M. Alrubaian, M. Al-Qurishni, A. Alamri, M. Hassam and G. Fortino, "Credibility in online social networks - a survey," *IEEE Access,* vol. 7, pp. 2828-2855, 2019.

[8] C. Slamet, R. Andrain, D. Maylawati, Suhendar, W. Darmalaksana and M. Ramdhani, "Web scraping and Naive Bayes classification for job search engine," in *IOP Conference Series*, 2018.

[9] D. Glez-Pena, A. Lourenco, H. Lopez-Fernandez, M. Reboiro-Jato and F. Fdez-Riverola, "Web scraping technologies in an API world," *Bioinformatics,* vol. 15, no. 5, pp. 788-797, 2013.

[10] R. Mitchell, Web scraping with python - collecting data from the modern web, O'Reilly Media Inc., 2015.

[11] D. Salt and A. Sellhorn, "Method, system and computer program product for a client application programming interface (API) in a service-oriented archecture". US Patent 8701128.

[12] L. Dongo, Y. Cardinale and A. Aguilera, "Credibility analysis for available information sources on the web - a review and a contribution," in *International Comnference on System Reliability and Safety*, 2019.

[13] Health GeoLab Collaborative, "ArcGIS Solution Coronavirus Case Dashboard," [Online]. Available: https://healthgeolab.net/DOCUMENTS/HGLC_Coronavirus_Case_D. [Accessed 07 05 2021].