COVID-19 School Dashboard Integrated Dataset and Data Visualization

A Data Management Plan created using DMP Assistant

Creator: Kelly Bairos

Principal Investigator: Prachi Srivastava

Data Manager: Nathan T.T. Lau

Project Administrator: Kelly Bairos

Affiliation: Western University

Funder: Digital Research Alliance of Canada (grant for data archiving and management)

Template: Alliance Template for Interdisciplinary Health Software/Technology Development

ORCID iD: 0000-0003-4865-8963

Project abstract:

The COVID-19 School Dashboard, an open-source, open-access site, reports and maps official public data on confirmed school-related COVID-19 cases in publicly funded elementary and secondary schools in Ontario, and connects these data to a public dataset on school-level demographic characteristics. Connecting previously disparate datasets produces a unique COVID-19 school-level dataset to generate equity-related analysis prompting better understanding of the locations of individual affected schools, geographic concentrations of school outbreak clusters, and social background characteristics of affected schools. Furthermore, schools with outbreaks are visualized, geocoded, and mapped on the dashboard website to provide accurate geospatial representation of outbreak locations.

Identifier: 8579

Last modified: 18-10-2022

Grant number / URL: https://alliancecan.ca/en

COVID-19 School Dashboard Integrated Dataset and Data Visualization - Software/Technology Management Plan

Software/Technology Development

What software/technology will be created in this study?

The COVID-19 School Dashboard is an open-source, open-access site that reports and maps official public data on confirmed school-related COVID-19 cases in publicly funded elementary and secondary schools in Ontario. This site connects these data to official public data on school administrative and school demographic characteristics. Connecting and visualizing otherwise disparate yet relevant datasets allows for better understanding of the locations of affected schools, the geographic concentrations of school outbreak clusters, and the social background characteristics of the populations of affected schools.

It reports infection data for the 2020-21 and 2021-22 school years as were available by official Ontario of Ministry of Education data sources. Data were available from September 2020 until December 2022. Reporting was suspended by the Ministry in January 2022.

The website is intended to provide: (a) the public, educators, researchers, and policy-makers an up-to-date overview of the impact of COVID-19 on schools; and (b) researchers and analysts a platform to analyze whether COVID-19 may affect schools differently in view of socio-economic factors. This is important in understanding potential equity effects.

What software/technology development framework or model will be used, if any?

Due to the fast-changing nature of the COVID-19 pandemic, it was not possible to foresee the tools/features that the site would require to visualize school-based infections. As such, a rapid prototyping framework was taken. Once the website could visualize school-based infections in a basic capacity, additional features/changes were undertaken by: (a) identifying the need, (b) coding the feature, (c) reviewing performance of the new feature.

Will you utilize any existing code to develop this software/technology?

The COVID-19 School Dashboard was initially developed in Shiny Apps. It was then migrated to a Google Cloud hosting platform. The site was developed by forking from the open source code for The COVID-19 Tracker https://vac-lshtm.shinyapps.io/ncov_tracker/ (Parker & Leclerc, 2020) which can be found here https://github.com/eparker12/nCoV_tracker.

Parker, E., & Leclerc, Q. (2020). *COVID-19 tracker.* [Web application]. https://vac-lshtm.shinyapps.io/ncov_tracker/

What test cases will you use to develop the software/technology?

A rapid prototyping framework was taken to beta test measures. Features and changes were undertaken by: (a) identifying the need, (b) coding the feature, and (c) reviewing performance of the new feature.

We implemented automatic test cases with the Google Cloud hosting platform to verify that any new changes to the site would dockerize correctly before going live. Having this protected branch on GitHub was to prevent errors from being uploaded and to ensure that the site would have the highest amount of uptime.

How will your software/technology and documentation adhere to disability and/or accessibility guidelines?

The software was checked using the WAVE Tool, an online software which checks websites for accessibility issues including contrast checks, screen reader issues, and more. Any issues found were fixed to ensure the widest access to the tool as possible.

What dependencies will be used in the development of this software/technology?

The following dependencies were used in development:

- R Packages including: shiny, shinydashboard, ggplot2, DT, reshape2, rgdal, shinythemes, sp, plotly, xts, data.table, ggmap, httr, leaflet, lubridate, readxl, Rcolorbrewer, rgeos, rvest, stringr, stringdist, git2r, Rcpp
- The development of this application relies on web hosting. It was first hosted using ShinyApps hosting and later migrated to Google Cloud hosting
- The version control software that will track all of the changes that are made and committed to the code will be GitLab and GitHub

Software/Technology Documentation

What information would be required for someone to understand and reuse your software/technology?

All information required for someone to understand and reuse the software is documented both in the archived code on Borealis Dataverse as well as on the formal documentation page on GitHub.
Borealis Dataverse here: https://doi.org/10.5683/SP3/Z9SNP0
GitHub here: https://github.com/connor-cozens/covid19-school-dashboard

What documentation will you develop to help others write and run tests on your software/technology?

Documented code is provided. See above for links.

How will you track changes to code and dependencies?

GitHub was used to track changes to the code. See above for links.

Software/Technology Preservation

How will the software/technology be updated and maintained over time?

The software includes a data downloader which checks the official government data. This was programmed to run each morning to capture the new daily numbers of school-related COVID-19 infections to update automatically on the site to coincide with the Ministry of Education release. This allowed for minimal maintenance on the site until the Ontario Ministry of Education stopped releasing new data (as of January 2022). After new data stopped being updated, we were able to focus on releasing new features and views including a time-slider to go back and view historical figures and visuals over the COVID-19 pandemic.

Describe the level of risk associated with the use of public web services/infrastructure/databases regarding their stability and sustainability.

No encryption was needed, as the website uses open-source de-identified data on school infections. The risks involved with cloud hosting are website downtime and lack of site access to users. We have put measures in place to mitigate these risks through the testing and coding processes.

Software/Technology Ethical and Legal Restrictions

Who will own the copyright to the software/technology?

Code attribution to Justin Marshall, Connor Cozens, Peter Taylor, and Prachi Srivastava.

What software/technology license will you choose?

Attribution 4.0 International (CC BY 4.0) will be used.

Software/Technology Responsible Parties

Who will have access to your software/technology throughout the project? Describe each collaborator's responsibilities in relation to having access to the data.

All collaborators had access to the raw data as they are publicly available. The project team developers had back-end access to the site for development purposes. This was a live site for public consumption. As such, the general public had access to the front-facing site once the beta testing phase was complete and upon external launch. Upon completion, the integrated dataset (further information the data management plan below) and website code are accessible via open access on GitHub and Borealis Dataverse.
Borealis Dataverse for datasets here: https://doi.org/10.5683/SP3/D0QXGQ

Who is responsible for reviewing and accepting each software/technology release?

A code review process was implemented on GitHub that required the developers Connor Cozens and Justin Marshall to review and accept changes as they were developed before being added to the website.

Software/Technology Sharing

Who are the intended users of your software/technology?

The intended users are researchers, educators, policy-makers and policy analysts, parents, and the general public.

What software/technology will be shared at the end of the study?

Code for the website is shared on GitHub and on Borealis Dataverse. See above for links.

Are there restrictions on how you can share your software/technology related to patents, copyright, or intellectual property?

N/A

Where will you share your software/technology?

Archived code for the software is stored on Borealis Dataverse and the original code and any updates are stored and maintained on GitHub. See above for links.