JMIR Publications
The Leading eHealth Publisher

COVID-19 Knowledge Resource Categorization and Tracking: A Conceptual Framework

Muhammad Afzal, Maqbool Hussain, Jamil Hussain, Sungyoung Lee

Submitted to: Journal of Medical Internet Research
on: April 18, 2021

Preprint
JMIR Publications

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on it's website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressively prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Preprint
JMIR Publications

COVID-19 Knowledge Resource Categorization and Tracking: A Conceptual Framework

Muhammad Afzal[1] PhD; Maqbool Hussain[1] PhD; Jamil Hussain[2] PhD; Sungyoung Lee[3] PhD

[1]Department of Software Sejong University Seoul KR
[2]Department of Data Science Sejong University Seoul KR
[3]Department of Computer Science and Engineering Kyung Hee University Yongin KR

Corresponding Author:
Muhammad Afzal PhD
Department of Software
Sejong University
Office 422 Innovation Center
Seoul
KR

Preprint JMIR Publications

Abstract

Background: Declaring the COVID-19 disease a global pandemic by the World Health Organization (WHO), it gained momentum as every day passed, and private and government sectors of different countries pushed funding towards research in various capacities. A great portion of efforts is devoted to information technology and service infrastructure development, including research to develop intelligent models and techniques for alerts, monitoring, early diagnosis, prevention, and other relevant services. As a result, tons of information resource have been created in the global space and are available for use. However, there is lack of a defined structure to organize these resources into categories or classes based on the nature as well the origin of data.

Objective: This study aims to organize COVID-19 information resources into a well-defined structure to facilitate easy identification of a resource, tracing information workflows, and a guide for contextual dashboards design and development.

Methods: A sequence of action research was performed that involve a review of COVID-19 efforts and initiatives on a global scale during the year 2020. Data is collected according to a defined structure of primary, secondary, and tertiary categories. Various descriptive statistical analysis techniques were employed to get insights of the data to help in developing a conceptual framework underlining the organization of resources and interactions among different resources.

Results: In this paper, we present a three-level structure of resource categorization that provides a gateway to access the global initiatives with enriched metadata, assists users in tracing the workflow of tertiary, secondary, and primary resources with relationships among various fragments of information. The results comprise mapping initiatives at the tertiary level to secondary and then to the primary level to reach the firsthand resource of data, research, and trials.

Conclusions: Adopting the proposed three-level structure enables a consistent organization and management of existing COVID-19 knowledge resources and provides a roadmap for classifying the futuristic resources. This study is one of the earliest studies to introduce an organized structure and demonstrate the placement of COVID-19 resources at the right place. By implementing the proposed framework according to the stated guidelines, this study facilitates the development of applications such as interactive dashboards to facilitate the contextual identification and tracking of interdependent COVID-19 information resources.

(JMIR Preprints 18/04/2021:29730)
DOI: https://doi.org/10.2196/preprints.29730

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Preprint JMIR Publications

Original Manuscript

Preprint
JMIR Publications

COVID-19 Knowledge Resource Categorization and Tracking: A Conceptual Framework

Muhammad Afzal[1], Maqbool Hussain[1], Jamil Hussain[2], Sungyoung Lee[3]
[1] Department of Software, Sejong University, Seoul, South Korea
[2] Department of Data Science, Sejong University, Seoul, South Korea
[2] Department of Computer Science and Engineering, Kyung Hee University, Yongin, South Korea

Background: Declaring the COVID-19 disease a global pandemic by the World Health Organization (WHO), it gained momentum as every day passed, and private and government sectors of different countries pushed funding towards research in various capacities. A great portion of efforts is devoted to information technology and service infrastructure development, including research to develop intelligent models and techniques for alerts, monitoring, early diagnosis, prevention, and other relevant services. As a result, tons of information resource have been created in the global space and are available for use. However, there is lack of a defined structure to organize these resources into categories or classes based on the nature as well the origin of data.
Objective: This study aims to organize COVID-19 information resources into a well-defined structure to facilitate easy identification of a resource, tracing information workflows, and a guide for contextual dashboards design and development.
Methods: A sequence of action research was performed that involve a review of COVID-19 efforts and initiatives on a global scale during the year 2020. Data is collected according to a defined structure of primary, secondary, and tertiary categories. Various descriptive statistical analysis techniques were employed to get insights of the data to help in developing a conceptual framework underlining the organization of resources and interactions among different resources.
Results: In this paper, we present a three-level structure of resource categorization that provides a gateway to access the global initiatives with enriched metadata, assists users in tracing the workflow of tertiary, secondary, and primary resources with relationships among various fragments of information. The results comprise mapping initiatives at the tertiary level to secondary and then to the primary level to reach the firsthand resource of data, research, and trials.
Conclusion: Adopting the proposed three-level structure enables a consistent organization and management of existing COVID-19 knowledge resources and provides a roadmap for classifying the futuristic resources. This study is one of the earliest studies to introduce an organized structure and demonstrate the placement of COVID-19 resources at the right place. By implementing the proposed framework according to the stated guidelines, this study facilitates the development of applications such as interactive dashboards to facilitate the contextual identification and tracking of interdependent COVID-19 information resources.
Keywords – information organization, resource management, knowledge graphs, interactive dashboard, dependency tracking, COVID-19

Introduction

The novel Coronavirus (SARS-CoV-2) first appeared in December 2019 and has quickly spread over other regions of the world. The World Health Organization (WHO) declared the COVID-19 disease, caused by the novel Coronavirus, a global pandemic [1]. It gained momentum as every day passed, and private and government sectors of different countries pushed funding towards research on COVID-19 in various capacities. A portion was dedicated to investing in vaccine discovery and personal protective equipment (PPE) manufacturing. The other portion was devoted to information

technology and service infrastructure development, including research to develop intelligent models and techniques for alerts, diagnosis, treatment, prognosis, prevention, and other relevant services. This study focuses on the second part of the research initiatives related to COVID-19 and provides a comprehensive review of such initiatives.

A rapid and timely response was required from the world to circumvent the challenge of COVID-19. Various organizations from different countries concentrated on investigations and discoveries in terms of data, information, and knowledge to support population health. We design a structure for categorizing efforts and initiatives related to COVID-19 into three levels or categories: primary, secondary, and tertiary. Primary level resources represent the initiatives that accrue raw data/research about COVID-19 patients and bring it to the global space. Secondary level resources encompass initiatives that analyze primary level resources making it more meaningful by adding metadata and filtering unnecessary items in data. Tertiary-level resources include initiatives that consolidate the first and second-level efforts by creating guidelines, code systems, standard resources, and vocabularies. This study provides a review of initiatives globally at three different levels and proposes a novel research dashboard for tracking novel corona resources with dependency workflows. The proposed dashboard will provide a gateway to global initiatives with enriched contextual metadata to help users track the information flow at different levels for validation and verification.

This study is the first of its kind to categorize efforts and initiatives in the form of resources related to COVID-19. It takes inspiration from the research sources explained in commentaries and libraries' handouts [2–6]. This study improvises the three types of information sources to map COVID-19 related initiatives to the individual level and devises a method of finding their interdependence for easy tracking of information in contextual trails and meta-information. Related to COVID-19, we can find studies that discuss datasets and techniques, but no such study is found to provide a categorization of initiatives and their dependencies.

This study is structured to describe three types of resources: primary resources in section 2, secondary resources in section 3, and tertiary resources in section 4. Section 5 describes the idea of a contextual dashboard as a future direction with associated challenges and prospect solutions. Section 6 concludes the study. The three sections, 2, 3, and 4 are primarily informative; however, an aspect-oriented comparative summary is described at the end of each section. Section 5 is an important section from the outcomes perspective that guides users to propose their solutions using the resource infrastructure discussed in this study.

Review of COVID-19 efforts and initiatives

This study explores initiatives through an informal search strategy to identify key initiatives at each level; however, the authors do not claim the list is exhaustive. The authors' experience in evidence-based medicine and translational research made it easy to reach out to well-known data resources and look for COVID-19 related efforts. The study does not distinguish the type of data and information pertinent to a specific initiative; instead, it focuses on resources. For different resources, different search strategies are employed. For example, the primary level studies are searched in top publisher networks by employing the procedure described in Figure 1. For secondary and tertiary level resources, a mixed method of free searching, word-of-mouth, and expert recommendations.

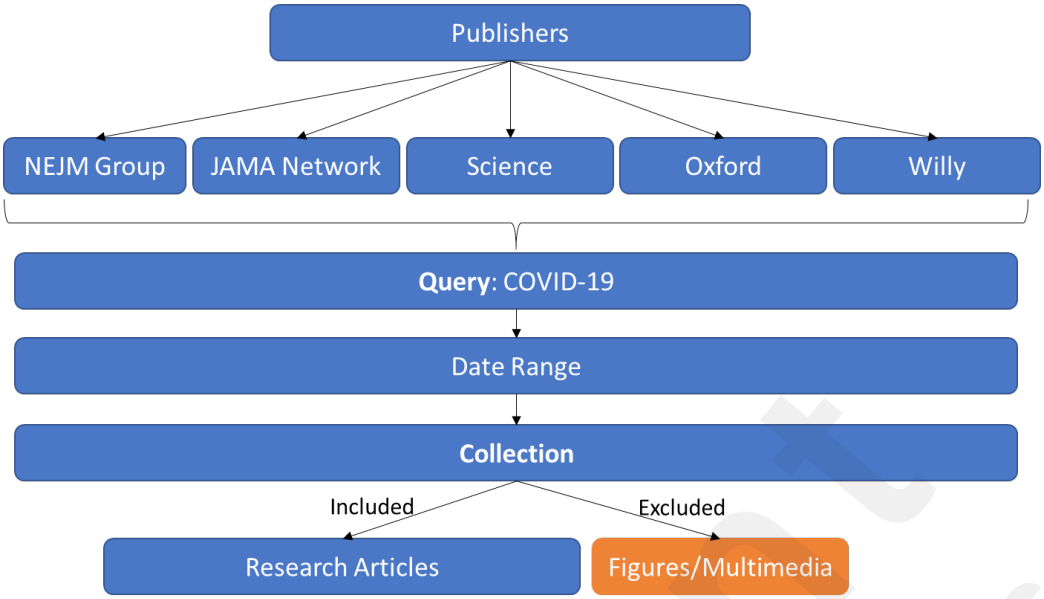

Figure 1: Publisher wise searching for published research articles on COVID-19

Primary level resources

The primary level includes resources that accumulate raw data/research about COVID-19 patients and bring it to the global space. The primary level resources are categorized into three groups: research literature, data, and clinical trials. The research literature includes published information like journal publications, raw data refers to firsthand patient information like demographic and clinical information, and clinical trials hold the in registered clinical trials for proving the effectiveness of treatment like drug or surgery. We provide a statistical overview of different resources in each group.

Research Literature

The research literature is divided into two categories: preprint and post-print. The preprint articles are “work in progress” research that is yet to be published in a peer-review journal. The post-print articles are the “submitted to journal” research that is considered published after going through copy editing and typeset formatting.

Post-print (published) literature

As of now, a vast number of publications mounted in the biomedical literature mentioning “COVID-19” in the titles, or the body of text, or the list of keywords and meta-information. From January 1, 2020, to December 31, 2020, about 12000 articles were published in prominent and focused research publishing groups that include Nature, Science, Cell, NEJM, JAMA, Lancet, and BMJ. As shown in Figure 2, Nature publishes a total of 3951 articles contributed about 34%, followed by Lancet, with about 19% in the given group of seven publishers. This quantitative representation does not discuss the quality of the articles. Later in this study, we will discuss the Covid19-relevant and Covid19-specific articles. The articles' progression was increasing every month initially, and a maximum number of articles appeared between May and October 2020. Starting with January 2020, we have hardly any substantive articles specific to COVID-19, but it increases to a two-digit number in February. A maximum number is observed in October with 1364 articles. The articles are still publishing at a steady rate.

