 doi:10.20944/preprints202209.0323.v1

Data Descriptor

CoviRx: A user-friendly interface for systematic down-selection of repurposed drug candidates for COVID-19

Hardik A. Jain [1], Vinti Agarwal [2,*] Chaarvi Bansal [2], Anupama Kumar [3], Faheem [4], Muzaffar-Ur-Rehman Mohammed [4], Sankaranarayanan Murugesan [4], Moana M. Simpson [5], Avinash V. Karpe [6,12], Rohitash Chandra [7], Christopher A. MacRaild [8], Ian K. Styles [8], Amanda L. Peterson [8], Matthew A. Cooper [9], Carl M. J. Kirkpatrick [10], Rohan M. Shah [11], Enzo A. Palombo [11], Natalie L. Trevaskis [8], Darren J. Creek [8], Seshadri S. Vasan [12,13,*] and on behalf of the sySTEMs Initiative [14]

[1] Birla Institute of Technology and Science, Department of Electrical and Electronics Engineering, Pilani, Rajasthan, India; f20190318@pilani.bits-pilani.ac.in (H.A.J.)
[2] Birla Institute of Technology and Science, Department of Computer Science and Information Systems, Pilani, Rajasthan, India; vinti.agarwal@pilani.bits-pilani.ac.in (V.A.); f20180913@pilani.bits-pilani.ac.in (C.B.)
[3] Commonwealth Scientific and Industrial Research Organisation, Land and Water, Waite Campus, South Australia 5064, Australia; anupama.kumar@csiro.au (A.K.)
[4] Birla Institute of Technology and Science, Department of Pharmacy, Pilani, Rajasthan, India; muzaffar.rehman@pilani.bits-pilani.ac.in (M-U-R.M.); faheem@pilani.bits-pilani.ac.in (F.); murugesan@pilani.bits-pilani.ac.in (S.M.)
[5] Compounds Australia, Griffith Institute for Drug Discovery, Griffith University, Nathan QLD 4111, Australia; m.simpson@griffith.edu.au (M.S); compounds-australia@griffith.edu.au
[6] Commonwealth Scientific and Industrial Research Organisation, Land and Water, Dutton Park, Queensland, Australia; akarpe@swin.edu.au (A.V.K.)
[7] University of New South Wales, School of Mathematics and Statistics, NSW 2052, Australia; rohitash.chandra@unsw.edu.au (R.C.)
[8] Monash University, Drug Delivery, Disposition and Dynamics, Monash Institute of Pharmaceutical Sciences, Parkville, Victoria 3052, Australia; chris.macraild@monash.edu (C.A.M.); ian.styles@monash.edu (I.K.S.); amanda.peterson@monash.edu (A.L.P.); natalie.trevaskis@monash.edu (N.L.C.); darren.creek@monash.edu (D.J.C.)
[9] The University of Queensland, Institute for Molecular Bioscience, Brisbane, Queensland, Australia; m.cooper@uq.edu.au (M.C.)
[10] Monash University, Centre for Medicine Use and Safety, Monash Institute of Pharmaceutical Sciences, Parkville, Victoria 3052, Australia, carl.kirkpatrick@monash.edu (C.M.J.K.)
[11] Swinburne University of Technology, Department of Chemistry and Biotechnology, Hawthorn, Victoria, Australia; rshah@swin.edu.au (R.S.); epalombo@swin.edu.au (E.A.P.); akarpe@swin.edu.au (A.V.K.)
[12] Commonwealth Scientific and Industrial Research Organisation, Health and Biosecurity, Geelong, Victoria 3220, Australia
[13] University of York, Department of Health Sciences, York, YO10 5DD, United Kingdom; prof.vasan@york.ac.uk (S.S.V.)
[14] sySTEMs Initiative, Commonwealth Scientific and Industrial Research Organisation, Australian Centre for Disease Preparedness, Geelong, Victoria, Australia

* Correspondence: Dr. Vinti Agarwal, E-mail: vinti.agarwal@pilani.bits-pilani.ac.in or Prof. Seshadri S. Vasan, E-mail: prof.vasan@york.ac.uk

Abstract: Although various vaccines are now commercially available, they have not been able to stop the spread of COVID-19 infection completely. An excellent strategy to quickly get safe, effective, and affordable COVID-19 treatment is to repurpose drugs that are already approved for other diseases as adjuvants along with the ongoing vaccine regime. The process of developing an accurate and standardized drug repurposing dataset requires a considerable level of resources and expertise due to the commercial availability of an extensive array of drugs that could be potentially used to address the SARS-CoV-2 infection. To address this bottleneck, we created the CoviRx platform. CoviRx is a user-friendly interface that provides access to the data, which is manually curated for COVID-19 drug repurposing data. Through CoviRx, the data curated has been made open-source to help advance drug repurposing research. CoviRx also encourages users to submit their findings after thoroughly validating the data, followed by merging it by enforcing uniformity and integrity-preserving constraints. This article discusses the various features of CoviRx and its design

 doi:10.20944/preprints202209.0323.v1

principles. CoviRx has been designed so that its functionality is independent of the data it displays. Thus, in the future, this platform can be extended to include any other disease X beyond COVID-19. CoviRx can be accessed at www.covirx.org.

Dataset: DOI number or link to the deposited dataset in cases where the dataset is published or set to be published separately. If the dataset is submitted and will be published as a supplement to this paper in the journal Data, this field will be filled by the editors of the journal. In this case, please make sure to submit the dataset as a supplement when entering your manuscript into our manuscript editorial system.

Dataset License: license under which the dataset is made available (CC0, CC-BY, CC-BY-SA, CC-BY-NC, etc.)

Keywords: COVID-19; Open-source dataset; Drug Repurposing; Database system; Web application development; software development; Drug fingerprints; Bulk upload

1. Summary

COVID-19 pandemic, caused by the SARS-CoV-2 virus, has created significant impacts throughout the entire world [1,2]. As of January 2022, more than 378 million cases have been reported, resulting in about 5.7 million deaths globally [3]. Despite recent advances in drug discovery, the treatment of viral infections remains a significant challenge for scientists worldwide. Although various vaccines are now available in the market, they have not been sufficient to stop the spread of COVID 19. An efficient and economic strategy for satisfying the urgent requirement of safe, effective, and affordable COVID-19 treatments is to repurpose the drugs already approved for other diseases [4,5]. This technique, however, comes with its own set of challenges. One of such challenges is data availability. Public access to high-quality, valuable information such as clinical trials data is still limited. Although there are a few open-access databases, such as the Drug Repurposing Hub [6], which provide drug repurposing candidates, very few are specific to COVID-19, such as Excelra [7] and NCATS [8]. Most of the databases present in public repositories offer data in a non-standard format, making it difficult to retrieve and interpret [9].

Additionally, significant sizes of these databases often create computational bottlenecks. Through this study, we explored to solve the problem of data accessibility, availability, and interpretability by developing CoviRx, a robust web-based application. The name of CoviRx is derived from the terms COVID-19 and Rx (a medical prescription), as it provides access to approved drugs repurposed for COVID-19.

