The LAISDAR project – hospital EHR harmonization in Rwanda through mapping to OMOP CDM; outcome, challenges and lessons learned

Lars Halvorsen[a], Jared Houghtaling[a], Emma Gesquiere[a], Lore Vermeylen[a], Benjamin Burke[a], Charles Ruranga[b], Marc Twagirumukiza[c], Gilbert Rukundo[d], Clarisse Musanabaganwa[d], Claude Mambo Muvunyi[d], Jean Claude Semuto Ngabonziza[d], Aurore Nishimwe[e]

[a] edenceHealth NV
[b] Center of Excellence in Data Sciences, University of Rwanda (UR), Kigali, Rwanda
[c] Faculty of medicine and Health sciences, Ghent University, Ghent, Belgium
[d] Rwanda Biomedical Center (RBC), Ministry of Health (MoH), Kigali, Rwanda
[e] College of medicine and health sciences (CMHS), University of Rwanda (UR), Kigali, Rwanda

Background

In response to the COVID-19 pandemic, a federated data network (FDN) of 15 hospitals was established in Rwanda; "Leveraging Artificial Intelligence and Data Science Techniques in Harmonizing, Sharing, Accessing and Analysing SARS-COV-2/COVID-19 Data in Rwanda (LAISDAR)" **[1, 2]**. The LAISDAR project combines EHR data and national COVID-19 testing and survey metrics. This project includes multiple Rwandan and Belgian institutions, is managed by the University of Rwanda (UR) and received funding from Canada's International Development Research Centre (IDRC) **[3]** as part of the Global South AI4COVID program **[4]**.

The project objective is to leverage the federated hospital data sets, extended with data from centralized COVID-19 test results and survey data, to support Rwandan government needs in monitoring and predicting the COVID-19 burden, including on hospital admissions and overall infection rates. The impact of various public health measures on the pandemic evolution, social-economic situation, and mental health are also key study objectives. Although the project was originally focused on COVID-19 research, the possible research topics have since widened to other disease areas.

Figure 1 Participating sites in LAISDAR network

Methods

Two different EHR systems are used across Rwanda, openClinic GA **[5]** and openMRS **[6]**, for which logic to transform to OMOP CDM was defined and implemented. An ETL was designed which can transform both source systems to the target format, and incorporate data from the national COVID-19 testing and COVID-19 survey results as part of the ETL process. Ares was installed on the central server for quality control across the network **[7]**.

Each hospital node was set up on a dedicated Mac Mini with the ETL (Extract-Transform-Load) script, OHDSI tooling and other supporting services installed. The OHDSI tools installed are Atlas, Achilles, DQD, AresIndexer, and Arachne node/execution engine. In addition R is installed, which is also used for report generation.

On a central server, OMOP CDM versions of the national COVID-19 testing and survey results are hosted and made available to the hospital ETLs through a secure access point. Atlas is installed to provide access to these data sources, as well as the data profiles of all the hospital node OMOP CDM instances. For a network-wide view on quality, mapping and aggregated statistics, the Ares web application is installed on the central server. Finally, Arachne Central is installed on the central server, as well as R.