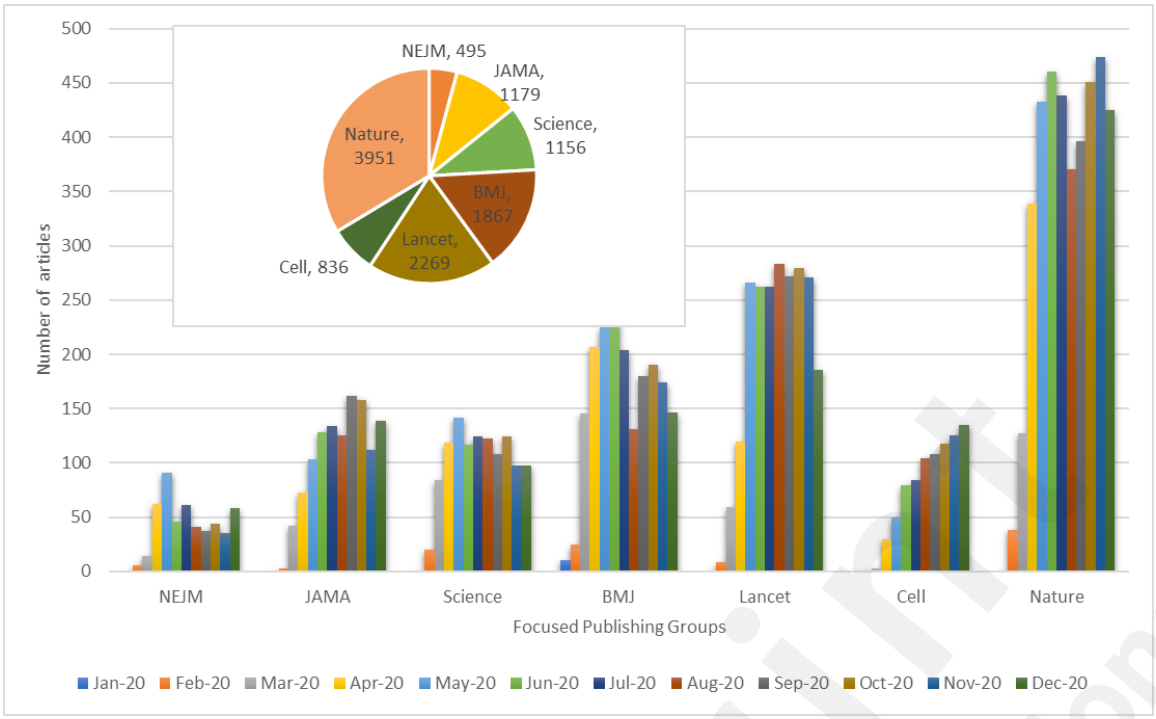

Figure 2: Monthly progress of publications on COVID-19 by focused publishing groups

Many articles are added to major libraries daily. We get a survey of eight major scientific libraries: ACM, Wiley, Springer, Oxford, SAGE, Elsevier, MDPI, and IEEE. All these databases are accessible to researchers; therefore, we opted to present their statistics. Note that for these searches, we use the term “COVID-19” to keep the search simple as we are not supposed to be wary about missing articles in searching; presumably, there must be a relative minimum in number. As shown in Figure 3, alone ACM identifies about 20000 entries to COVID-19 related articles. These numbers indicate that COVID-19 related articles are published everywhere, including the IEEE Explore though its focus is more on engineering discipline.

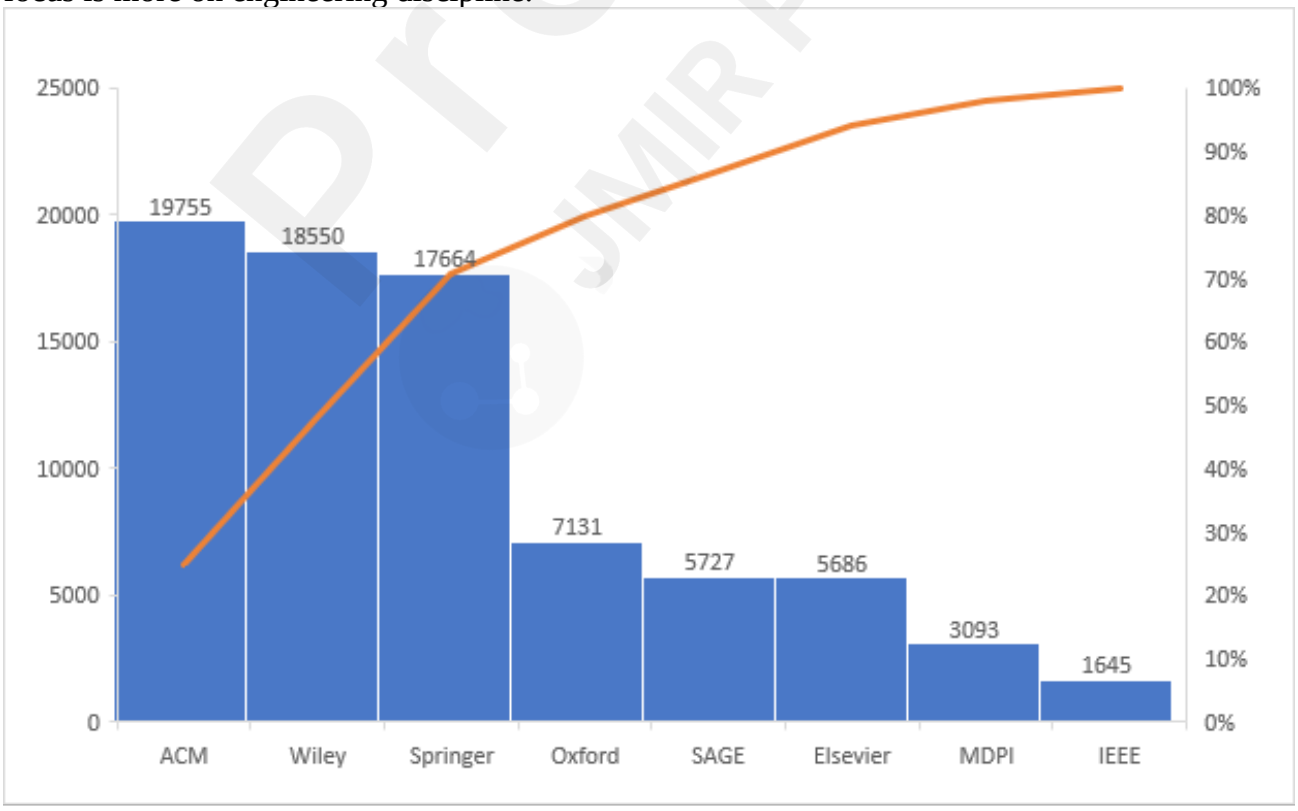

Figure 3: Showing the articles that appeared in major scientific libraries accessible to a researcher in decreasing order between

January 2020 ~ December 2020.

Preprint literature

Alongside published research literature, many research articles are submitted to preprint repositories that are unreviewed. In subsequent times they may be published in some referred journals. Due to unpredicted review time, most of the researchers first submit their earlier version of the paper to a preprint database to get an earlier response on their research outcomes. Among the popular preprint, arrive [7] is an extensive multidisciplinary archive for about 1.7 million scholarly articles in physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, electrical engineering, and systems science economics. Regarding COVID-19, it published about 3000 articles in 2020 as illustrated in Figure 4. Related to the medical and biology domain, two archives, midrib and bioRxiv [8], are at the forefront to collect unreviewed literature. Currently, altogether about 12000 COVID-19 preprints are available for download from these two archives. Other archives like chemRxiv [9]- a service for preprints in chemistry and related areas-also publish a considerable number of pre-reviewed articles on COVID-19.

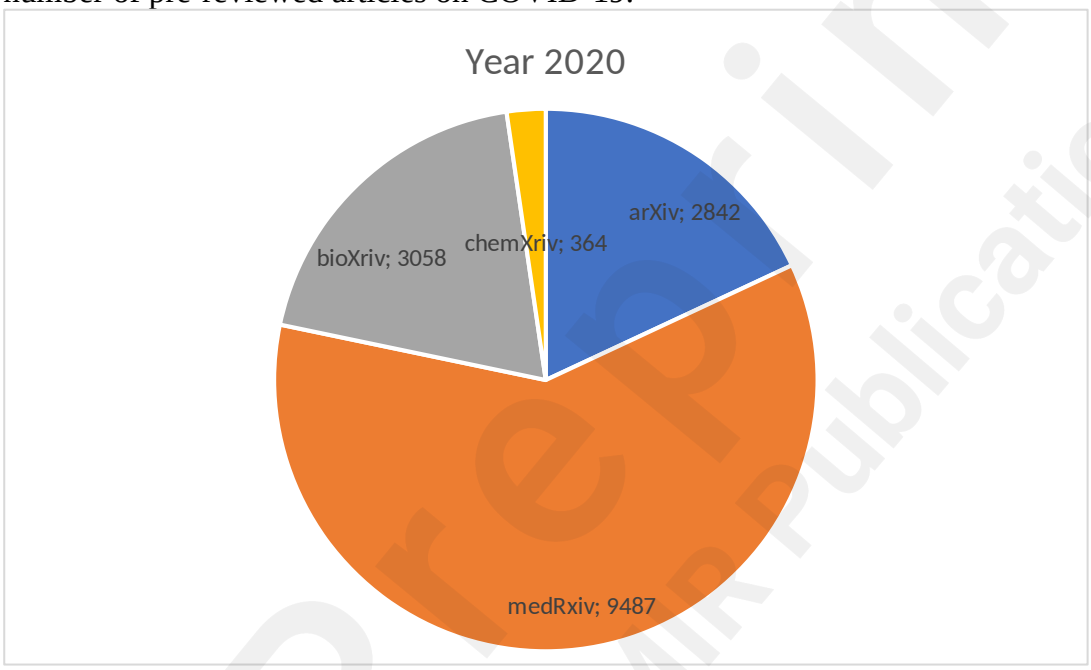

Figure 4: Submission of research articles about COVID-19 on preprint servers.

Patient Dataset

Data plays a crucial role in the ability to research, study, and explore population health and safety, and now this is truer in the case of a global pandemic. Access to datasets and associated tools that can examine that data is increasingly crucial to the research process and is particularly necessary for the global response to the novel Coronavirus.

To aid researchers, developers, and analysts in the struggle to fight COVID-19, different platforms offer datasets about COVID-19 patients like Johns Hopkins Center for Systems Science and Engineering (JHU CSSE) [9], the Google COVID-19 Public Dataset Program [10], the Kaggle [11], and the GitHub [12]. Only on Kaggle, there are 972 datasets of small (773), medium (168), and large (24) sizes. Most of these datasets are available in CSV (comma-separated value) or MS Excel formats. On GitHub, which is primarily a repository for hosting project code files developed in different languages, e.g., such as Python, data is also stored for the code reproducibility alongside the code files. Until present (June 10, 2020), 4073 public repositories are added to GitHub, matching the topic of COVID-19. Several other resources include data.world [13], AMiner [14], and IEEEDataPort [15] provide free datasets related to COVID-19. Hundreds of regional and country-

specific datasets are available through different channels; however, this study does not accommodate to describe them here. This study rather aims at presenting platforms that host datasets across the regions/countries.

For public health safety and security, a few global level information resources such as WHO (World Health Organization) [16] and CDC (Centers for Disease Control and Prevention [17] are worthful to mention as they provide an important piece of information for public awareness. WHO is a global entity that accumulates information from the member organization and disseminates it on its website and other public interest channels and awareness. The WHO developed a specialized tool so-called the Coronavirus Disease (COVID-19) Dashboard [9], which provides a rich spectrum of meta-information and analytics about COVID-19. The CDC fights diseases whether they start in the USA or abroad, are chronic or acute, curable or preventable, human error or deliberate attack. It provides a rich source of information on COVID-19 about the symptoms, risk factors, guidelines on social distancing, and important information to different stakeholders like policymakers, travelers, businessmen, schools, health professionals, the general public, and high-risk populations. In Table 2, we provide summarized information about the informative resources about COVID-19.

To facilitate users to understand the nature of the data, we describe popular dataset platforms in Table 1, reporting characteristics that are useful for developing statistical and machine learning models for further analysis and research. In addition to describing the platform's aim and scope, we also provide a brief explanation of the nature of the datasets hosted on that platform to help users spend their time appropriately. Some platforms provide data that can be used solely for statistical analysis and reporting purposes, while others could be used for assistance in clinical decision-making of patient diagnosis, treatment, and prognosis.