CoviRx provides users with an interactive interface to access more than 7000 drugs for COVID-19 drug repurposing. The web application (Henceforth mentioned as either App or app) presents physical and chemical properties, original indication, assay data on multiple assays, COVID clinical trials, and red flags (pregnancy concerns and contraindications, among others) for these drugs. It dynamically calculates drugs similar to the query drug using the Tanimoto coefficient [10]. It displays it along with 11 filters designed by the research team that the drug has passed through. This work has been submitted as an interlinked manuscript "Systematic down-selection of repurposed drug candidates for COVID-19" to the *International Journal of Molecular Sciences;* MacRaild et al.,2022. To increase the collaboration with the other researchers, features have been incorporated in CoviRx that allow the registered users to submit their drugs and findings to the CoviRx web app. Followed by a thorough peer review process, these submissions will be merged with the primary database, which we herein report. The user can download the data available in JSON and PDF formats or share it using persistent and unique URLs. This provides an elegant solution to the problem of data non-standardization and lack of data mining, interpretation, and manipulation-friendly datasets [11].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2022 doi:10.20944/preprints202209.0323.v1

CoviRx has been designed according to current industry standards. Stringent checks are placed to ensure data consistency, security, and uniformity. Unique identifiers have been added for the individual drugs to ensure that they are easy to search and data for all the drugs is easily accessible. Modern frameworks such as Django [12,13], jQuery [14], Bootstrap5 [15] have been used for development, and cloud hosting services like Azure have been used to ensure that the web app is scalable, i.e., able to handle an increase in website traffic. It supports different user profiles to protect the integrity of data. Only registered users can submit to the website to maintain the data uniformity and reliability and prevent malicious activities suffered by the prominent open-source platforms. Only the users with administration rights have modification privileges. Numerous graphical representations [16] **are present throughout the web app to give the users an overview of the dataset. This adds to the user experience. Its modular architecture ensures that it is flexible enough to be extended to any other disease** 'X' without disrupting its underlying functionality. CoviRx thus is a one-stop application for accessing high-quality data related to COVID-19 drug repurposing and, in the future, any other disease X.

2. Data Description

Drug data (with assays) was manually curated by our team from various research journals and publicly available drug databases. Regulatory approval data were obtained from U.S. Food and Drug Administration (FDA) [17], Therapeutic Drug Administration, Australia (TGA) [18], Inxight Drugs websites [19], and clinical trials data was acquired from clinicaltrials.gov [20] for more than **7000 drugs**.

These drugs were then passed through a series of 11 filters, and the resultant drugs were screened for possible action against COVID-19 to narrow down to ten drugs for further studies on COVID-19 drug repurposing (submitted as an interlinked manuscript, MacRaild et al.,2022 to the *International Journal of Molecular Sciences*).

Manual curation of data and datasets requires considerable time, labor, and financial resources. Therefore, to economise these resources, it was decided to open source this dataset and make it available with a user-friendly graphical user interface so that other researchers working on their filters for down-selecting drugs can use it.

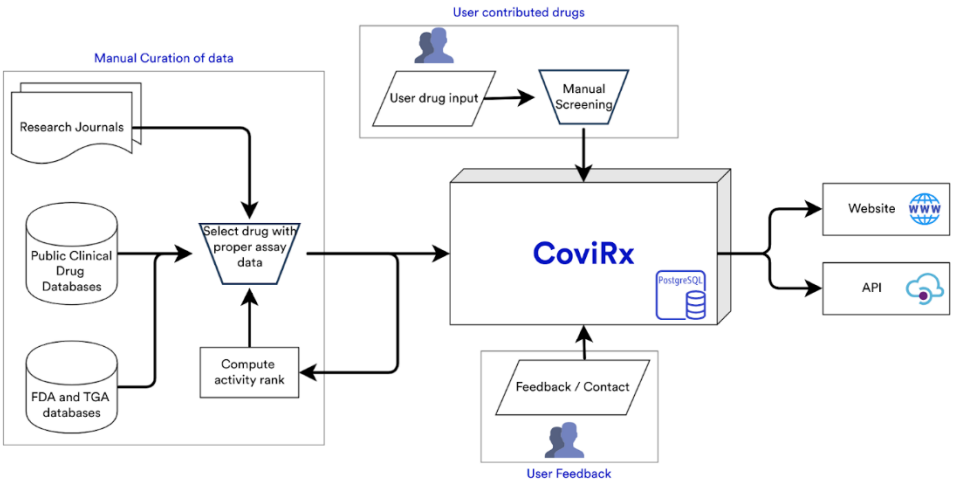

Figure 1: CoviRx components and workflow

The CoviRx website consists of various identifiers such as synonyms, Chemical Abstract Service (**CAS) number, PubChem ID, ChEMBL id**, among others, which are hyperlinked to provide a seamless transition across different databases. Various drug-likeness properties like molecular weight, Partition Coefficient (logP), hydrogen bond donors

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2022 doi:10.20944/preprints202209.0323.v1

(HBD), hydrogen bond acceptors (HBA), among others, are also included for each compound to assist with early drug discovery programs. A drug's original indication, mechanism of action (MOA), relevant pathways it interferes with, and associated targets are also included in the CoviRx website to provide users with a brief background about its original application. CoviRx provides users with SARS-CoV-2 drug repurposing data from several COVID-19 drug repurposing studies and provides the status of such drugs in COVID-19 trials. The website's pharmacokinetics (PK) section includes information regarding a compound's route of administration, the volume of distribution, clearance, protein binding, etc. The CoviRx website also consists of a specially curated section called "red flags," which sheds light on drug-drug interactions, contraindications, pregnancy category of a drug, breastfeeding concerns, and severe side effects or black box warnings. To avoid early confounders in drug repurposing programs, predicted CAD (Cationic Amphiphilic Drugs) induced phospholipidosis and Pan Assay Interference Compounds (PAINS) information are also included in the red flags section. Specific data points related to drug activity on the original target have not been included. They do not have well-defined values or data uniformity but can be added in the future.

The Data has been segmented into different tables for storing it efficiently. Figure 2 depicts different tables and the data fields used to store the drug data, and Figure 3 illustrates the same for user data. The data types of each field and the type of constraints enforced (if any) are also highlighted.

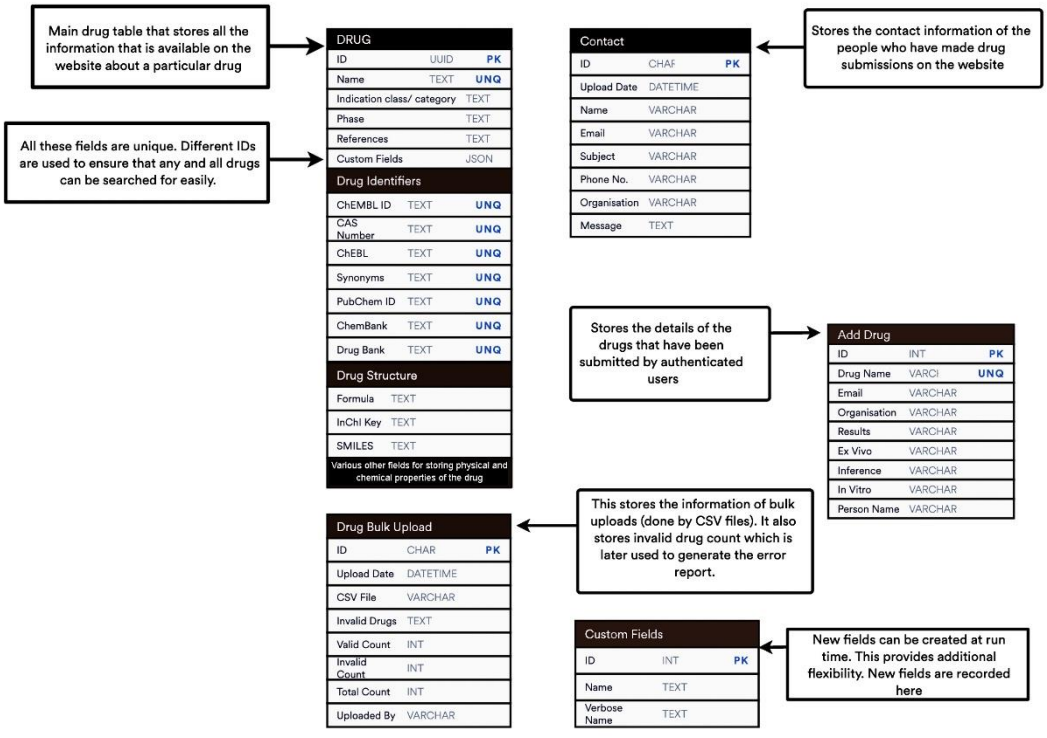

Figure 2 Tables and data fields used for storing drug information

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2022 doi:10.20944/preprints202209.0323.v1

