

## Privacy preserving health analytics of patient data: use case

FAIRified patient data has huge potential in terms of analytics. They could, for instance, be used for the plotting of disease (co)-occurrence or medication use. Importantly, it ensures that health analytics are possible, whilst ensuring that privacy is preserved by allowing data visiting through queries, whilst the data itself stays in the health facilities. We tested out analytics of FAIR patient data in Uganda, by converting raw data into a FAIR format, by assigning metadata and converting it into a machine-readable format. This allows querying across the data. In this case, queries were run on the prevalence of diseases and medicine use per location.

**FAIR** stands for **F**indable, **A**ccessible, **I**nteroperable and **R**eusable. By making data FAIR, data reuse is improved, by ensuring that one can easily find and access data, and that the data is interoperable so that it is usable in different places and systems.

One of the primary challenges of this application is limitations in the volume and quality of the data, which patient data being fragmented, the lack of electronic health records, and high aggregation in terms of location and time.

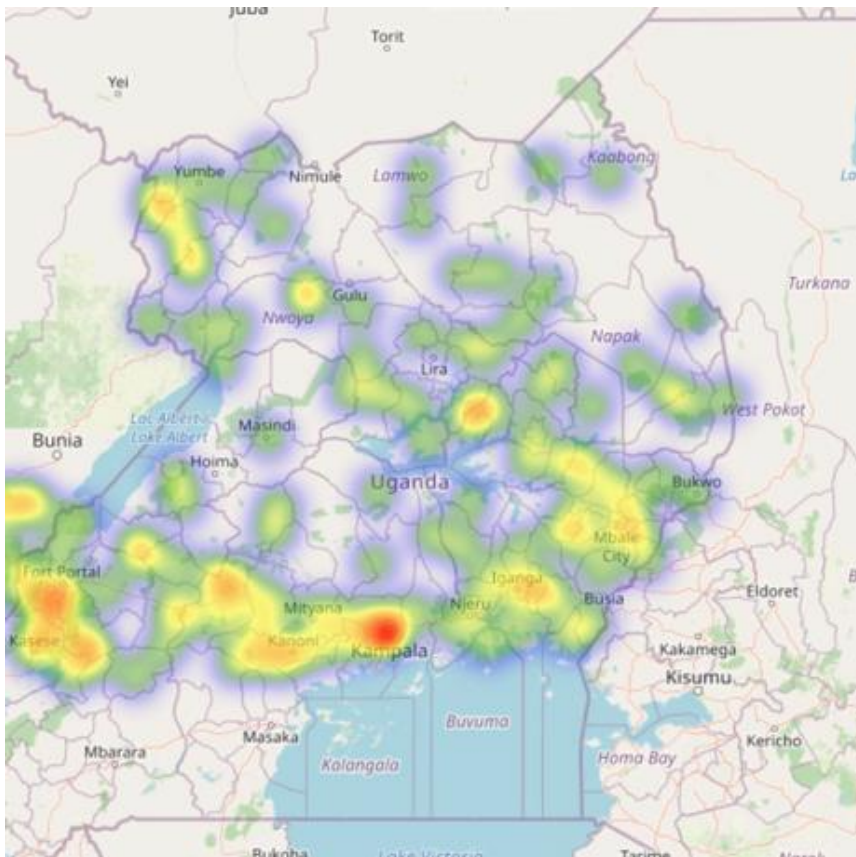


Figure 1. Example of a heat map of HIV infection in Uganda.

## RECOMMENDATIONS

1. Establish a national policy for digitizing patient health records and ensuring that all new health data is collected in electronic formats compatible with the FAIR principles (Findable, Accessible, Interoperable, Reusable).
2. Develop and enforce national health data standards to ensure data interoperability across health facilities, systems, and regions. This should include a common standard for metadata, data formats, and coding systems, such as ICD codes for diseases.
3. Offer financial and technical incentives to healthcare facilities for adopting and maintaining FAIR data practices. This could include grants, technology infrastructure support, or tax incentives for facilities that convert raw data into FAIR-compliant formats.
4. Provide training programs for healthcare workers and data scientists on data management, FAIR principles, and advanced analytics techniques, such as machine learning. This should include support for the collection, maintenance, and analysis of patient data.
5. Deploy advanced data collection tools, such as mobile health applications, to capture real-time, detailed patient data. These tools should allow for capturing more granular data on patient location, disease occurrence, and medication use over time.
6. Establish a national system for monitoring and reporting data quality at health facilities. The system should automatically flag incomplete or inconsistent data, and facilities should be required to submit corrective measures. Foster partnerships between the public health sector, academia, and private technology companies to develop innovative data analytics tools and platforms for querying, visualization, and machine learning on FAIR health data.