Table 1: Brief description of COVID-19 Platforms that offer an open repository of datasets and their scope and usefulness

Dataset Platform	Brief Description	COVID-19 Datasets/ repositories
Kaggle [11]	Kaggle is the world's largest data science community with powerful tools and resources to achieve data science goals.	972 datasets
GitHub [12]	GitHub is primarily a host for software development and version control; however, alongside the code files, the associated data is also available for code reproducibility	4073 repositories
data.world [13]	Data.world is an open data resource hub about COVID-19 contributed by thousands of users and organizations worldwide.	40 datasets
AMiner [14]	AMiner collects all kinds of open datasets about COVID-19 and keeps updating every day. The data is open and available for download.	438 datasets
IEEEDataPort [15]	IEEE DataPort is free dataset storage and hosts different types of datasets. It provides space to host datasets related to COVID-19. It hosts a large set of Tweets data containing 174,573,543 global Tweets in the English language.	15 datasets
Google COVID-19 Public Dataset Program [10]	To make data more accessible to researchers, data scientists, and analysts, Google created a COVID-19 Public Datasets program that hosts a repository of public datasets -- free to access and analyze.	Three datasets (JHU CSSE dataset, Global Health Data from the World Bank, and OpenStreetMap data.)

Table 2: Brief description of COVID-19 related informative resources

Organization	Brief Description	Key Information Services
WHO [18]	WHO works with 194 Member States across six regions and from more than 150 offices, aimed striving to combat diseases – communicable diseases like influenza and HIV, and non-communicable diseases like cancer and heart disease.	✓ **COVID-19 Dashboard** that provides up-to-date case information, including the number of deaths and recovery. ✓ **Generate advice** for public awareness to reduce the chances of being infected or spreading COVID-19. ✓ **Situation Reports** are released everyday that provide the current COVID-19 epidemiological situation and presents official case and death counts and transmission classifications. Until Jun-16-20, 148 situation reports have been released. ✓ Other services include Travel advice, training and exercise, technical guidance, response funds, and others**.**
CDC [17]	CDC is America based organization aimed at providing health safety and security threats, both foreign and domestic.	✓ **Symptoms** information ✓ **Risk factors** information ✓ **Social distancing** information ✓ A **bank of answers** to important questions and many other guidance materials for various stakeholders includes travelers, healthcare professionals, and others.
EU Open Data Portal [19]	European Centre for Disease Prevention and Control publishes a dataset contains the latest available public data on COVID-19 worldwide through screening up to 500 relevant sources every day between 6.00 and 10.00 CET.	✓ **COVID-19 cases worldwide** for download ✓ **Visualizations** of cases geographically distributed, situation dashboard, and other graphical representations**.** ✓ **Documentation** as to how the data is collected, a script of R software, and webinars.
JHU COVID-19 Dashboard [20]	Johns Hopkins experts designed a rich and interactive dashboard as an international response to COVID-19 to inform the public and help policymakers creating awareness and save lives.	✓ **An interactive dashboard** for tracking global COVID-19 cases. ✓ **Animated maps** that show total cases, deaths, and new cases. ✓ **Critical Trends** on how the novel Coronavirus is spreading around the globe ✓ Worldwide **Mortality Analysis**
The World Bank-COVID-19 case data [21]	The World Bank provides an array of real-time data, statistical indicators, and other types of data relevant to COVID-19, particularly on the economic and social impacts of the pandemic and the World Bank's efforts to address them.	✓ **Global poverty** estimates of the impact of COVID-19. ✓ **Health Nutrition and Population Statistics** ✓ Understanding the COVID-19 pandemic through data through **indicators and worldwide cases** ✓ **Map** of World Bank's **Operational Response** to the Coronavirus and relevant services.
DXY [22]	DXY provides timely, accurate, and authoritative COVID-19 global pandemic real-time reports through global mapping and knowledge.	✓ **Global Mapping** of corona cases ✓ **COVID-19 Knowledge** for all, public, and doctors.
NIH Coronavirus (COVID-19) [23]	National Institute of Health offers a specialized service that provides the latest research information on COVID-19.	✓ **ACTIV** (Accelerating COVID-19 Therapeutic Interventions and Vaccines) ✓ **Treatment Guidelines** ✓ **Grants and Funding** information ✓ **COVID-19 Testing** information

Preprint JMIR Publications

Clinical Trials

Clinical trials are conducted to evaluate medications or medical devices' effectiveness and safety by monitoring their effects on a selected population. Clinical trials pass through two stages: registration and publishing. Preferably, every trial should have at least one result article; even if the results are not significant or with negative findings; however, sometimes it may be harder to publish due to publication bias [24]. A trial can be linked to a journal article through an unstructured trial-article link (may not involve unique identifiers) or a structured trial-article link (a computable link assigned with unique identifiers such as the ClinicalTrials.gov ID or the PubMed ID) [25].

Several platforms are functional such as those that register primary clinical trials related to COVID-19 and are freely accessible and searchable. The WHO ICTRP (International Clinical Trials Registry Platform) [26] was established to facilitate a network of international clinical trial registers to ensure a single point of access and the unambiguous identification of trials have public accessibility. From Jan-20 to Jun-20, a total of 3163 clinical trials have been registered in the WHO ICTRP database from 18 different sources that include ClinicalTrials.gov, chiCTR (Chinese Clinical Trials Registry), ANZCTR (Australia New Zealand Clinical Trial Registry), CRIS (**Clinical Research Information Service**, Republic of Korea), CTRI (The Clinical Trial Registry- India), EU Clinical Trials Register, German Clinical Trials Register, IRCT (Iranian Registry of Clinical Trials), ISRCTN (International Standard Randomised Controlled Trial Number), JPRN (Japan Primary Registries Network), LBCTR (Lebanese Clinical Trial Registry), Netherlands Trial Register, PACTR (Pan African Clinical Trials Registry), REBEC (**Registro Brasileiro de Ensaios Clinicos**- Brazilian Clinical Trials Registry), REPEC (Peruvian Clinical Trial Registry), RPCEC (Cuban Public Registry of Clinical Trials), SLCTR (Sri Lanka Clinical Trials Registry), TCTR (Thai Clinical Trials Registry).

The US-based ClinicalTrials.gov [27] is perhaps the major database of privately and publicly funded clinical studies conducted worldwide. In the ClinicalTrials.gov database, a total of 2172 Studies were found for COVID-19 on June 18, 2020. The second-largest clinical trials database is the chiCTR [28] – a non-profit Chines platform that has registered 721 clinical trials. All the registries are region or country-specific except the ISRCTN, which was initially stood for 'International Standard Randomised Controlled Trial Number'; however, the scope has widened to include other study types to evaluate the efficacy of human-related health interventions [29]. As of June 18, 2020), the ISRCTN database registered 70 clinical trials, out of which only six are completed, four suspended, and 60 are in progress.

Figure 5 shows a progress bar of trials registered for the first six months of 2020. We can observe the highest number of trials (N=967) have been registered in April. The number starts dropping after April is generally applied to all databases except the chiCTR, where we can see the decline start after February. This may be because, in China, the peak number of COVID-19 cases was observed in February. The number of trials in May and June should higher than in April as the number of positive cases is still on the rise globally.

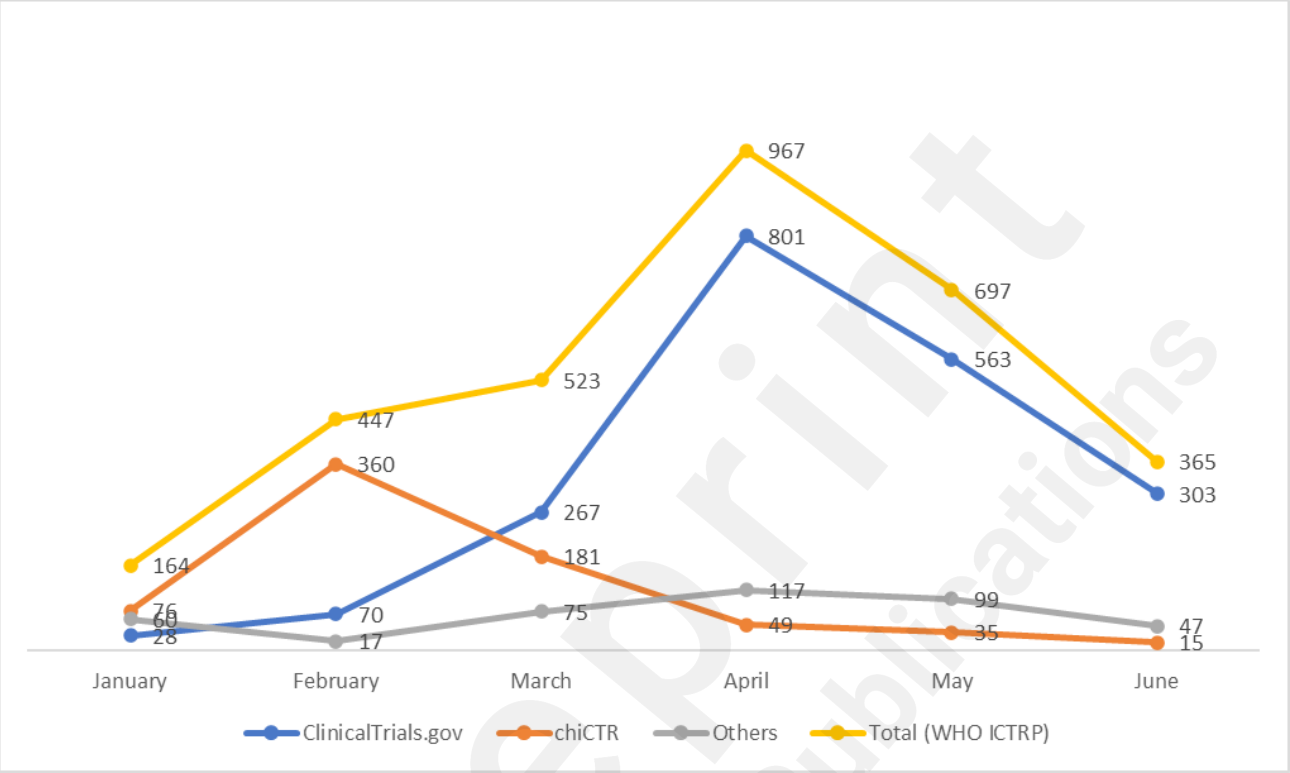

Figure 6 provides a statistical summary of ClinicalTrials.gov that includes the number of trials in 2020. The data shows that only 572 trials have been completed, out of which only 19 have been reported with results. A Phase IV is also called a post-marketing surveillance trial, occurs after FDA has approved a drug for marketing, and is the only trial that is eligible to watching drug use in public. At current, the number of Phase IV trials (N=98) is far less than the Phase III trials (N=432) and other trials in other phases; however, it is a reasonable number to raise the hope of getting significant results from watching the drug effects on a large number of COVID-19 patients around the world.