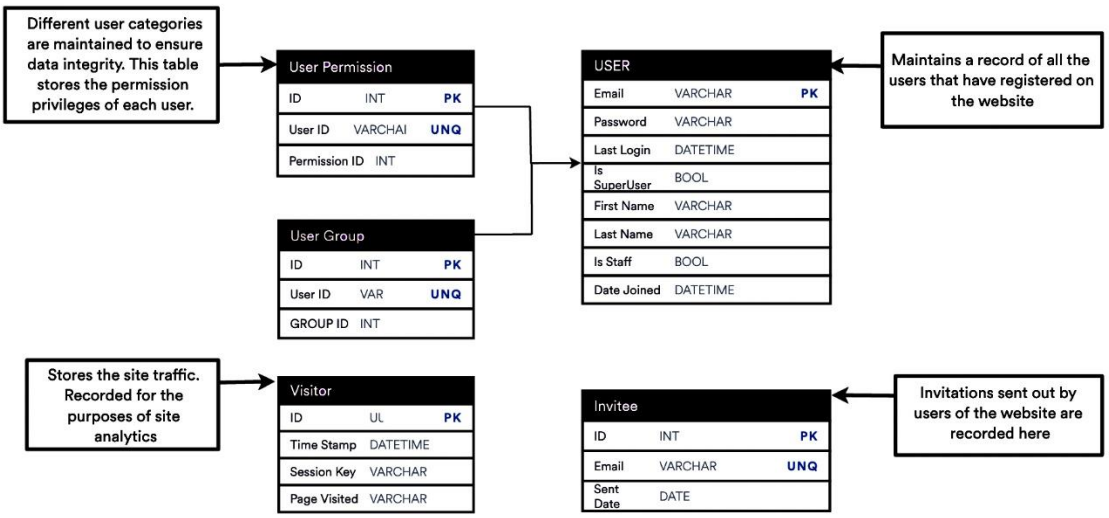

Figure 3 Tables and data fields used for storing user information

3. Methods

3.1 Tech Stack and Design Principles

Modern technologies were employed to develop the front and back end of the web app. Figure 4 demonstrates the interaction between different components like front-end and back-end servers and the frameworks we have used to develop these. These frameworks are open-source, well documented, and highly flexible to use. Back-end servers such as Apache [21] and cloud hosting platforms such as **Azure** were preferred over other alternatives as these provided support for both vertical and horizontal scalability. Currently, SQLite [22], a relational database management system, is being used to handle website data; however, it will be migrated to PostgreSQL soon [23,24] as it offers better concurrency control [25]. Google OAuth [26] (login/logout), Google reCAPTCHA [27], and Google simple mail transfer protocol (SMTP) [28] have been used during development. CoviRx sends emails to its users for various purposes, such as a copy of responses filled by them in contact forms and contribute drug forms, an invitation to the admin panel, and reporting errors that occur during drug upload. Since emails do not support CSS or JS files, and to deliver well-styled HTML emails to our users, a python library called Premailer was used [29].

Few properties of database systems and websites [30–32] form the foundation for CoviRx's architecture. These design principles have been discussed in detail in Table 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2022 doi:10.20944/preprints202209.0323.v1

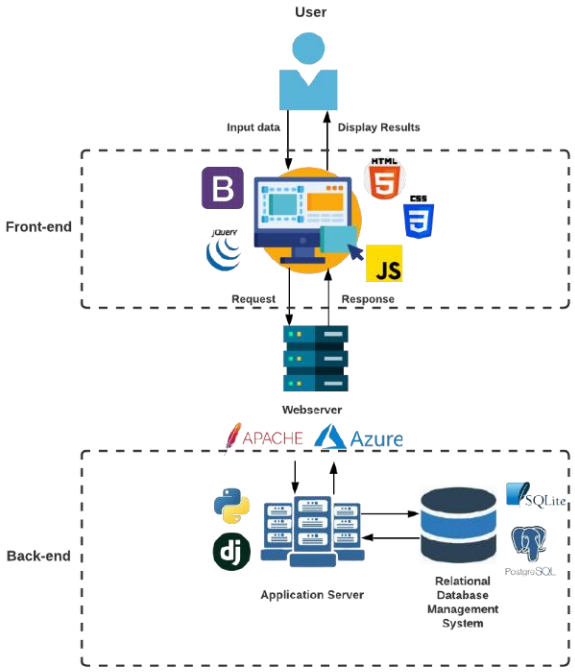

Figure 4: Tech stack used for developing CoviRx and the interaction between these frameworks. User interacts with the front-end, built using HTML, CSS, JQuery3 [14], Bootstrap 5 [15], to request data. Front end then interacts with the webservers, i.e., Azure and Apache. These fetch data from the application servers powered by Django [12,13,33], which in turn interacts with the database management system, i.e., SQLite [22]. Once this Data is extracted, it follows a reverse path back to the user.

 doi:10.20944/preprints202209.0323.v1

Table 1: Design Principles adopted during the development of CoviRx

Design Principles	Implementation	Advantages	Framework Support
Data integrity	To ensure that no unauthorized access happens, users were divided into three broad categories: 1. ***Unauthenticated users*:** This type of user can read from the database but cannot modify the database. 2. ***Users with limited privileges**:* This type of user can read and submit drugs that users with admin privileges will verify before merging. 3. ***Admin users**:* This type of user has all the privileges, i.e., read, write, update, delete.	• Ensures data security. • Maintains data quality and data consistency. • Ensures system reliability.	---
Data security	Protection against cross-site scripting (XSS), cross-site request forgery (CSRF), SQL injection, and clickjacking is provided by Django. Secure communication between the web app and the server is established as only encrypted Data is shared via the HTTPS protocol.	• Protects the data from attack. • Ensures data privacy.	Django, HTTPS Protocols
Scalability	Modern cloud hosting providers and backend servers support both vertical and horizontal scaling.	• Ability to manage increasing website traffic without disturbing end users. • Hosting becomes less expensive. • Provides multiuser support.	Django, Apache backend server, AWS, Azure, Google Cloud Platform (GCP)
Modularity	The entire website has been broken down into multiple independent components which are interchangeable. The database components are well encapsulated from program components of the web app.	It adds a layer of abstraction that provides data program independence. • The underlying database can be changed without altering the code. • Easy to maintain. • Easy to recognise and fix bugs.	Django

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2022 doi:10.20944/preprints202209.0323.v1

Design Principles	Implementation	Advantages	Framework Support
Integrity Constraints	Various constraints are in place to ensure that only data that passes all quality checks are merged with the primary database. Unique identifier constraint is enforced in multiple fields like drug name, ID, etc. Char fields with maximum length constraints over text fields are used to optimize memory usage.	• Maintains correctness of data. • Maintains data uniformity.	---
Auto Versioning of Static Files	A string specifying the version of the static file is appended to the URL pointing to the static files. This version is updated, forcing the browser to load the latest version instead of the one from the cache.	• Quicker developmental iterations. • Faster load time.	---
Flexibility and Responsiveness	The web app is responsive to a wide range of screen sizes. Django template language makes designing dynamic frontend components easy.	• Multi-device support makes it accessible to a broader audience. • Better user experience.	Django, Bootstrap5, jQuery3, jQuery3 UI
Backup and Recovery	Data backups are done weekly.	Protects the data from system failure, both hardware, and software.	---
Concurrency Control	SQLite is not optimized for performing concurrent operations and often results in Database is locked exceptions. So, we intend to shift to PostgreSQL in the future so that simultaneous operations do not lead to abnormalities.	• Reduces wait time • Reduces response time • Increases resource utilisation • Increases performance and efficiency • Maintains data consistency	SQLite (shifting to PostgreSQL in future)
Data Persistency	If an exception happens between a transaction (write operation), the data to be written might get destroyed. We plan to protect our system in rare cases of failure like this to ensure durability	• Data is never at risk.	Celery [34] (to be implemented)

 doi:10.20944/preprints202209.0323.v1

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2022 doi:10.20944/preprints202209.0323.v1

3.2 CoviRx Architecture

An iterative software development model was followed [35] while designing CoviRx. First, this was achieved by creating a wireframe [36] for every front-end page, followed by a basic application implementation. The design components were then modified in each version by adding new features and functionalities. Each component was tested individually [37] before merging with the main framework. Upon its completion, the entire application was tested [38]. Figure 5 depicts various pages/features of CoviRx.