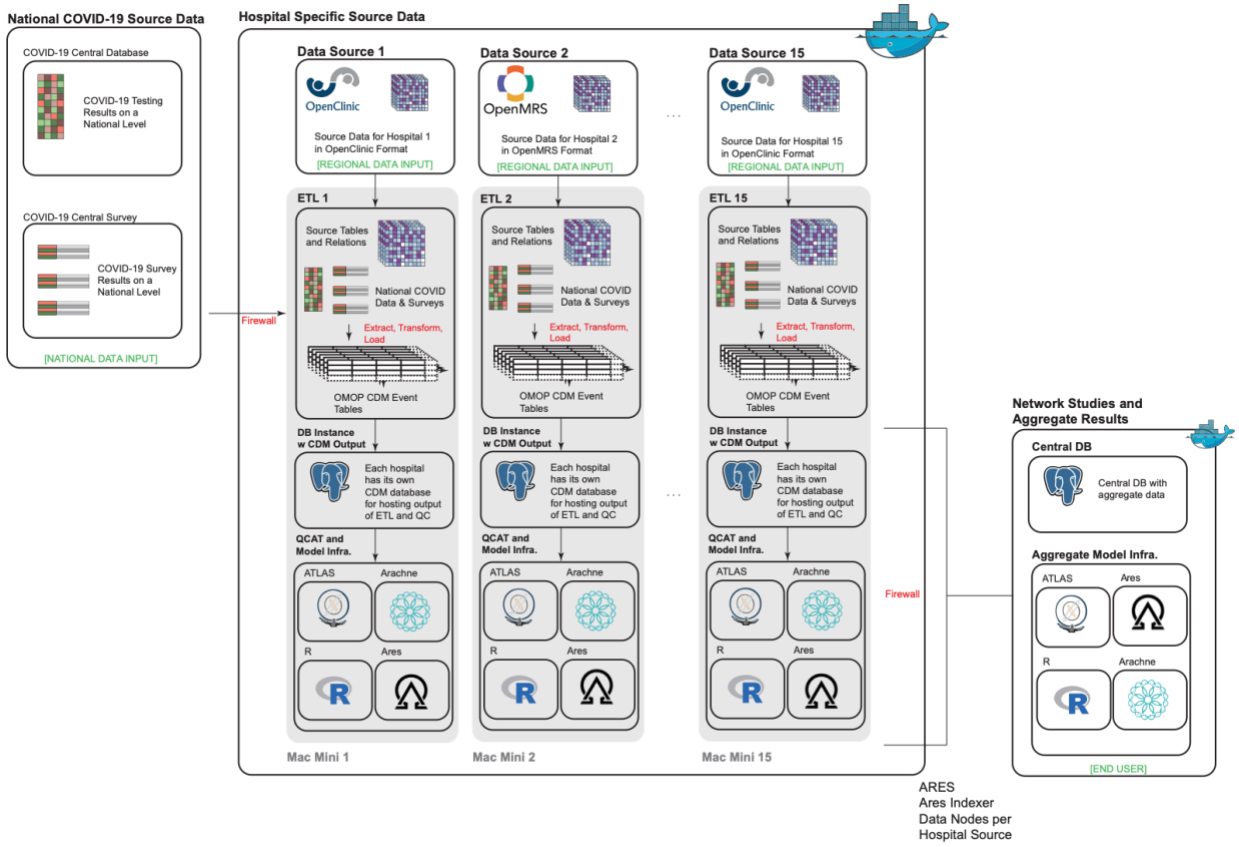

Figure 2 The LAISDAR overall architecture

Results

As of April 2023, the ETL to transform the hospital EHR data to OMOP CDM has been run at least once at 14 of the hospitals. The deployment of Ares has allowed a centralized view on domain and mapping coverage across the network, which has aided the planning of the next steps for concept mapping and ETL improvement. Some of the data quality issues encountered were related to inconsistencies with how birth dates were filled in, and gender specific clinical events inconsistent with the patient's gender. Other challenges were related to different configurations of the same EHR system at different sites, necessitating additional logic in the ETL to handle these mapping differences. Finally, the concept mappings did not provide the level of coverage foreseen for some domains, for example for drug mappings. All of the above issues are being further followed up, a task which Ares has been a great tool for facilitating.

In the initial phase of the project, most tasks related to the deployment and setup of the hospital nodes and central server were supported remotely, which was not always an optimal approach. In the last 6 months of the project, onsite visits by edenceHealth and Ghent University personnel helped finalize the node setups, and in the process solved different technical challenges. For example, the deployments were originally set up to pull the different Docker container images when a new version of the ETL or OHDSI tool was made available. This turned out to be an inefficient approach due to limited Internet download speeds and occasional power outages; instead, the deployment approach has now been changed to automated build processes on the Mac Mini nodes themselves, using server instances in Rwanda whenever possible.

The Arachne node and server instances were built and deployed using Docker containers, but have not yet been activated due to some remaining challenges with the execution engine configuration. We chose to prioritize the deployment of Ares at this stage, and will continue with the Arachne deployment at a later stage.

Finally, a proof-of-concept for a reporting solution has been developed, through which the mandatory monthly reports from the hospitals to the Ministry of Health can be, at least partially, automated based on OMOP CDM.

Conclusion

The LAISDAR project is concluding in June 2023, but efforts are underway to continue the work started during this project. The LAISDAR project has accomplished much; 14 hospital nodes with EHR data transformed to OMOP CDM, with a total of about 3,5M patients represented. The national COVID-19 test results have been converted to OMOP CDM, as has the results of a COVID-19 related survey from 2022 that included 10000 participants. A sustainable infrastructure for regular updates of the hospitals' OMOP CDM database instances have been established, with centralized quality assurance and data coverage overviews based on Ares.

Training has been provided on OMOP CDM and Atlas, to enable local researchers and data managers to utilize the FDN for research, modelling, monitoring, and reporting.

The LAISDAR project has also provided important lessons and experiences that will enable further expansion of OMOP CDM in Rwanda and the rest of Africa.

References

[1] Halvorsen, L. et al. “LAISDAR – A federated data network to support COVID-19 research in Rwanda”. *OHDSI 2021 Global Symposium*

[2] Nishimwe, A., Ruranga, C., Musanabaganwa, C. et al. “Leveraging artificial intelligence and data science techniques in harmonizing, sharing, accessing and analyzing SARS-COV-2/COVID-19 data in Rwanda (LAISDAR Project): study design and rationale.” *BMC Medical Informatics and Decision Making 22, 214 (2022).*

[3] International Development Research Centre (IDRC), Canada. https://www.idrc.ca/en

[4] Global South AI4COVID Program. https://covidsouth.ai

[5] OpenMRS: open-source platform to build customized EMR system. https://openmrs.org

[6] OpenClinic GA: open-source integrated hospital information management system. https://sourceforge.net/projects/open-clinic/

[7] Houghtaling, J. et al. “Implementation of the ARES application to monitor network-wide data quality and mapping coverage for 16 unique OMOP sources across Rwanda”. *OHDSI Europe Symposium 2023*