Preprint JMIR Publications

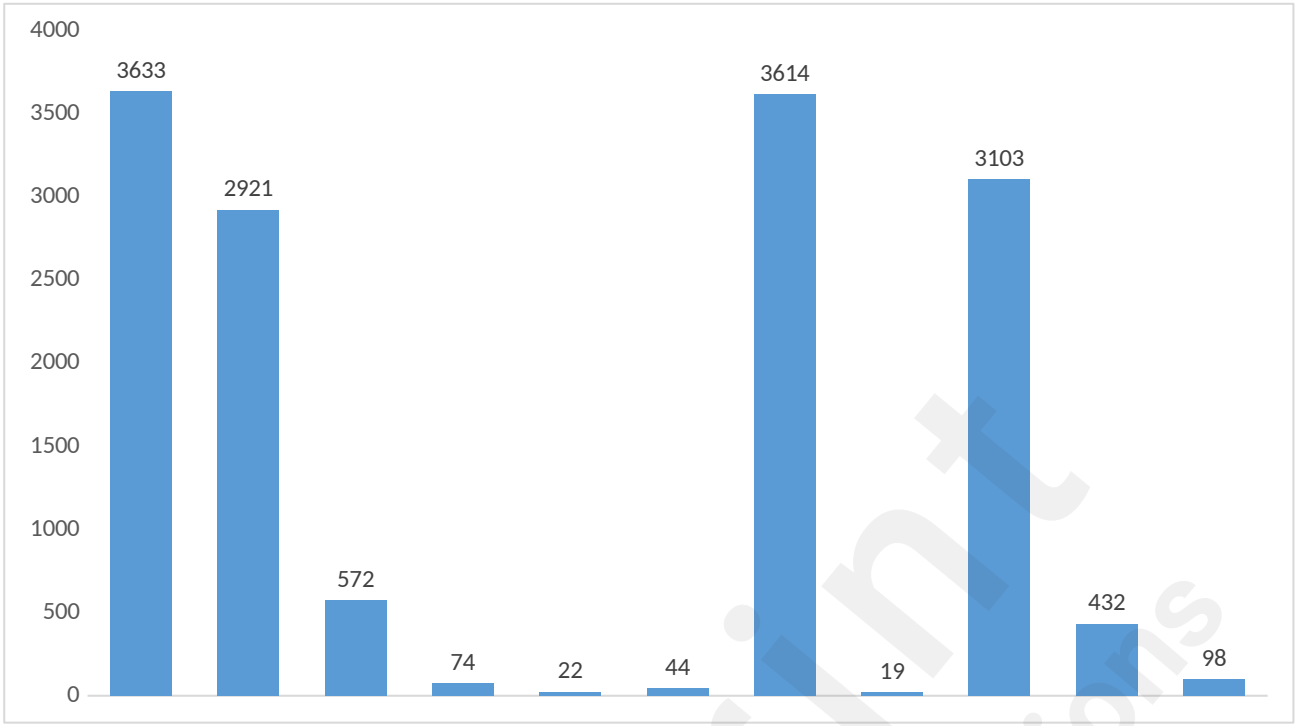

Figure 6: Statistical summary of clinical trials registered in the ClinicalTrials.gov database accounting for the trials' status either they are [completed, withdrawn, suspended, terminated or in progress], trials with results, and trials registered in different Phases.

Secondary Resources

We present global scale initiatives founded on the data and information acquired from the primary level sources at the secondary level. These initiatives are discussed without any specific order. Most of these initiatives are pertinent to creating metadata over the research literature articles for subsequent analysis and research.

PrePrint JMIR Publications

CORD-19: The COVID-19 Open Research Dataset

CORD-19 is a free, open research data resource consisting of 130,000 scholarly articles about the novel Coronavirus available for the global research community [30]. The CORD-19 resource is updated every week with newly published research to facilitate the development of text mining and information retrieval systems, and it has been downloaded over 75,000 times in the first month of its release [31]. The articles in CORD-19 are derived from four primary level repositories that include PubMed Central (PMC) [32], bioRxiv and medRxiv preprint servers, and the World Health Organization (WHO) Covid-19 Database. The significant accomplishment of CORD-19 is the cleaning of metadata and machine readability of full text. A simple deduplication logic of creating clusters for retaining similar articles unless there is a conflict is applied while one paper metadata from each source is cleaned and formatted into CORD-19. After cleaning, they parse content from PDFs formatted papers into a JSON schema which is simple to utilize for different text mining tasks. One of the CORD-19 dataset features is the article's source representing the name of the publishing source. Overall, seven unique sources are enlisted; however, multiple sources are mentioned for articles published in more than one source. The seven individual sources hold articles either unique to them or share with other sources. In Figure 7, the articles shared among different resources are visualized using a Battle Venn diagram where the common area represents the number of shared articles. The intensity of color shows the high number of sources that share those articles. For instance, Medline has 84399 articles that are not published anywhere else. However, it shares several articles with other sources; for example, with PMC, it shares 62808 articles. No article is common in all the sources; however, four articles shared among the highest number of sources means all the sources except Elsevier and ArXiv.

Preprint JMIR Publications

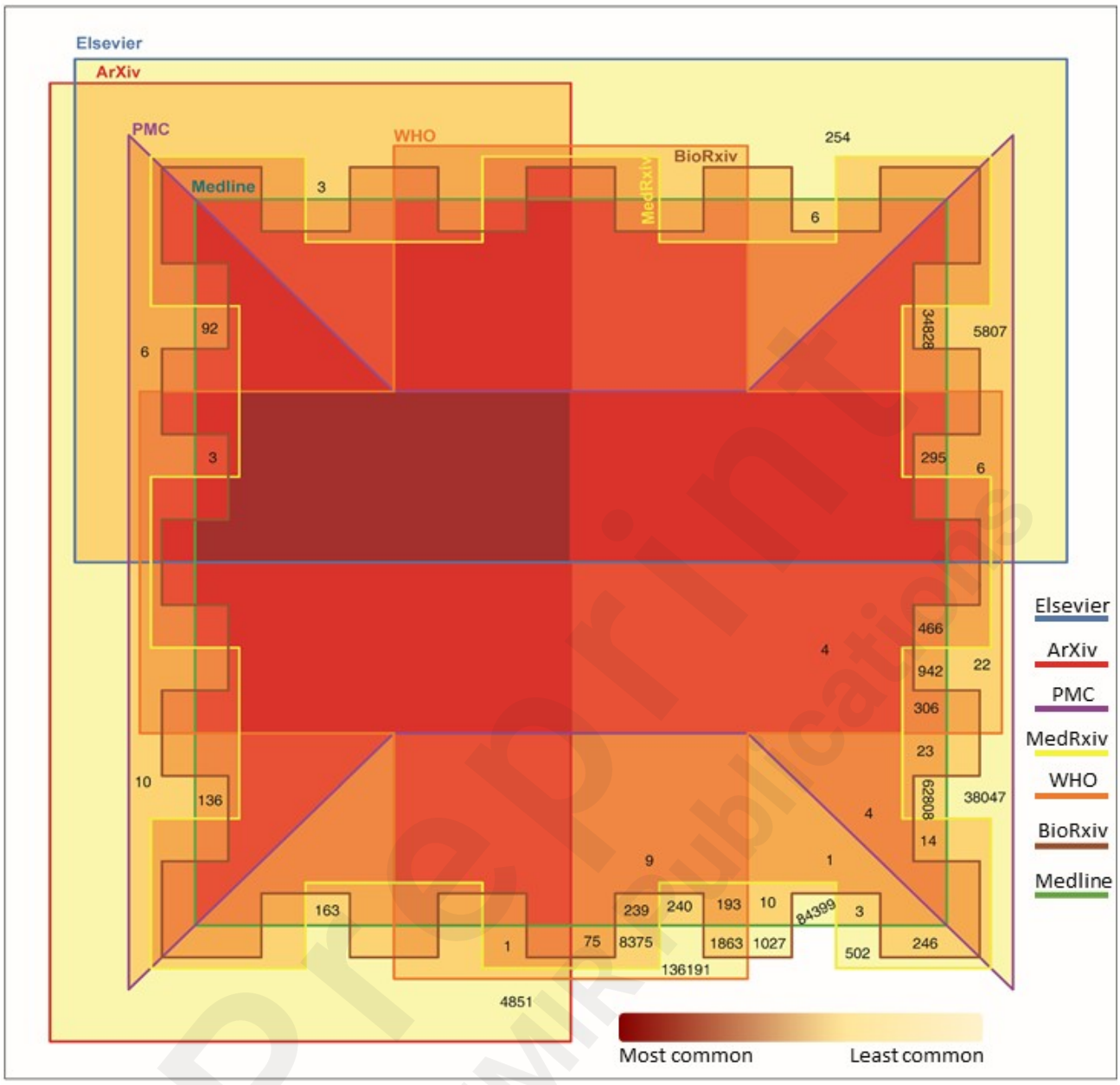

Figure 7: Number of articles published unique to each resource with no intersection, i.e., these articles are not published in more than one resource.

LitCovid

LitCovid is a hub of curated literature of scientific articles about COVID-19 [33–35]. The source is updated daily and has access to about 29000 (still growing) articles in PubMed. One unique feature of this effort is identifying relevant articles that are 35% better than the conventional keyword-based searches [35]. For improved accessibility, additional information is added, such as the articles are categorized by research topics that include general information, mechanism, treatment, case report, transmission, disease diagnosis, prevention and epidemic forecasting, and geographical locations [35]. Under each topic, three more important pieces of information are indexed – chemicals (the name of chemical products like Remdesivir used in different trials), journals, and countries (the host country of research).

An associated effort extended by PubTator Central [36] derived articles from LitCovid and annotate them with six entity types, also called bio concepts – Gene, Disease, Chemical, Mutation, Species, and CellLine. The annotations of these six entity types are made in different colors (Figure 8), such

as if a disease concept appears in title or abstract, it is highlighted in yellow color. Similarly, chemical names appear in green color, a gene in purple. These annotations support text and data mining activities are available for download in XML format.

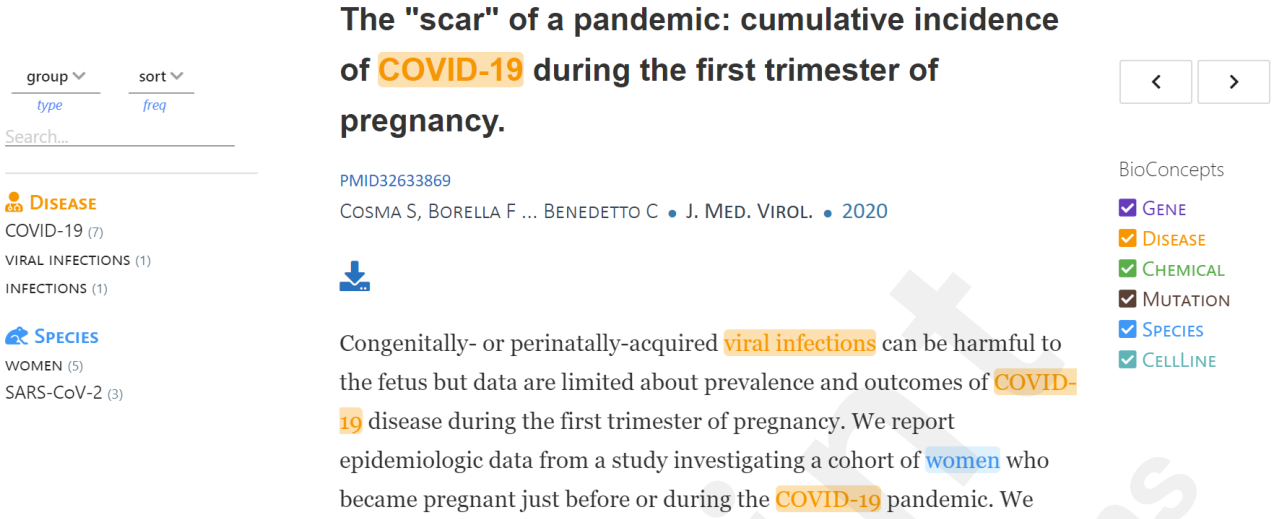

Figure 8: Annotated entity types highlighted in different colors in the title and abstract.

COVID-19 Evidence Alerts

COVID-19 Evidence Alerts is a McMaster University service that alerts users to current best evidence about COVID-19 [37]. It notifies users about the reports published in MEDLINE-based journals, which are critically appraised on scientific merit. Each appraised study report is assigned a study category that includes diagnosis, etiology, treatment, prognosis, guideline, and clinical prediction.

i. The study reports with sufficient scientific merit are appraised into three categories. “higher-quality studies for clinical attention.”
ii. The study reports that belong to the study category but provide lower-quality scientific evidence are posted with at least one reason for their lower quality, e.g., “not a randomized controlled trial.”
iii. A third category, called “studies that do not belong to a study category we appraise,” as its name suggests, includes those study reports that are relevant but may be assigned to a specified study category.

This service is beneficial for applications that involve clinical decisions. As of June 30, 2020, about 2700 study reports are posted accumulatively in all three categories, where the list is growing and updated every weekday. Out of all the appraised study reports, about 70 reports belong to higher quality, which establishes that less than 3% of reports meet the scientific criteria and clinical relevance. As shown in Figure 9, the study reports that do not belong to a study category are double compared to those study reports belonging to at least one of the study categories.