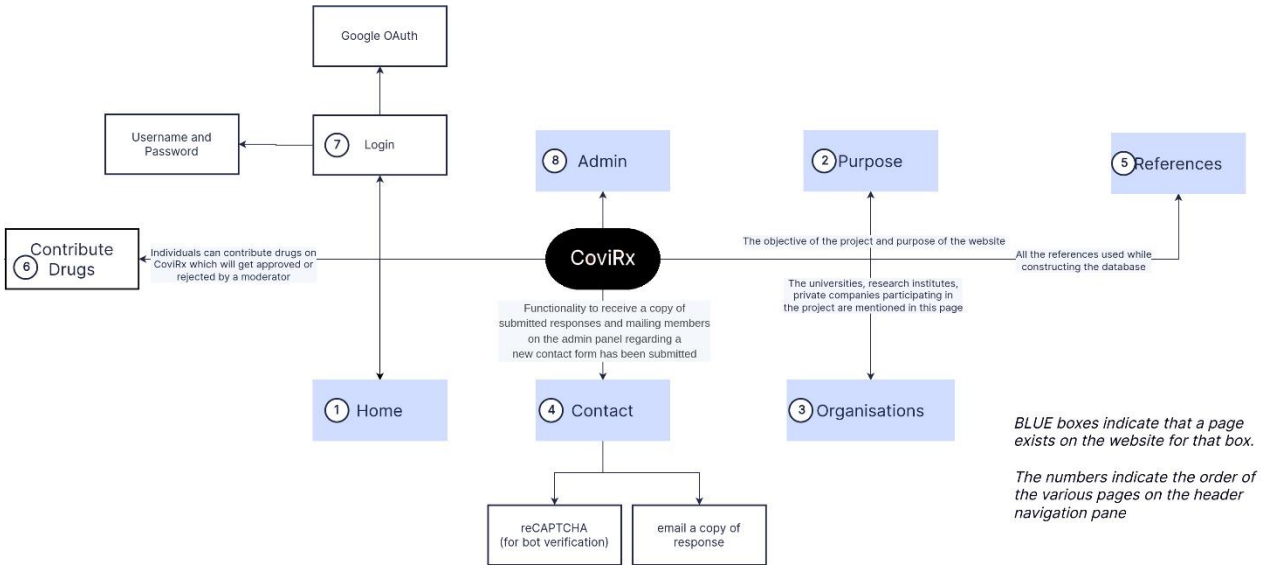

Figure 5: Overview of CoviRx architecture

The CoviRx web app's top hierarchy consists of the admin panel [39] **and the user interface.** It was a design decision to keep the user interface interactive and straightforward. Users can interact with our tool using features like the search engine, drug overviews, contact forms, drug submission forms, etc. Features of the website home page are depicted in Figure 6.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2022 doi:10.20944/preprints202209.0323.v1

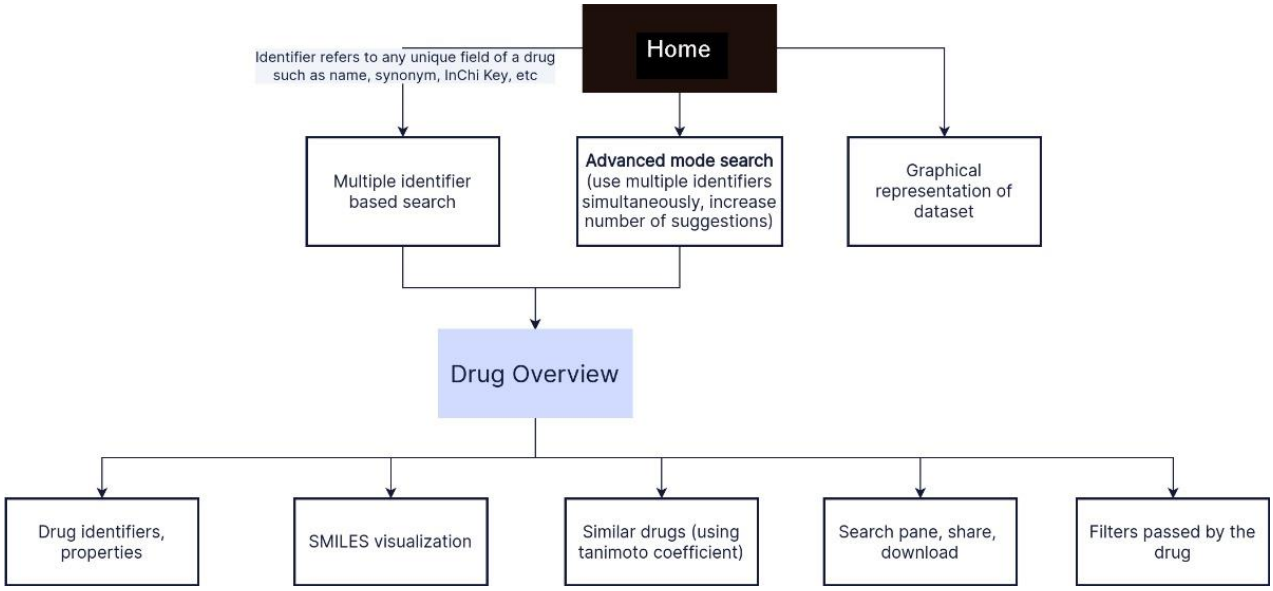

Figure 6: Home page overview

The admin panel is for users with admin rights for the web app. It has two sections, accounts (associated with the users) and main. Admins can use the main section to contribute/upload drugs access website analytics data and custom fields. A feature to invite other users to the admin panel has been added, and Admins can use it to send out invites with a seven-day expiry. Features of the admin panel are depicted in Figure 7.

