

Introduction to the Special Issue: Data Intelligence on Patient Health Records

Mirjam van Reisen^{1†}, Barend Mons²

¹Leiden University Medical Centre (LUMC), Leiden University, 1310 Leiden, the Netherlands

²Leiden University Medical Centre, Poortgebouw N-01, Rijnsburgerweg 10 2333 AA Leiden, the Netherlands

Citation: van Reisen, M., Mons, B.: Introduction to the Special Issue: Data Intelligence on patient health records. *Data Intelligence* 4(4), 671–672 (2022). doi: 10.1162/dint_e_00165

Submitted: March 10, 2021; Revised: June 10, 2022; Accepted: July 15, 2022

Data Intelligence is the ultimate purpose of FAIR data management. FAIR as in data that is Findable, Accessible (under well defined conditions), Interoperable and Reusable. FAIR also as in ethical data; data that fulfils the requirements of Personal Data Protection, is collected for well defined purposes and is held and curated within ownership of the location where the data is produced.

In this Special Issue, we ask how more data intelligence can be derived from Medical Patient Records. All health facilities associated with the reported studies produce Medical Patient Records, and these records are highly structured and employing vocabularies that are internationally understood and used. The researchers of the Virus Outbreak Data Network-Africa were interested to understand whether these data would be a good source for FAIRification and, once FAIRified, would provide a data pipeline that can be the future of data intelligence with a global scope and at a scale that is currently lacking. The programme therefore focused on Africa, a continent with a minimum of legacy in digital health records data intelligence and therefore an excellent place to build a pioneer, novel FAIR system.

The Special Issue introduces the preparatory work to establishing a workable Minimal Viable Product based on FAIR machine-actionable patient health records. The work comprised the following steps. First an exploration of the regulatory frameworks internationally and in each of the participating countries. This exploration was positive, in all of the countries the policy-framework was positively geared towards the FAIR principles.

The second step was focused on the FAIRification process of the data as well as the localisation of data repositories. The outcome of this phase was positive, a data-visiting exercise across countries and even across continents, worked out. This demonstrated that the basic principles of the FAIRification process were realistic and operational.

[†] Corresponding author: Mirjam van Reisen (Email: mirjamvanreisen@gmail.com; ORCID: 0000-0003-0627-8014).

In the third step the research team set out to understand the different conditions for deployment in different places. For this purpose the team identified places that varied widely in terms of remoteness, connectedness, health services provided and conditions for operations. The result was a list of specifications and requirements providing the basis for engineering software, that could be installed and serviced locally in these widely different conditions in an African context.

In a final stage the working conditions as pertaining to digital patient health records in facilities were studied, with a view to understanding the usefulness in the daily practice of the health facilities. Four parallel use-cases were identified and these were integrated in a Virtual Image that was deployed in 88 health facilities in 8 countries.

The result is a data intelligent system that relates to four purposes: (i) monitoring by Ministry of Health; (ii) monitoring of data generated and stored in each of the facilities (iii) visualisation of common statistics derived from data-visited by algorithms in the health facilities and (iv) a FAIR-store of data that can be used for dynamic use-cases, based on permission granted by the facility.

FAIR works in Africa and this test-case forms an excellent basis for further development and deployment—it combines data ownership with a data-visiting capacity, and this can form an ethical pipeline for innovation of ML and AI supported health systems. The lessons learned in the African setting can be of high value for implementation in other continents as well, including regions where health information systems and ICT infrastructure might be more advanced, but nevertheless not enabling FAIR based distributed analytics and learning by visiting local data. The research demonstrated that it is possible to include data from areas that are poorly connected and that quality data for ML and AI pipelines can and should include diverse data. This Special Issue documents that this is an achievable goal: the development of an inclusive, high quality, and ethical pipeline based on stewardship of data according to FAIR Guidelines of medical patient records.

AUTHOR CONTRIBUTION STATEMENT

Mirjam Van Reisen (0000-0003-0627-8014, mirjamvanreisen@gmail.com); conceptualization, methodology, formal analysis, visualization, writing original draft preparation, writing—review and editing, supervision, project administration, funding acquisition Barend Mons (0000-0003-3934-0072, barendmons@gmail.com) conceptualization, writing—review and editing.