Preprint JMIR Publications

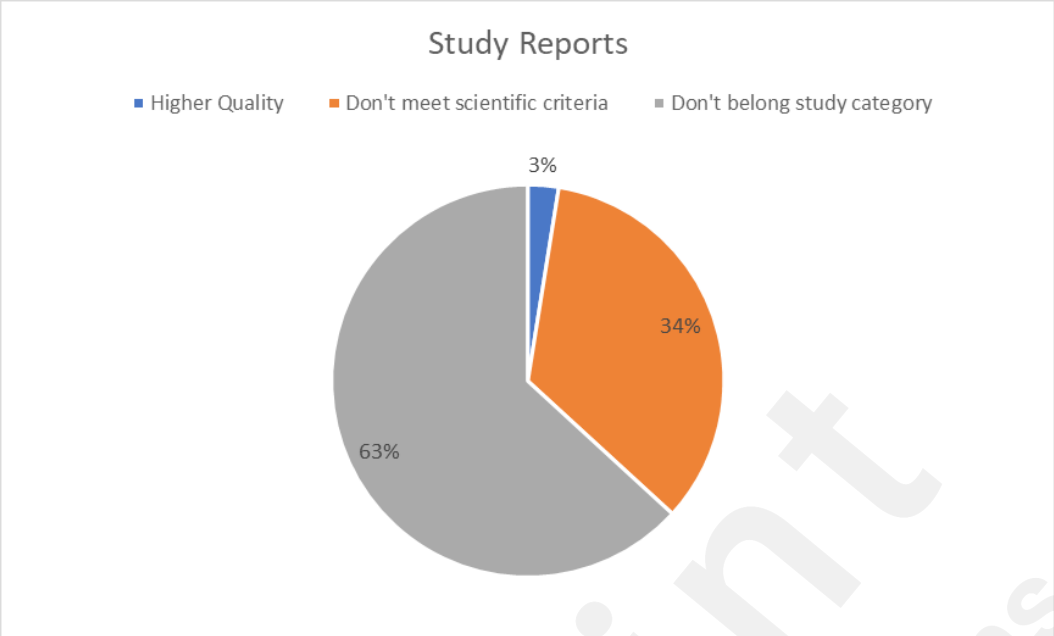

Figure 9: Study reports categorization on scientific merit and clinical relevance.

CALC-19

COVID-19 Advanced Literature Classifier (CALC-19) is a classifier of medical literature about COVID-19. The data is updated every week; as of July, the dataset includes more than 150000 scholarly articles where each article is tagged by country, year, source, topic, and keywords [38]. CALC-19 includes all its articles from the CORD-19 dataset and adds metadata like mesh keywords, author, date of publication, and journal of publication in addition to title and abstract. Users can search the articles of interest using different filters and tags provided by the service. The searched articles can be downloaded in Excel, CSV, Endnote, RIS, BibTex, and a new HL7 FHIR JSON format. The CALC-19 service is powered by the PICO Portal that provides a platform to accelerate research and innovation by leveraging AI solutions and creating efficient systematic reviews of studies.

Evidence Aid COVID-19

Evidence Aid [39] is a collection of evidence in the form of summaries created over high-quality research studies. The collection is available in English with translation in five languages: French, Spanish, Portuguese, Arabic, and Chinese. This collection holds summaries of systematic reviews intending to provide a quick overview of relevant and impactful information about COVID-19 concerning health conditions, outcomes, and other aspects necessary for the recovery period.
Each summary of a systematic review consists of subsections such as “citation,” “what is this?” and “what was found.” The citation section includes information about the systematic review title, authors, publishing venue, date, and page information. The “what is this?” section summarizes the information about population, experimental setup, environment, location, and others. The “what was found” section describes the findings and outcomes of the trials included in the systematic review, such as the impact of a drug on COVID-19 patients as positive, negative, or no effect.

Comparative summary of initiatives

The initiatives discussed as a part of the secondary level resources are summarized in Table 3 to provide a comparative analysis of different aspects, such as dependent resource information, applications, language support, human validations, and download formats.

Table 3: Comparative summary of secondary level resources based on different features important for researchers to expand their research on COVID-19.

Initiative\Feature	Baseline Resource	Direct Applications	Update Frequency	Multilingual Support	Human Expert Verified	Export Format
CORD-19	Original Research	IR & TM	Weekly	NA	NA	CSV
LitCovid	PubMed	IR & TM	Every day	NA	Yes	RIS and TSV.
COVID-19 Evidence Alerts	(CORD-19)	SR	Every weekday	NA	Yes	NA
CALC-19	(CORD-19)	IR & TM	Weekly	NA	Users can filter records and tags to find records relevant to their work.	Excel, CSV, Endnote, RIS, BibTex & HL7 FHIR JSON.
Evidence Aid COVID-19	Systematic Review	EBM, CDS, Guidelines	NA	Yes	Yes	NA

Tertiary Sources

The tertiary level initiatives use the outcomes of secondary resources for creating clinical guidelines, standards, and vocabularies to provide direct assistance to the medical professionals and the implementers of clinical decision support systems. WHO and CDC are the two comprehensive resources that provide general and all technical guidelines to the public and health professionals. On top of this, other resources add, refine, and customize their findings. WHO publishes technical guidelines in 14 different categories: clinical care, infection prevention and control, laboratory and diagnosis, and others [40]. CDC offers guidelines for different stakeholders; for health workers alone, it guides 12 categories: testing, clinical care, infection control, and others [41]. Numerous initiatives derive guidance from WHO and CDS guidelines, such as Duodecim: EBM Guidelines Coronavirus infections [42] and COVID-19 Guidelines listed in DynaMed [43].

COVID-19 Knowledge Accelerator

COVID-19 Knowledge Accelerator (COKA) [44] is first initiated at the end week of March 2020. Led by Brian Alper, COKA is an add-on effort to an ongoing project, Evidence-Based Medicine on FHIR (EBMonFHIR) [45], to accelerate processing massive research data of COVID-19 to summarize and synthesize the evidence in a standard format for computable expression. It aims at resolving inefficiencies in current scientific dissemination systems in which research data is transformed into various non-computable forms for human displays [46]. This initiative rightly identified the problem area of non-computable communication and channelized its efforts to construct computable (structured) results directly from research publications, thus accelerated evidence synthesis.

The initiative gained momentum over a short period of time, and as of July 2020, COKA had more than 150 working meetings with more than 40 active participants from more than 25 organizations from academia, industry, government, and non-profits in 7 countries. Activities under this initiative are divided into three-team setups: project, process, system. Under each setup, four workgroup meetings are held every week (accumulatively twelve meetings/week). The participants actively participate and contribute to different meetings of their choice every week, and a report of discussion is shared at the end of the day. Table 4 describes the structure of weekly workgroup meetings

explains the purpose and team setup.

Table 4: Team set and working group meetings with weekly frequency

Team Set	Name of Work Group	Frequency (weekly)
Project	Project Coordination	2
	Recommendations/CDS Liaison	1
	Communications (Awareness, Scholarly Publications)	1
Process	Summarization Process Development	1
	Classifying Content Process Development	1
	Statistics Process Development	1
	Quality/Certainty Rating Process Development	1
System	Standards Development (FHIR, Terminologies)	2
	Evidence Evaluation and Reporting Tools Development	1
	Content Citation and Classification Tools Development	1

One of the key achievements of this initiative is the development of Citation resource schema and instances. As of July, COKA has created more than 36,000 citation resources for the biomedical publications in the CORD dataset. Moreover, a profile resource, "EvidenceReport," is another important outcome of this initiative, an extension of the "Composition" resource. The EvidenceReport resource provides a comprehensive report refer to one or more than one resource. As of July, more than 30 example reports related to COVID-10 have been generated. In addition to resource schemas, COKA had tremendous efforts in vocabulary mappings for evidence-related resources. In a short time, COKA developed a 13-step Code System Development Protocol in September 2020. HL7 FHIR is used as the underlying standard to meet the interoperability needs and support the global development of terminologies for the exchange of scientific evidence [47].

ACTS COVID-19 Guidance-to-Action Collaborative

AHRQ evidence-based Care Transformation Support (ACTS) initiative of COVID-19 Guidance-to-Action Collaborative aims to improve the development, dissemination, and use of "living" COVID-19 guidance [48]. The Collaborative supports the COVID-19 'knowledge supply chain,' i.e., data-to evidence-to knowledge-to guidance-to action to make the processes of guidance development, workflow integration, and knowledge supply chain more efficient and effective. Among its primary functions, the Collaborative provides current answers to urgent clinical questions faced by health professionals, helps guidance developers in tracking COVID-19-related recommendations, fosters collaborations among implementers of COVID-19 guidance summaries, and facilitates coordination to optimize the flow of COVID-19 "evidence-to-action."

The collaborative produces rapid guidance summaries that provide a comprehensive description of existing evidence and guidance from various sources such as the CDC, WHO, and Europe CDC (ECDC). The guidance summaries are not clinical practice guidelines (CPGs); therefore, they should not be used or interpreted as CPGs; they can help develop local recommendations and policies. As of July, 20 guidance summaries have been produced in two categories: patient care and operations. The guidance summary is structured in a question form put at the top with answers to major recommendations followed by a list of evidence collected from various sources.

NIH COVID-19 Treatment Guidelines

In collaboration with other organizations, the National Institutes of Health (NIH) have developed treatment guidelines to support clinicians in COVID-19 patients' care [49]. These guidelines are updated frequently due to the quickly evolving clinical information about the new Coronavirus. The guidelines' recommendations, which are based on scientific evidence and expert opinion, possess two ratings: the strength of the recommendations indicated in a letter (A, B, or C), and the quality of the evidence indicated in a Roman numeral (I, II, or III). A panel composed of experienced

representatives from 14 different organizations and societies such as the American College of Chest and Emergency Physicians, Food and Drug Administration, and Society of Critical Care Medicine is established to develop these guidelines. The panel utilizes data from the published scientific literature on COVID-19 and their experience to develop the recommendations in these guidelines.
The panel develops the recommendations majorly in clinical care areas, i.e., care of critically ill patients with COVID-19, including antiviral therapy, immune-based therapy, and adjuvant therapy guidelines for special populations such as pregnancy and children. The panel's approach to publishing the recommendation in these guidelines can be learned from the example recommendation provided about Chloroquine or Hydroxychloroquine in the category of antiviral therapy.
*"The Panel recommends against the use of high-dose chloroquine (600 mg twice daily for ten days) for the treatment of COVID-19 (**AI**)."*
This example tells us that the strength of the recommendation is **A**, which means it is a strong recommendation for the statement, the quality of evidence is **I**, which means it is supported with data of one or more randomized trials with clinical outcomes and validated laboratory endpoints.

ACS Elective Case Triage Guidelines for Surgical Care

American College of Surgeons (ACS) has developed recommendations for surgeons to identify which procedure should be curtailed [50]. The ACS releases twice-weekly newsletters to update recommendations on curtailing the performance of surgical procedures continuously. As of July, the ACS has published guidelines in 14 categories: cancer surgery, gynecology, neurosurgery, urology, vascular surgery, and others. Most of the guidelines are provided in a descriptive form, but a few include vascular surgery and orthopedic procedures. An example of COVID-19 Guidelines for Triage of Vascular Surgery Patients is provided in Table 5, describing each tier class's meaning (1, 2a, 2b, and 3) concerning surgery postponement.

Table 5: COVID-19 Guidelines for Triage of Vascular Surgery Patients.

Category	Condition	Tier Class
AAA	Ruptured or symptomatic TAAA or AAA …	3 Do not postpone …
Aneurysm peripheral	Peripheral aneurysm, Asymptomatic …	2a Consider postponing …
Bypass graft complications	Revascularization for high-grade re-stenosis of previous intervention …	2b Postpone if possible …
Carotid	Asymptomatic carotid artery stenosis …	1 Postpone …
…	…	…

COVID-19 – SNOMED CT

SNOMED International -- a leading healthcare terminology organization, took steps to identify codes for different terms related to COVID-19 [51]. In the March 2020 interim release, 24 records have been added, which is increased to 49 records as of August 11, 2020. A complete list of these concepts with URI, fully specified name and preferred term is now available on confluence [52]. The two most important terms, SARS-CoV-2 (organism) and COVID-19 (disorder), are given identifiers 840533007 and 840539006, respectively. A map of SNOMED CT to ICD-10 is provided for two concepts: COVID-19 and Exposure to SARS-CoV-2, as described in Table 6.