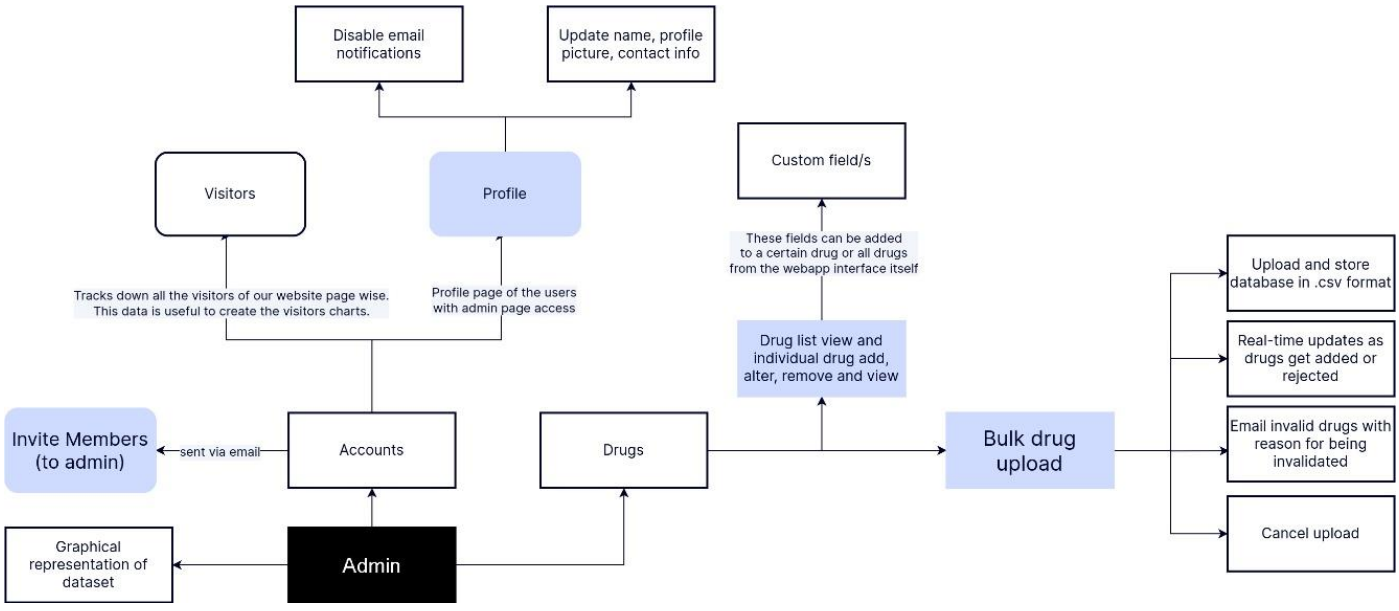

Figure 7: Admin panel overview

Graphical representations have been used throughout the website to give the users a visual representation of the database, enhancing the user experience. CoviRx offers multi-language support [40] so that our open-source dataset is accessible to a broader range of users [41].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2022 doi:10.20944/preprints202209.0323.v1

3.2.1 User Interface

1. *The search engine*

The standard and advanced search modes are two modes of usage, as depicted by Figures 8 and 9, respectively. Both the modes offer search-as-you-type [42] functionality. The significant difference between the modes is that only one identifier can be used to search for a drug in the case of normal mode. In advanced search, multiple identifiers can be used simultaneously. The search results have been capped at five for the normal mode to reduce search time, and this limit is adjustable in the advanced search mode. The search results comprise drug metadata, i.e., drug synonyms, indication class, Simplified molecular-input line-entry system (SMILES) [43] representation, etc.

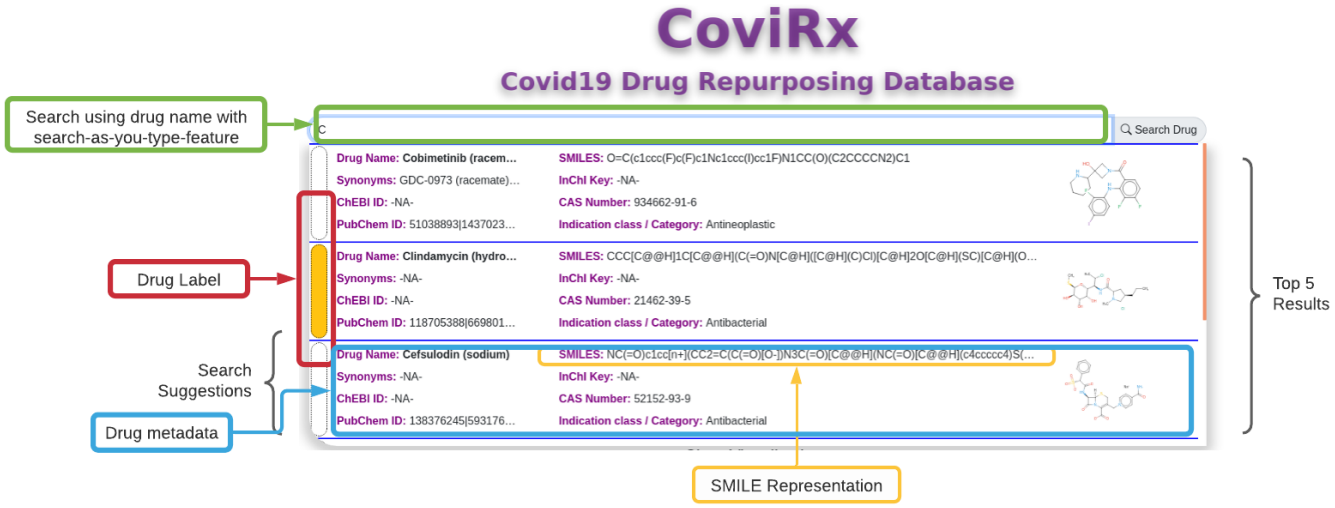

Figure 8:

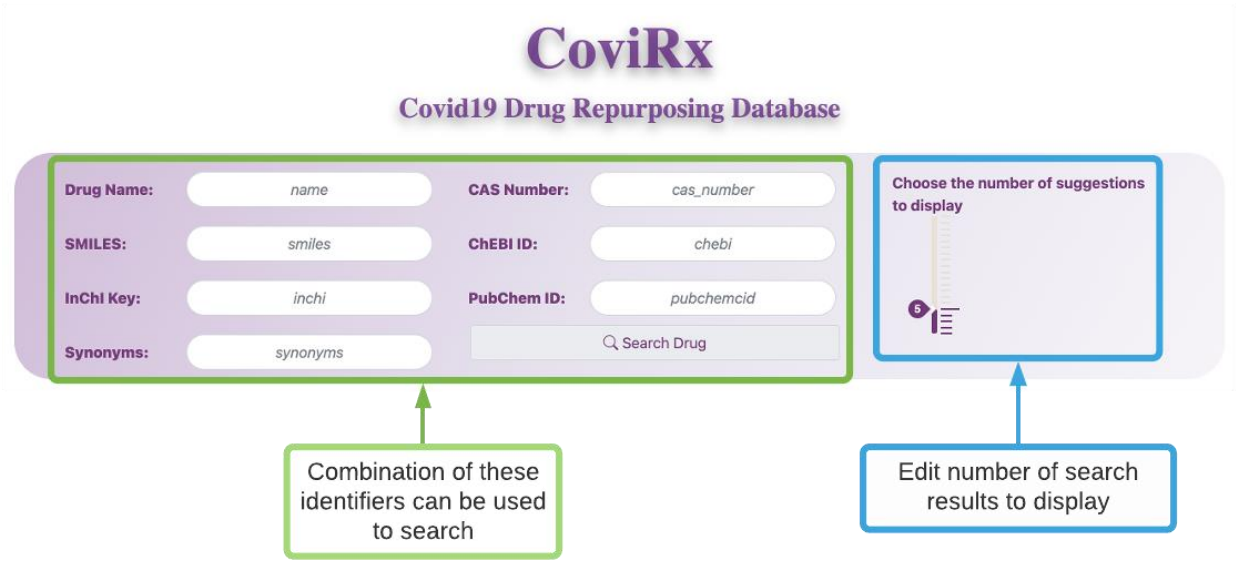

Figure 9: Advanced search mode

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2022 doi:10.20944/preprints202209.0323.v1

2. *Drug overview page*

This section of the website displays information about the query drug. The Data has been divided into distinct categories like identifiers, drug-likeness, original indication, target models, clinical trials, red flags, filters passed, etc. The structure of the webpage is depicted in Figure 10. Specific data points related to drug activity on the original target have not been included. They do not have well-defined values or data uniformity but can be added in the future. The data points such as original indications and red flags have been added to aid researchers. In contrast, other fields like pharmacokinetics (PK) would be helpful to individuals who wish to model their filters and down select drugs. Visual representation of the filters used to down select drugs and the filters passed by each drug has been added for those who would like to investigate it further. Drugs are ranked based on activity scores calculated using the assay data. More details regarding this can be found in MacRaild et al.,2022 (submitted as an interlinked manuscript to the *International Journal of Molecular Sciences*).