Table 6: SNOMED-CT to ICD-10 Map

Preferred Term	Source SNOMED CT Identifier	Target ICD-10 Identifier

COVID-19	840539006	U07.1
Exposure to SARS-CoV-2	840546002	Z20.8

Summary of tertiary resources

This section provides an inexhaustive list of initiatives that contributes to creating tertiary resources. As shown in Table 7, the initiatives are categories into three categories: standard, guidelines, and terminologies. The supporting group or organization information is provided to support future applications cataloging these resources in their services.

Table 7: Summary of tertiary resources related to COVID-19 based standards, guidelines, and terminologies.

Category	Organization/ Group	Initiatives
Standards	HL7	Health Level Seven (HL7) is a health information exchange standard. Using its current standard—Fast Health Interoperability Resource (FHIR), different initiatives have been taken place related to COVID-19 that include COKA [44] and SANER Project Using HL7 FHIR to Enable Easier Reporting for Public Health [53].
	ISO	International Organization for Standardization (ISO) has compiled a list of freely available standards to support global efforts in dealing with the COVID-19 crisis [54]. It also featured a list of national resources developed by ISO members in different countries to support the fight against COVID-19 [55].
	ANSI	In response to the COVID-19 pandemic, the American National Standards Institute (ANSI) has initiated the Standards Alliance Phase 2 (SA2) [56] that aims to reduce importation and regulatory barriers of COVID-19 testing kits and train resources on the use of the medical devices and testing equipment.
Guidelines and Recommendations	NIH	National Institute of Health (NIH) has developed a comprehensive set of COVID-19 Treatment Guidelines [49]
	ACS	American College of Surgeons (ACS) Elective Case Triage Guidelines for Surgical Care [50] aims to develop surgeons' recommendations to identify which procedure should be curtailed.
	WHO	WHO is perhaps the largest source of information, guidance, and recommendations to support the fight against COVID-19. In addition to general public health recommendations, it provides technical guidance on different topics [40].
	CDC	Like the WHO, the CDC offers advice, recommendations, and guidelines for different stakeholders. For health workers alone, it guides 12 categories that include testing, clinical care, infection control, and others [38].
	AGA	The American Gastroenterological Association (AGA) has developed recommendations based on systematic review and meta-analysis of 47 studies and 10,890 unique patients with gastrointestinal symptoms [57].
	FDA	The U.S. Food and Drug Administration (FDA) provides guidance related to drug development programs and the food industry impacted by COVID-19, such as donating COVID-19 Plasma and facilitating the development and availability of medical devices therapeutics to combat COVID-19 [58].
Vocabularies and Terminologies	SNOMED CT	As of August 11, 2020, SNOMED International has added 49 records related to COVID-19 [51].
	ICD-10	The International Classification of Diseases, 10th Revision (ICD-10) provides critical information needed to understand the usage of ICD-10-CM in an official coding and reporting guidelines report April 1, 2020 through September 30, 2020 [59]. These codes can be

		searched in their newly released user-friendly browser [60].
	UMLS	The Unified Modeling Language System (UMLS) provides a set of COVID-19 related terms [61] mostly mapped to SNOMED CT and MeSH.
	MeSH	A new Medical Subject Headings (MeSH) Supplementary Concept Record (SCR Class 3-Disease) was added on February 13, 2020, to the 2020 MeSH Browser in response to COVID-19 [62]. The most current updates are found on the MeSH Browser [63]. Using MeSH terms, the recommended search strategy for retrieving COVID-19 related biomedical studies is *"2019-nCoV OR 2019nCoV OR COVID-19 OR SARS-CoV-2 OR ((Wuhan AND coronavirus) AND 2019/12[PDAT]:2030[PDAT])"*.
	LOINC	Logical Observation Identifiers Names and Codes (LOINC) introduced codes 264 terms in response to COVID-19 caused by novel coronavirus [64]. These codes are distributed in five categories: SARS-CoV-2 Lab Tests (84 terms), LOINC terms for SARS-CoV-2 AOE questions (9 terms), Convalescent plasma (2 terms), LOINC terms related to public health case reporting (63 terms), and COVID-19/Telehealth Documents (106 terms).

COVID-19 dashboard applications

An array of dashboards is proposed and implemented to portray the influx of data and information related to COVID-19. Here we first discuss a set of well-known dashboards in service and provide information in different forms that include structured, unstructured, plane, and graphical.

John Hopkins interactive COVID-19 dashboard

In response to the COVID-19 emergency, the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU), USA, developed an online interactive dashboard to visualize and track Corona cases in real-time [20,52]. Starting on January 22, 2020, the dashboard data was updated manually, but with the increased number of cases, the manual reporting process became unsustainable; therefore, a semi-automated living data stream strategy was started to follow on February 1. This dashboard relies on several data sources, such as for the new case identification, it seeks data from DXY China (initially it was the only data source), Twitter feeds, online news and direct communication sent to the dashboard. The case numbers are duly confirmed with regional and local health departments, including the respective CDCs, health departments, and WHO. This dashboard's main feature is the real-time interactivity, enabling users to see the cumulative cases, active cases, and other information like incidence rate on the global map. The users can click and select their chosen region on the map and see the region's statistics. A U.S. map is separately provided with county-wise confirmed cases and deaths. Moreover, it supports data in motion that show daily trends of cases and deaths in different regions.

WHO COVID-19 dashboard

Like the CSSE, the JHU interactive dashboard, the WHO provides a live dashboard providing COVID-19 updated case numbers and deaths on the map with mouse hovered feature that saves an extra click for the users to have an overview of the cases in a different part of the world [16]. It provides two kinds of maps: a bubble map and a choropleth map. Alongside the maps, it provides data tables structured on country-wise cumulative confirmed cases, deaths, and newly reported cases and deaths in the last 24 hours. It facilitates searches region-wise such as Europe, Asia, and Africa, as well as individual country-wise.

Worldometer Coronavirus updates

Run by an international team of developers, researchers, and volunteers, Worldometer provides global COVID-19 real-time statistics on data collected, analyzed, and validated from thousands of sources worldwide [65,66]. The data is claimed to be trusted and used by different governments and private organizations such as the UK, Thailand, Pakistan, Sri Lanka, Vietnam, and organizations like Johns Hopkins CSSE, the BBC, the New York Times, and many others. It provides country-wise statistics of new cases, deaths, recovered cases, critical cases, tests per million, and other important information. The most informative feature is the searchable and clickable country-wise data table that can be customized to include/exclude columns such as new cases, total cases, deaths, tests per million along with the total population of that country.

Other than the dashboards mentioned above, dozens of dashboards are available for use scoped by a specific country, territory, or region. Coronaboard [67] provides COVID-19 statistics globally for the USA and three other countries: the Netherlands, France, and South Korea in their respective languages. Almost all countries affected by the Coronavirus pandemic have put forward their efforts to inform their people using easy-to-use and understandable communication methods. For instance, Japan has developed its dashboard [68] and supported the Japanese language and other languages such as English, Deutsch, Français, and Arabic. Moreover, different media groups such as CNN, the BBC, and New York Times come up with sophisticated dashboards, which rely heavily on Johns Hopkins University data.

Research Design

This research reports on Phase 1 of an action research project, where we design a conceptual framework for categorizing COVID-19 resources and survey the efforts and initiatives that contributed to creating these resources. This first phase establishes the infrastructure for locating and placing the right information at the right place. It provides a schema guideline for developing implementation models and systems.

Phase 2 is developing a software system and testing the viability of the proposed approach by undertaking the software development process model for each component and deploying on to an open repository for global access.

Figure 10 provides an overview of the two phases of the action research design. Planning includes brainstorming sessions, informal and formal meetings, surveying the efforts and initiatives of COVID-19 to collect data. The organization is the initial data arrangement to get a sense of the data elements through a visual technique that includes bar graphs, pie charts, and histograms. Conceptual design is the design of an architecture framework for the categorization of resources. Algorithm design involves knowledge graphs, data management and querying, data-driven approaches, and user interface & user experience (UI/UX). Development is the actual implementation of the designed methods in chosen programming languages. Deployment is the reflection and availability of developed methods and models to the community via open-access platforms like GitHub. As described, at any stage of Phase 2, the process of Phase 1 can be accessed and replicated.

JMIR Preprints JMIR Publications

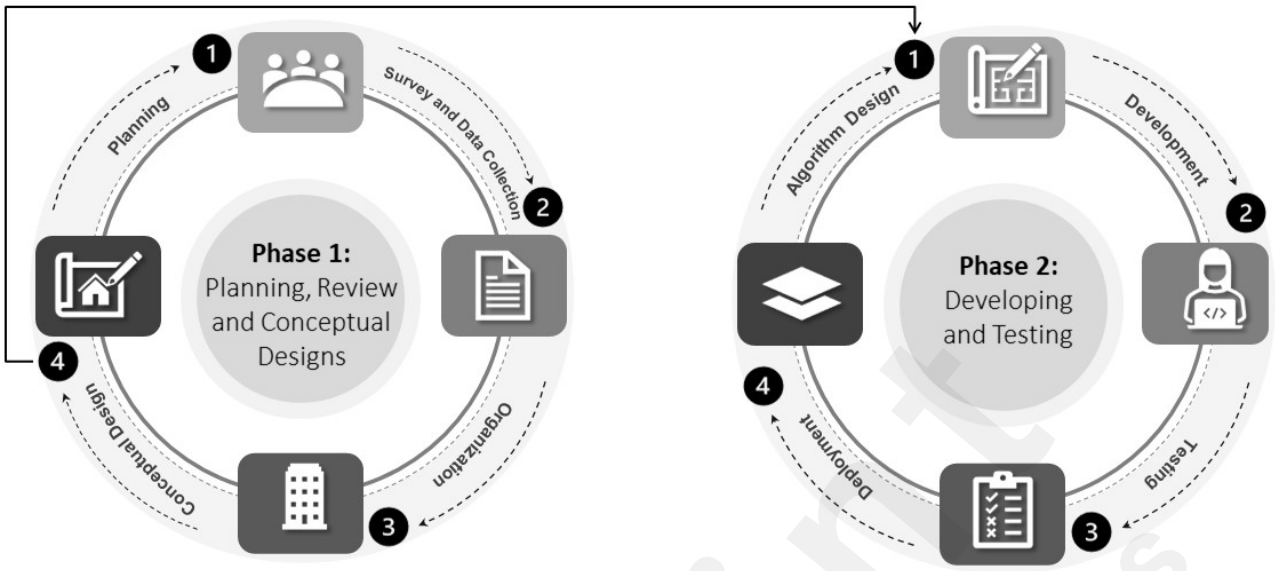

Figure 10: Action research phases

Results and discussion

Investigating diverse information at the three levels discussed earlier enabled us to develop an enterprise architecture framework for COVID-19 related knowledge resources. As illustrated in Figure 4, the framework comprises the categorization of efforts and initiatives that organize resources at three levels; primary, secondary, and tertiary, where resources are interconnected and show the flow of information.

The framework's key aspect is to enable the tracking of a trail of context in terms of information resources at different levels. For instance, if a user is interested in some information that is an outcome of a secondary level initiative, the dashboard shall provide the required information resource(s) along with the dependent resource(s) at the first level. The resources are then visualized in the form of dependency graphs generated automatically in response to the user query.

Figure 11 is a high-level conceptual illustration of the organization of COVID-19 resources in three logical layers connected to analytics that enable the creation of user applications such as analytical dashboards. In the following sections, we discuss the physical design of resources at different levels with relationships and properties, which realizes the approach's practicality.

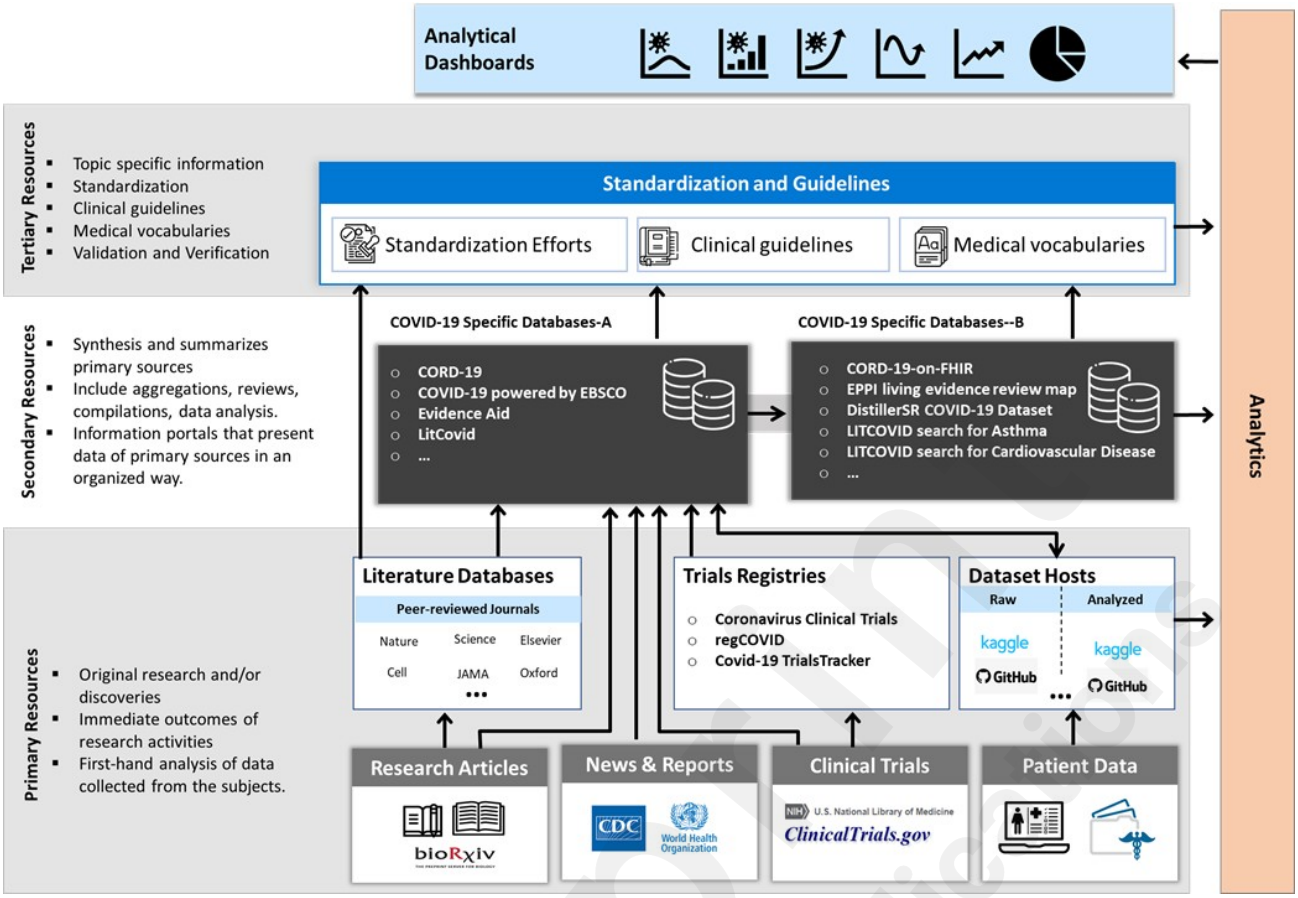

Figure 11: An enterprise architecture framework for COVID-19 related efforts and initiatives involving primary, secondary, and tertiary knowledge resources

Conceptual modeling using a knowledge graph

To manage the highly interlinked COVID-19 information resources and metadata, a robust and semantically enriched model is needed. Based on the proposed framework, we created a knowledge graph design for the representation of interlinked resources to put data in context via linking and semantic information. An example scenario is presented in Figure 12 to realize the suitability of the knowledge graph. Let us have six resources named A, B, C, D, E, and F, where three resources (A, B, and C) belong to primary, two resources (D and E) to secondary, and 1 (F) to tertiary. Let resource F is dependent on resource D, which depends on resources B and C. Similarly, let resource E is dependent on resource A. The corresponding nodes in the knowledge graph for representing the dependency among resources is represented using a relationship node "Depends On," which has two attributes, "Source Category" and "Target Category," that refer to the category of resources it depends on and the category of resource created as a result, respectively. Each resource has at least three attributes: identifiers, name, and type, where the type represents the sub-category within primary, secondary, or tertiary. For instance, there are resources whose type is data resources, and others are research publication resources in the primary category.

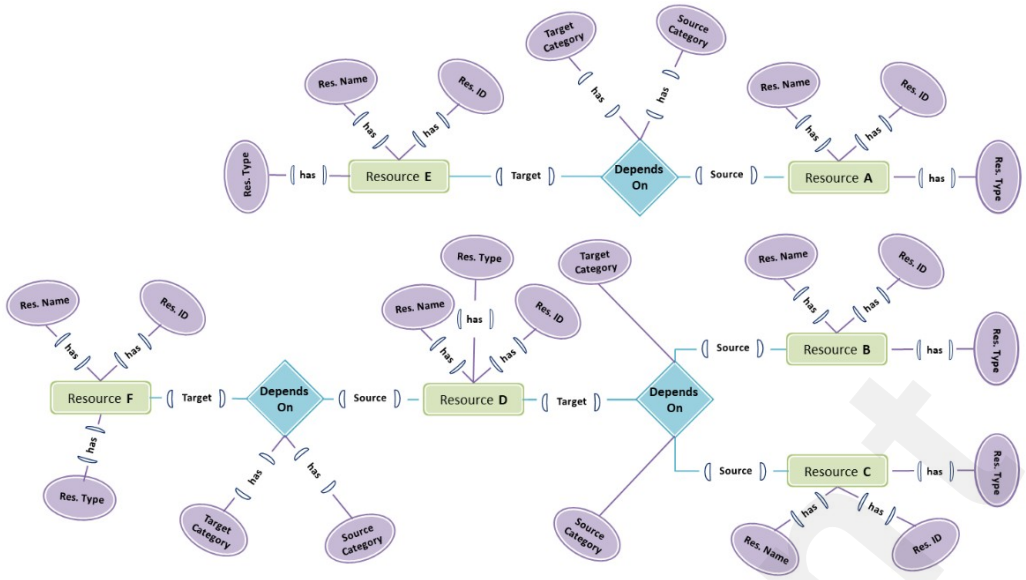

Figure 12: Multiple linked list-based data structure design for three levels represented as list parent and resources in belong to three levels (L1: level 1, L2: level 2, and L3: level 3) are represented in the list child (Nodes: A~F) with pointers to other nodes (resources).

Maintenance in the knowledge graph

The rich feature set of knowledge graphs that include classes, relationship types, and categorization support are of great help to resolve the resource maintenance challenge. Adding a new resource to the database by assigning the correct class of primary, secondary, and tertiary requires the users to input meta-information about the resource. The input data may consist of dichotomy questions to support structured queries run on the underline graph database to infer the target class for the resource. While saving a new resource must check for duplication by checking multi-parameters.

Visualization through the knowledge graph

Current COVID-19 dashboards typically represent time-series visuals and geographic maps, with exceptions of dashboards that respond to the pandemic by showing clinical trials, policy and finance-related interventions, and social distancing directives [69]. Some dashboards include contact tracing data; however, resource dependency tracking dashboards are not seen in the literature. Novel knowledge graph visualization algorithms are emerging, which could be customized to support resource tracking with their dependencies and the metadata semantics.

Future applications

This study's resource categorization can help develop applications for various purposes for combatting the Novel Coronavirus. First, it will provide a baseline for further categorization of the resources in each level. Further classifications can be introduced in each level to make its use more specialized and customized. Second, it will allow developing customized search engines for users' guidance to get results more precisely. Third, it will enable specialized dashboards constructed over the information structured in three levels. Finally, the proposed framework is extendable to bring clinical and genomic resources together by highlighting their associations to help disease progression and tracking.

Conclusion

We conducted this study to classify COVID-19 resources into a three-level structure. We designed a

two-phase research design approach to accomplish the task of locating COVID-19 resources, their placement, and conceptualizing the architecture in the first phase, while developing and testing of the contextual dashboards in the second phase. We surveyed various efforts and initiatives worldwide, provide descriptive statistics, and classified them into the primary, secondary, and tertiary categories. The proposed categorization led us to design knowledge graph-based models for developing contextual dashboard. A dependency graph theory is incorporated to visualize the results in the form of information resources and their interdependence. The proposed work enables other applications such as search engines, interactive dashboards, and tracking systems. Using knowledge graph models for the application development will add transparency to the information infrastructure, thus increases the trust factor on the applications.

Acknowledgement

This work was supported in part by the Ministry of Science and ICT (MSIT), South Korea, through the Information Technology Research Center (ITRC) Support Program under Grant IITP-2017-0-01629, and in part by the Institute for Information & Communications Technology Promotion (IITP), through the Korea Government (MSIT) under Grant 2017-0-00655 and Grant NRF-2019R1A2C2090504.

Preprint JMIR Publications

References

1. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. Acta Biomed. Mattioli 1885; 2020. p. 157–160. PMID:32191675
2. User Education Services. Primary, Secondary, and Tertiary Sources. Univ Maryl Univ Libr [Internet] 2012 [cited 2020 Sep 18]; Available from: http://www.lib.umd.edu/ues/guides/primary-sources
3. UMC. Primary, Secondary, and Tertiary Sources | University of Minnesota Crookston [Internet]. Univ Minnesota Crookst. [cited 2020 Sep 18]. Available from: https://www.crk.umn.edu/library/primary-secondary-and-tertiary-sources
4. Q. What are primary, secondary, and tertiary sources? - LibAnswers [Internet]. [cited 2020 Sep 18]. Available from: https://apus.libanswers.com/faq/2299
5. Primary, Secondary and Tertiary Literature in the Sciences - Pharmacy - Research Resources - LibGuides at Ohio Northern University [Internet]. [cited 2020 Sep 18]. Available from: https://library.onu.edu/c.php?g=357773&p=2720217
6. Learning T&, Libraries U. Primary, Secondary & Tertiary Sources. The Ohio State University;
7. arXiv.org e-Print archive [Internet]. [cited 2020 Jun 26]. Available from: https://arxiv.org/
8. Cold Spring Harbor Laboratory. bioRxiv COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv [Internet]. bioRxiv Preprunt Serv Biol. 2020 [cited 2020 Jun 26]. Available from: https://connect.biorxiv.org/relate/content/181
9. JHU Coronavirus COVID-19 Global Cases, by country | Marketplace - Google Cloud Platform [Internet]. [cited 2020 Jun 10]. Available from: https://console.cloud.google.com/marketplace/details/johnshopkins/covid19_jhu_global_cases
10. About COVID-19 Public Datasets | Marketplace - Google Cloud Platform [Internet]. [cited 2020 Jun 10]. Available from: https://console.cloud.google.com/marketplace/details/bigquery-public-datasets/covid19-public-data-program?_ga=2.50179686.-1918167096.1591790334
11. Search | Kaggle [Internet]. [cited 2020 Jun 10]. Available from: https://www.kaggle.com/search?q=COVID-19
12. Search · COVID-19 · GitHub [Internet]. [cited 2020 Jun 10]. Available from: https://github.com/search?q=COVID-19
13. covid-19 data on data.world | 34 datasets available [Internet]. [cited 2020 Jun 11]. Available from: https://data.world/datasets/covid-19
14. COVID-19 Open Datasets [Internet]. [cited 2020 Jun 11]. Available from: https://www.aminer.cn/data-covid19/
15. COVID-19 | IEEE DataPort [Internet]. [cited 2020 Jun 11]. Available from: https://ieee-dataport.org/topic-tags/covid-19
16. World Health Organization. WHO Coronavirus Disease (COVID-19) Dashboard [Internet]. Covid-19 Dashboard. 2020 [cited 2020 Jun 11]. p. 1–1. Available from: https://covid19.who.int/?gclid=Cj0KCQjwiYL3BRDVARIsAF9E4GdZ0og0f2VbDJL6b6GGh9uzAKSWemLaej8CEN01Sv-_MHMu8eprTpYaAligEALw_wcB
17. Singh J. Centers for disease control and prevention [Internet]. Indian J Pharmacol. 2004 [cited 2020 Jun 17]. p. 268–269. [doi: 10.1097/01.phm.0000189841.18419.16]
18. World Health Organization. About WHO [Internet]. 2019. 1948 [cited 2020 Jun 17]. Available from: https://www.who.int/about
19. COVID-19 Coronavirus data - Datasets [Internet]. [cited 2020 Jun 11]. Available from: https://