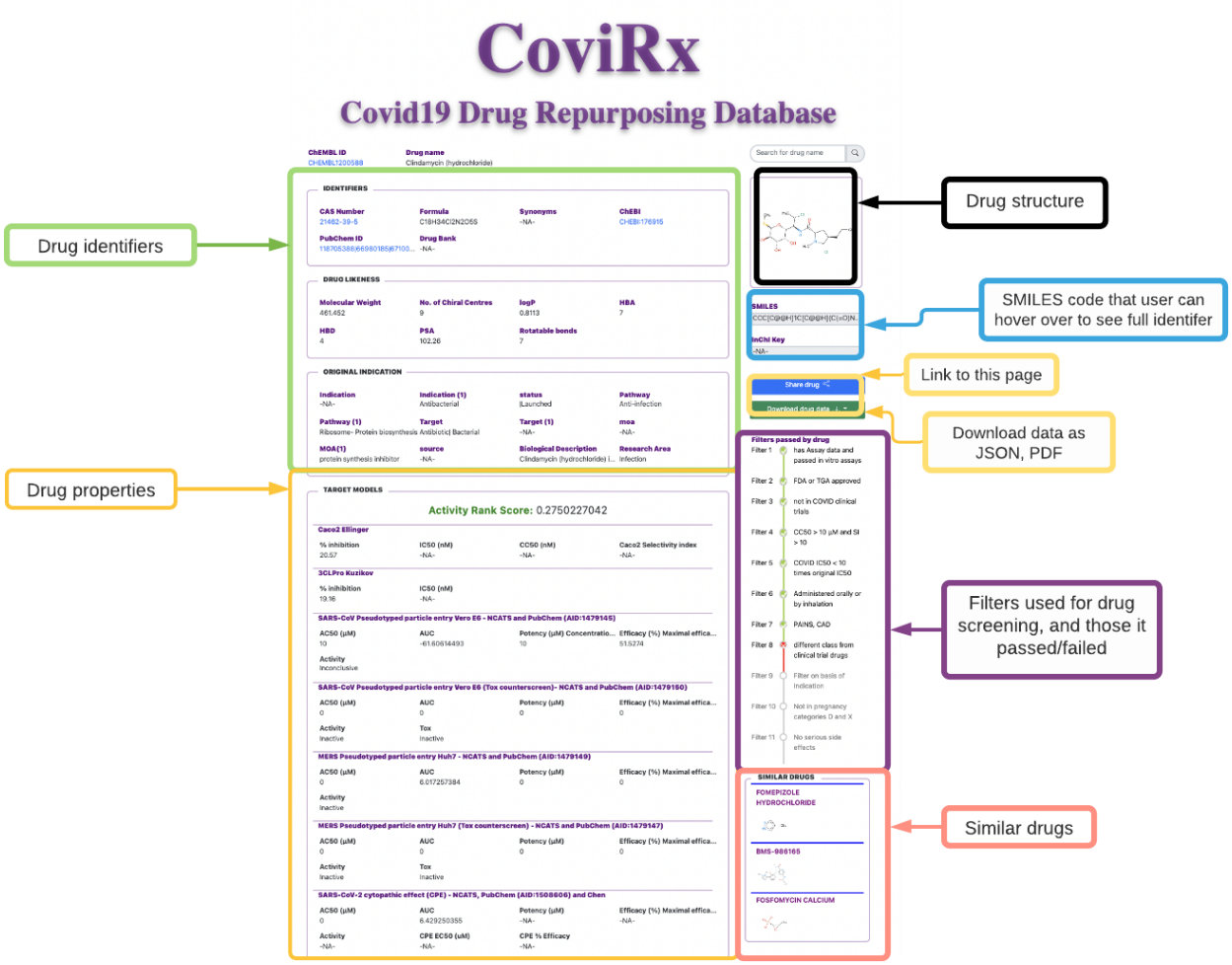

Figure 10: Drug overview webpage

Another exciting feature of this page is the dynamic calculation and display of similar drug candidates. Drug similarity is calculated using the Tanimoto coefficient [10,44]**.** Tanimoto similarity between two drugs is the ratio of common fingerprints and the number of total fingerprints. The RDKit python library [45] was used to generate the fingerprints of the drugs and calculate the Tanimoto similarity between them [46]. The threshold for similarity is kept at a value of 0.7, less than the suggested value of 0.85 [47], to increase the number of similar drugs. Once the similarity is

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2022 doi:10.20944/preprints202209.0323.v1

calculated, the page displays the top 5 (i.e., highest five Tanimoto coefficients) candidates.

3. *Contribute Drug*

CoviRx supports and encourages the community to help us enlarge this open-source database. Individuals can contribute to the web app using the contribute drug feature, an overview of which is depicted in Figure 11. However, only registered people can make these contributions. This additional authentication step was added to ensure data integrity and uniformity. All contributed drugs would be verified manually by our team to ensure data authenticity and reliability.

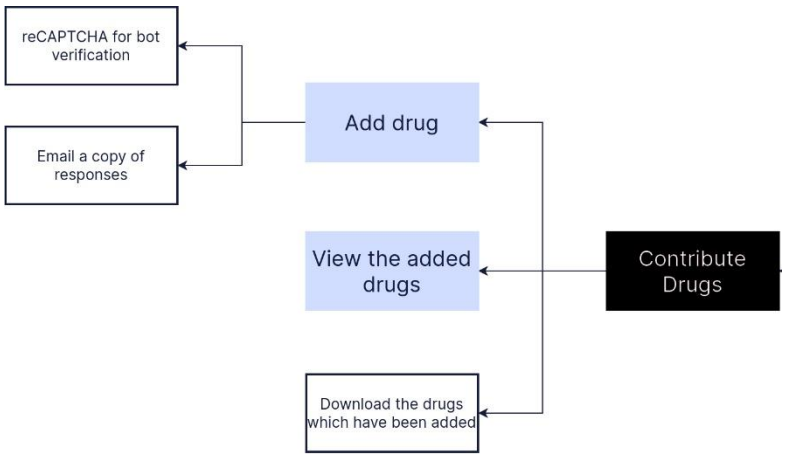

Figure 11: Contribute Drugs Feature

4. *Contact*

Using the contact page, users can contact our team and give us valuable feedback, request access to contribute drugs, etc. A reCAPTCHA [27,48] check has been added to prevent robots from submitting the form.

3.2.1 *Admin Panel*

1. *Bulk upload support*

To speed up the process of drug upload and make the upload process hassle-free, CoviRx has a feature called Bulk Drug upload. It enables the non-IT admin users to upload substantial amounts of data in one go by directly submitting a CSV file. A single drug upload/delete/modify functionality is also available. CoviRx evaluates every drug against multiple constraints before adding it to the primary database. These constraints ensure the validity of the drugs, and drugs that fail these checks, the invalidated drugs, are shared with the drug selection committee to check for the correctness of the data. Figure 12 depicts the various sub-features of the bulk upload feature.

2. *Website Analytics*

Mechanisms have been developed to measure the website traffic, the details of which are available in the admin panel. According to the website load, this was added to plan the resources for hosting the web application. It will also help track the trends about the most popular drug in the database. While monitoring the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2022 doi:10.20944/preprints202209.0323.v1

website traffic, only a session ID cookie [49] is used, and no other details like IP address are saved to ensure users' privacy [50].