data.europa.eu/euodp/en/data/dataset/covid-19-coronavirus-data
20. Johns Hopkins CSSE. Coronavirus COVID-19 (2019-nCoV) [Internet]. Coronavirus COVID-19 Glob Cases by Johns Hopkins CSSE. 2020 [cited 2020 Jun 18]. p. 1. Available from: https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6
21. World Bank Group. Understanding the Coronavirus (COVID-19) pandemic through data [Internet]. 2020 [cited 2020 Jun 11]. Available from: http://datatopics.worldbank.org/universal-health-coverage/coronavirus/
22. The latest real-time epidemic map of global new coronavirus [Internet]. [cited 2020 Jun 18]. Available from: https://ncov.dxy.cn/ncovh5/view/pneumonia
23. U.S Department of Health & Human Services. Coronavirus (COVID-19) | National Institutes of Health (NIH) [Internet]. 2020. 2020 [cited 2020 Jun 18]. Available from: https://www.nih.gov/coronavirus
24. Simes RJ. Publication bias: The case for an international registry of clinical trials. J Clin Oncol 1986 Sep 22;4(10):1529–1541. [doi: 10.1200/JCO.1986.4.10.1529]
25. Huser V, Cimino JJ. Linking ClinicalTrials.gov and PubMed to Track Results of Interventional Human Clinical Trials. PLoS One [Internet] 2013 [cited 2020 Jun 18];8(7). PMID:23874614
26. WHO | Welcome to the WHO ICTRP [Internet]. [cited 2020 Jun 18]. Available from: https://www.who.int/ictrp/en/
27. Search of: COVID-19 - List Results - ClinicalTrials.gov [Internet]. [cited 2020 Jun 18]. Available from: https://clinicaltrials.gov/ct2/results?cond=COVID-19&term=&type=&rslt=&age_v=&gndr=&intr=&titles=&outc=&spons=&lead=&id=&cntry=&state=&city=&dist=&locn=&rsub=&strd_s=&strd_e=&prcd_s=&prcd_e=&sfpd_s=&sfpd_e=&rfpd_s=&rfpd_e=&lupd_s=&lupd_e=&sort=
28. ChiCTR-PRC-13003310. Chinese Clinical Trial Register (ChiCTR) - The world health organization international clinical trials registered organization registered platform [Internet]. 2018 [cited 2020 Jun 19]. p. 3–4. Available from: http://www.chictr.org.cn/abouten.aspx
29. ISRCTN trial registry. ISRCTN - Search results [Internet]. BioMed Cent Ltd, Part Springer nature. [cited 2020 Jun 18]. Available from: https://www.isrctn.com/search?q=COVID-19
30. SemanticScholar. CORD-19 | Semantic Scholar [Internet]. 2020 [cited 2020 Jun 26]. Available from: https://www.semanticscholar.org/cord19
31. Lu Wang L, Lo K, Chandrasekhar Y, Reas R, Yang J, Eide D, Funk K, Kinney R, Liu Z, Merrill W, Mooney P, Murdick D, Rishi D, Sheehan J, Shen Z, Stilson B, Wade AD, Wang K, Wilhelm C, Xie B, Raymond D, Weld DS, Etzioni O, Kohlmeier S. CORD-19: The Covid-19 Open Research Dataset. ArXiv [Internet] arXiv; 2020 Apr 22 [cited 2020 Jun 26]; PMID:32510522
32. US National Library of Medicine - National Institutes of Health. Home - PMC - NCBI [Internet]. www.ncbi.nlm.nih.gov. 2017 [cited 2020 Jun 26]. [doi: 10.1166/asl.2017.8751]
33. Chen Q, Allot A, Lu Z. LitCovid: An open database of COVID-19 literature. Nucleic Acids Res [Internet] Oxford University Press; 2021 Jan 8 [cited 2021 Mar 22];49(D1):D1534–D1540. PMID:33166392
34. NIH. LitCovid - NCBI - NLM - NIH [Internet]. 2020 [cited 2020 Jul 9]. Available from: https://www.ncbi.nlm.nih.gov/research/coronavirus/
35. Chen Q, Allot A, Lu Z. Keep up with the latest coronavirus research. Nature. NLM (Medline); 2020. p. 193. PMID:32157233
36. PubTator Central - NCBI - NLM - NIH [Internet]. [cited 2020 Jul 9]. Available from: https://www.ncbi.nlm.nih.gov/research/pubtator/?view=docsum&query=$LitCovid
37. COVID-19 Evidence Alerts from McMaster PLUS | About [Internet]. [cited 2020 Jul 9].

Available from: https://plus.mcmaster.ca/COVID-19/About
38. COVID-19 Advanced Literature Classifier - PICO Portal [Internet]. [cited 2020 Jul 10]. Available from: https://covid.picoportal.net/
39. Coronavirus (COVID-19) - Evidence Aid Evidence Aid [Internet]. [cited 2020 Jul 10]. Available from: https://evidenceaid.org/evidence/coronavirus-covid-19/
40. World Health Organization; Technical guidance publications [Internet]. 2020 [cited 2020 Aug 24]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance-publications?healthtopics=b6bd35a3-cf4f-4851-8e80-85cb0068335b&publishingoffices=aeebab07-3d0c-4a24-b6ef-7c11b7139e43&healthtopics-hidden=true&publishingoffices-hidden=true
41. CDC. Information for Healthcare Professionals about Coronavirus (COVID-19) [Internet]. Cent Dis Control Prev. 2020 [cited 2020 Aug 24]. Available from: https://www.cdc.gov/coronavirus/2019-nCoV/hcp/index.html
42. Duodecim: EBM Guidelines Coronavirus infections | EBSCO COVID-19 Clinical Care [Internet]. [cited 2020 Aug 24]. Available from: https://covid-19clinicalcare.ebscomedical.com/node/59935
43. Iacoviello V. COVID-19 (Novel Coronavirus) [Internet]. DynaMed. 2020 [cited 2020 Aug 24]. Available from: https://www.dynamed.com/condition/covid-19-novel-coronavirus#GUIDELINES
44. COVID-19 Knowledge Accelerator [Internet]. [cited 2020 Aug 20]. Available from: https://www.gps.health/covid19_knowledge_accelerator.html
45. EBMonFHIR - Clinical Decision Support - Confluence [Internet]. [cited 2020 Aug 20]. Available from: https://confluence.hl7.org/display/CDS/EBMonFHIR
46. Richardson JE, Alper BS, Lehmann HP, Subbian V. It is time for computable evidence synthesis: The COVID-19 Knowledge Accelerator Initiative. J Am Med Informatics Assoc [Internet] Oxford University Press (OUP); 2020 May 22 [cited 2020 Aug 20]; [doi: 10.1093/jamia/ocaa114]
47. Alper BS, Dehnbostel J, Afzal M, Subbian V, Soares A, Kunnamo I, Shahin K, McClure RC. Making Science Computable: Developing Code Systems for Statistics, Study Design, and Risk of Bias. J Biomed Inform [Internet] Academic Press; 2021 Jan 21 [cited 2021 Feb 5];103685. PMID:33486066
48. Workflow Tool Examples | AHRQ Digital Healthcare Research: Informing Improvement in Care Quality, Safety, and Efficiency [Internet]. [cited 2020 Aug 21]. Available from: https://digital.ahrq.gov/covid-acts
49. Roy M. Gulick, H. Clifford Lane, Henry Masur et al. Guidelines Introduction | Coronavirus Disease COVID-19 [Internet]. [cited 2020 Aug 25]. Available from: https://www.covid19treatmentguidelines.nih.gov/introduction/
50. Surgeons AC of. COVID-19: Elective Case Triage Guidelines for Surgical Care (Outline published March 24, 2020) [Internet]. 2020. Available from: https://www.facs.org/covid-19/clinical-guidance/elective-case
51. SNOMED - COVID-19 [Internet]. [cited 2020 Aug 25]. Available from: https://www.snomed.org/snomed-ct/covid-19
52. SNOMED CT COVID-19 Related Content - Announcements - SNOMED Confluence [Internet]. [cited 2020 Aug 25]. Available from: https://confluence.ihtsdotools.org/display/snomed/SNOMED+CT+COVID-19+Related+Content

53. SANER Project Using HL7 FHIR to Enable Easier Reporting for Public Health [Internet]. [cited 2020 Sep 3]. Available from: https://blog.hl7.org/saner-project-using-hl7-fhir-to-enable-easier-reporting-for-public-health-reports
54. International Organization for Standardization. ISO - COVID-19 response: freely available ISO standards [Internet]. 2020 [cited 2020 Sep 4]. Available from: https://www.iso.org/covid19
55. ISO. COVID-19: National Resources [Internet]. [cited 2020 Sep 4]. Available from: https://www.iso.org/covid19-members
56. Standards Alliance Phase 2 Supports COVID-19 Response and Recovery [Internet]. [cited 2020 Sep 4]. Available from: https://www.ansi.org/news_publications/news_story?menuid=7&articleid=d4868f3d-e393-4baa-94e5-9ea1e239b228
57. Sultan S, Altayar O, Siddique SM, Davitkov P, Feuerstein JD, Lim JK, Falck-Ytter Y, El-Serag HB. AGA Institute Rapid Review of the Gastrointestinal and Liver Manifestations of COVID-19, Meta-Analysis of International Data, and Recommendations for the Consultative Management of Patients with COVID-19. Gastroenterology [Internet] 2020 [cited 2020 Sep 4];159(1):320-334.e27. PMID:32407808
58. Coronavirus Disease 2019 (COVID-19) | FDA [Internet]. [cited 2020 Sep 4]. Available from: https://www.fda.gov/emergency-preparedness-and-response/counterterrorism-and-emerging-threats/coronavirus-disease-2019-covid-19
59. Centers for Disease Control and Prevention. ICD-10-CM Official Coding and Reporting Guidelines, April 1, 2020 through September 30, 2020. 2020 [cited 2020 Sep 4];19(i):19–21. Available from: https://www.cdc.gov/nchs/data/icd/COVID-19-guidelines-final.pdf
60. ICD-10-CM [Internet]. [cited 2020 Sep 4]. Available from: https://icd10cmtool.cdc.gov/?fy=FY2020&q=covid-19
61. Covid-19 Terms [Internet]. [cited 2020 Sep 4]. Available from: https://metamap.nlm.nih.gov/Covid19Terms.shtml
62. New MeSH Supplementary Concept Record for Coronavirus Disease 2019 (COVID19). NLM Tech Bull [Internet] 2020 [cited 2020 Sep 4];1–2. Available from: https://www.nlm.nih.gov/pubs/techbull/jf20/brief/jf20_mesh_novel_coronavirus_disease.html
63. MeSH Browser [Internet]. [cited 2020 Sep 4]. Available from: https://meshb.nlm.nih.gov/record/ui?ui=C000657245
64. SARS-CoV-2 and COVID-19 related LOINC terms – LOINC [Internet]. [cited 2020 Sep 4]. Available from: https://loinc.org/sars-cov-2-and-covid-19/
65. Worldometer. Coronavirus Update (Live): Cases and Deaths from COVID-19 Virus Pandemic [Internet]. Worldometers. 2020 [cited 2020 Sep 17]. p. 1. Available from: https://www.worldometers.info/coronavirus/
66. Worldometer. Coronavirus Cases [Internet]. Worldometer. 2020 [cited 2020 Sep 17]. p. 1–22. [doi: 10.1101/2020.01.23.20018549V2]
67. COVID-19 Dashboard USA [Internet]. [cited 2020 Sep 17]. Available from: https://coronaboard.com/
68. Reustle. (15791) Coronavirus Disease (COVID-19) Japan Tracker [Internet]. [cited 2020 Sep 17]. Available from: https://covid19japan.com/
69. Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, Rees G, Emery VC, Stevens MM, Keegan N, Short MJ, Pillay D, Manley E, Cox IJ, Heymann D, Johnson AM, McKendry RA. Digital technologies in the public-health response to COVID-19. Nat Med [Internet] Nature Research; 2020 Aug 1 [cited 2021 Jan 27];26(8):1183–1192. PMID:32770165