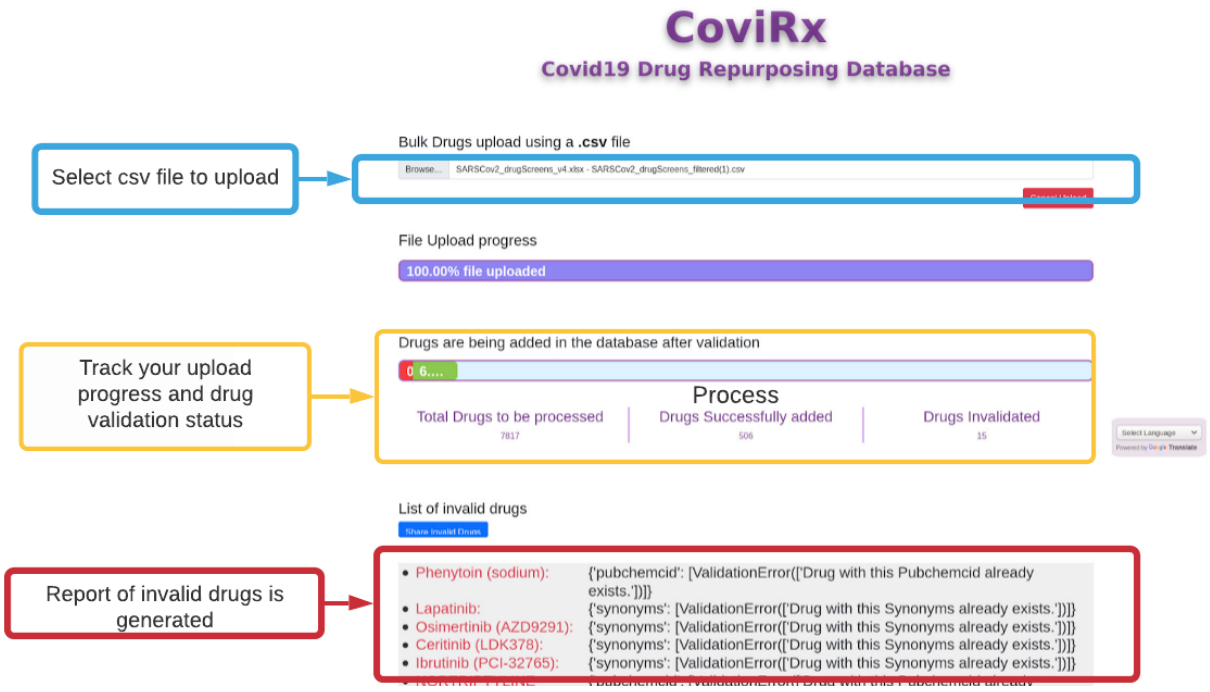

Figure 12: Bulk upload webpage

3. *Custom Fields*

Every data field of a drug displayed on the CoviRx platform maps to an attribute in the database. This implies that the code needs to be altered if a field is added, deleted, or modified for a particular drug. Custom fields have been added to overcome this problem of repeatedly changing the code. Custom Fields feature is one of the critical components in the admin panel. It is a JSON field in the database that stores key-value pairs with the key holding the name of the drug field and the value holding the value for that field [51]**. This feature would make** CoviRx extensible and flexible in the long run.

4. User Notes

The supplementary information provides additional information to the users. The external libraries and external APIs used are listed in Tables S1 and S2, respectively. A few more screenshots for the User Interface, Admin panel and the formulas used are provided in S3, S4, S5, respectively. Additional information for contributing drugs and privacy policy of CoviRx is covered in S6; S7 includes a link to the GitHub repository for CoviRx; S8 links to more detailed technical documentation, and S10 provides information on reusability for developers. All the API endpoints provided by CoviRx are covered in detail, with example in S9.

5. Conclusions

CoviRx provides users with SARS-CoV-2 drug repurposing data from several COVID-19 drug repurposing studies and provides the status of such drugs in COVID-19 trials. Its user-friendly interface and multi-language support improve data accessibility. Various drug-likeness properties like molecular weight, logP, hydrogen bond donors

 doi:10.20944/preprints202209.0323.v1

(HBD), hydrogen bond acceptors (HBA), among others, are included for each compound to assist with early drug discovery programs. The users who wish to design their filters for drug down selection can benefit from the website's pharmacokinetics (PK) section, which provides information regarding a compound's route of administration, the volume of distribution, clearance, protein binding, etc. This platform also encourages its users to help expand this open-source dataset by submitting their drug repurposing data. CoviRx has a wide variety of applications and thus will help advance the field of drug repurposing.

 doi:10.20944/preprints202209.0323.v1

Supplementary Materials: The supporting information can be downloaded at: www.mdpi.com/ TO BE PROVIDED

Author Contributions:

Conceptualization, H.A.J., A. K., M.M.S., R.C., M.A.C., N. L. T., D.J.C., V.A., and S.S.V.; methodology, H.A.J., A.K., M.M.S., M.A.C., C.M.J.K., R.M.S., E.A.P., V.A., and S.S.V.; software, H.A.J. and V.A.; validation, H.A.J., C.B., F., M-U-R.M., A.V.K., C.A.M., I.K.S., E.A.P., N.L.T., D.J.C., and V.A.; formal analysis, H.A.J., A.V.K., E.A.P., and V.A.; investigation, H.A.J. and E.A.P.; resources, S.M., V.A., and S.S.V.; data curation, S.M., C.A.M., I.K.S., A.L.P., N.L.T., D.J.C., and V.A.; writing—original draft preparation, H.A.J., C.B., S.M., V.A., S.S.V.; writing—review and editing, All authors.; visualization, H.A.J. and V.A.; supervision, A.K., D.J.C., V.A., and S.S.V.; project administration, S.M., I.K.S., V.A., S.S.V.; funding acquisition, S.S.V.

All authors have read and agreed to the published version of the manuscript.

Funding: This research was kindly funded (Principal Investigator: S.S.V.) by the Australian Department of Health through its Medical Research Future Fund (MRFF) via the National Health and Medical Research Grant Number: MRF2009092). Additional support was provided by the United States Food and Drug Administration Medical Countermeasures Initiative (75F40121C00144). The article reflects the views of the authors and does not represent the views or policies of the funding agencies, including the FDA.

Institutional Review Board Statement:

Not applicable

Informed Consent Statement:

Not applicable

Data Availability Statement:

Any underlying data not presented can be provided by the corresponding author upon reasonable request.

Acknowledgments:

The authors would like to acknowledge the contributions made to the project by the broader sySTEMs Initiative: Aditya V Vashi, Alexander J McAuley, Carmel M O'Brien, Carol Lee, Darcie Cooper, David Beale, Eugene Athan, Elizabeth Vincan, James Hudson, Jenny O'Connell, Jian-Wei Liu, John Noel Viana, Kanta Subbarao, Kathie Burkett, Kim R Blasdell, Laurence Wilson, Lee Trinidad, Mary Tachedjian, Matthew P Bruce, Michael Kuiper, Nagendrakumar Balasubramanian Singanallur, Nathan J Gödde, Petrus Jansen van Vuren, Sarah Goldie, Shane Riddell, Simone Clayton, Simran Chahal, Stephanie Keating, Trevor Drew and Vinod Sundaramoorthy

Conflicts of Interest:

The authors declare no conflict of interest

References

1. Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine* **2020**, *382*, 727–733. 10.1056/NEJMOA2001017/SUPPL_FILE/NEJMOA2001017_DISCLOSURES.PDF.
2. Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* **2020**, *395*, 497–506. 10.1016/S0140-6736(20)30183-5/ATTACHMENT/D5332CA1-83D8-4C4C-BC57-00A390BF0396/MMC1.PDF.
3. Dong, E.; Du, H.; Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* **2020**, *20*, 533–534. 10.1016/S1473-3099(20)30120-1/ATTACHMENT/9DE52FFD-61D4-4C3F-826F-841D978D469D/MMC1.PDF.
4. Wang, X.; Guan, Y. COVID-19 drug repurposing: A review of computational screening methods, clinical trials, and protein interaction assays. *Medicinal Research Reviews* **2021**, *41*, 5–28. 10.1002/MED.21728.
5. Pushpakom, S.; Iorio, F.; Eyers, P.A.; Escott, K.J.; Hopper, S.; Wells, A.; Doig, A.; Guilliams, T.; Latimer, J.; McNamee, C.; et al. Drug repurposing: progress, challenges and recommendations. *Nature Reviews Drug Discovery 2018 18:1* **2018**, *18*, 41–58. 10.1038/nrd.2018.168.
6. Corsello, S.M.; Bittker, J.A.; Liu, Z.; Gould, J.; McCarren, P.; Hirschman, J.E.; Johnston, S.E.; Vrcic, A.; Wong, B.; Khan, M.; et al. The Drug Repurposing Hub: a next-generation drug library and information resource. *Nature Medicine 2017 23:4* **2017**, *23*, 405–408. 10.1038/nm.4306.
7. Excelra | Covid-19-Drug-Repurposing-Database Available online: https://www.excelra.com/covid-19-drug-repurposing-database/ (accessed on Jan 31, 2022).
8. Brimacombe, K.R.; Zhao, T.; Eastman, R.T.; Hu, X.; Wang, K.; Backus, M.; Baljinnyam, B.; Chen, C.Z.; Chen, L.; Eicher, T.; et al. An OpenData portal to share COVID-19 drug repurposing data in real time. 10.1101/2020.06.04.135046.
9. Pawar, A.Y. Combating devastating COVID-19 by drug repurposing. *International Journal of Antimicrobial Agents* **2020**, *56*, 105984. 10.1016/J.IJANTIMICAG.2020.105984.
10. Tanimoto, T.T. Elementary mathematical theory of classification and prediction. **1958**.
11. Talevi, A.; Bellera, C.L. Challenges and opportunities with drug repurposing: finding strategies to find alternative uses of therapeutics. *https://doi.org/10.1080/17460441.2020.1704729* **2019**, *15*, 397–401. 10.1080/17460441.2020.1704729.
12. Django (Version 3.2) Available online: https://www.djangoproject.com/ (accessed on Jan 26, 2022).
13. Holovaty, A.; Kaplan-Moss, J. *The definitive guide to Django: Web development done right*; Apress, 2009;
14. jQuery 3.6.0 Released! | Official jQuery Blog Available online: https://blog.jquery.com/2021/03/02/jquery-3-6-0-released/ (accessed on Jan 26, 2022).
15. Bootstrap 5: Documentation Available online: https://getbootstrap.com/docs/5.0/getting-started/introduction/ (accessed on Jan 26, 2022).
16. Charts | Google Developers Available online: https://developers.google.com/chart (accessed on Feb 7, 2022).
17. Drugs@FDA: FDA-Approved Drugs Available online: https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm (accessed on Jan 26, 2022).
18. TGA Search -- Available online: https://tga-search.clients.funnelback.com/s/search.html?query=&collection=tga-artg (accessed on Jan 26, 2022).
19. Inxight Drugs Available online: https://drugs.ncats.io/ (accessed on Jan 26, 2022).
20. Home - ClinicalTrials.gov Available online: https://clinicaltrials.gov/ (accessed on Jan 26, 2022).

21. Fielding, R.T.; Kaiser, G. Collaborative work: The apache http server project. *IEEE Internet Computing* **1997**, *1*, 88–90. 10.1109/4236.612229.
22. Owens, M. *The definitive guide to SQLite*; Apress, 2006;
23. Group, T.P.G.D. *Documentation PostgreSQL 10.3*; Group, T.P.G.D., Ed.; 2018;
24. Douglas, K.; Douglas, S. *PostgreSQL: a comprehensive guide to building, programming, and administering PostgresSQL databases*; SAMS publishing, 2003;
25. SQLite vs MySQL vs PostgreSQL: A Comparison Of Relational Database Management Systems | DigitalOcean Available online: https://www.digitalocean.com/community/tutorials/sqlite-vs-mysql-vs-postgresql-a-comparison-of-relational-database-management-systems (accessed on Jan 26, 2022).
26. Using OAuth 2.0 to Access Google APIs Available online: https://developers.google.com/identity/protocols/oauth2 (accessed on Jan 26, 2022).
27. reCAPTCHA Available online: https://www.google.com/recaptcha/about/ (accessed on Jan 26, 2022).
28. IMAP, POP, and SMTP | Gmail IMAP | Google Developers Available online: https://developers.google.com/gmail/imap/imap-smtp (accessed on Jan 26, 2022).
29. GitHub - peterbe/premailer: Turns CSS blocks into style attributes Available online: https://github.com/peterbe/premailer (accessed on Jan 26, 2022).
30. Ramakrishnan, R.; Gehrke, J.; Gehrke, J. *Database management systems*; McGraw-Hill New York, 2003; Vol. 3;.
31. Silberschatz A; Korth H F; Sudarshan S *Database System Concepts*; 7th ed.; 2019;
32. Elmasri, R.; Navathe, S.B.; Elmasri, R.; Navathe, S.B. *Fundamentals of Database Systems*; Springer, 2000;
33. Django Template language Available online: https://docs.djangoproject.com/en/3.2/topics/templates/ (accessed on Jan 26, 2022).
34. celery · PyPI Available online: https://pypi.org/project/celery/ (accessed on Feb 7, 2022).
35. Larman, C.; Basili, V.R. Iterative and incremental development: A brief history. *Computer* **2003**, *36*, 47–56. 10.1109/MC.2003.1204375.
36. Becker, S.A.; Berkemeyer, A. Rapid application design and testing of Web usability. *IEEE Multimedia* **2002**, *9*, 38–46. 10.1109/MMUL.2002.1041947.
37. Weyuker, E.J. Testing component-based software: A cautionary tale. *IEEE Software* **1998**, *15*, 54–59. 10.1109/52.714817.
38. Pezze, M.; Young, M. *Software testing and analysis: process, principles, and techniques*; John Wiley & Sons, 2008;
39. django-admin-interface · PyPI Available online: https://pypi.org/project/django-admin-interface/ (accessed on Feb 7, 2022).
40. Cloud Translation documentation | Google Cloud Available online: https://cloud.google.com/translate/docs (accessed on Jan 26, 2022).
41. Melitz, J.; Toubal, F. Native language, spoken language, translation and trade. *Journal of International Economics* **2014**, *93*, 351–363. 10.1016/J.JINTECO.2014.04.004.
42. Ji, S.; Li, G.; Li, C.; Feng, J. Efficient interactive fuzzy keyword search. *WWW'09 - Proceedings of the 18th International World Wide Web Conference* **2009**, 371–380. 10.1145/1526709.1526760.
43. Weininger, D. SMILES, a Chemical Language and Information System: 1: Introduction to Methodology and Encoding Rules. *Journal of Chemical Information and Computer Sciences* **1988**, *28*, 31–36. 10.1021/CI00057A005.
44. Bajusz, D.; Rácz, A.; Héberger, K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *Journal of Cheminformatics* **2015**, *7*, 1–13. 10.1186/S13321-015-0069-3/FIGURES/7.
45. RDKit Available online: https://www.rdkit.org/ (accessed on Jan 26, 2022).
46. Willett, P.; Barnard, J.M.; Downs, G.M. Chemical similarity searching. *Journal of chemical information and computer sciences* **1998**, *38*, 983–996.

 doi:10.20944/preprints202209.0323.v1

47. Two similar compounds with a low (smaller than 0.85) Tanimoto coefficient? Available online: https://www.researchgate.net/post/Two_similar_compounds_with_a_low_smaller_than_085_Tanimoto_coefficient2 (accessed on Jan 26, 2022).
48. Zhang, Y.; Gao, H.; Pei, G.; Luo, S.; Chang, G.; Cheng, N. A survey of research on CAPTCHA designing and breaking techniques. *Proceedings - 2019 18th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/13th IEEE International Conference on Big Data Science and Engineering, TrustCom/BigDataSE 2019* **2019**, 75–84. 10.1109/TRUSTCOM/BIGDATASE.2019.00020.
49. Lacroix, K.; Loo, Y.L.; Choi, Y.B. Cookies and Sessions: A Study of What They Are, How They Work and How They Can Be Stolen. *Proceedings - 2017 International Conference on Software Security and Assurance, ICSSA 2017* **2018**, 20–24. 10.1109/ICSSA.2017.9.
50. Chung, W.; Paynter, J. Privacy issues on the Internet. *Proceedings of the Annual Hawaii International Conference on System Sciences* **2002**, *2002-January*, 9. 10.1109/HICSS.2002.994191.
51. django-flat-json-widget · PyPI Available online: https://pypi.org/project/django-flat-json-widget/ (accessed on Feb 7, 2022).