Exploring the Design of Low-End Technology to Increase Patient Connectivity to Electronic Health Records

Rens Kievit[1][a], Abdullahi Abubakar Kawu[2][b], Mirjam van Reisen[1][c], Dympna O'Sullivan[2][d] and Lucy Hederman[3][e]

[1] *Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Leiden, The Netherlands*

[2] *Technological University, Dublin, Ireland*

[3] *School of Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland*

Keywords: Patient Generated Health Data, FAIR Data, Interactive Voice Response, Home Monitoring Devices, Patient Connectivity.

Abstract: The tracking of the vitals of patients with long term health problems is essential for clinicians to determine proper care. Using Patient Generated Health Data (PGHD) communicated remotely allows patients to be monitored without requiring frequent hospital visits. Issues might arise when the communication of data digitally is difficult or impossible due to a lack of access to internet or a low level of digital literacy as is the case in many African countries. The VODAN-Africa project (van Reisen et al., 2021) started in 2020 and has greatly increased the capabilities of clinics in different countries in both Africa and Asia, but currently no systems are in place for the integration of external data from patients with long term health problems. In this article we outline our investigation into methods to increase the connectivity of patients with long term health problems with their clinics, and propose a solution in the form of a data pipeline prototype based on an Interactive Voice Response (IVR) system.

1 INTRODUCTION

With the creation of the VODAN-Africa network (van Reisen et al., 2021), the health data capabilities of many non-western countries have been improved tremendously. At the time of this writing, there are a total of 67 health facilities with patient instances spread out over 8 countries on the African continent. The VODAN-Africa network is a novel healthcare framework that is entirely based on the FAIR principles. These were introduced by (Wilkinson et al., 2016) and describe guidelines to improve the Findability, Accessibility (under well-defined constraints), Interoperability and Reproducibility of digital objects. The main goal of the development of these guidelines was to increase the capability and actionability of machine algorithms to better deal with the quickly increasing volume and complexity of data in the modern world.[1] Within the context of the VODAN-Africa project, these guidelines create the possibility of the analysis of healthcare trends on large geographical scales while being able to maintain data ownership within the clinics. The former happens through the standardization of the data flow and creation of controlled vocabularies which allows interoperability between clinics in various countries. Data ownership is ensured by a smart aggregating data visiting system (van Reisen et al., 2021; Plug et al., 2022).

The data collection and registration currently only happens within clinics, where medical staff treat the patients and collect data, and trained data clerks insert this data into the previously mentioned systems. Traveling to, or communication with, these clinics, however, might not always be trivial. In poorer or more rural regions, proper access to infrastructure such as roads or the internet might be complicated. This can be especially difficult for patients with long term physical-conditions such as asthma, diabetes,

[a] https://orcid.org/0000-0003-2328-4117
[b] https://orcid.org/0000-0003-2531-9539
[c] https://orcid.org/0000-0003-0627-8014
[d] https://orcid.org/0000-0003-2841-9738
[e] https://orcid.org/0000-0001-6073-4063

[1] A more detailed description of the guidelines and their use is described at https://www.go-fair.org/

Kievit, R., Kawu, A., van Reisen, M., O'Sullivan, D. and Hederman, L.
Exploring the Design of Low-End Technology to Increase Patient Connectivity to Electronic Health Records.
DOI: 10.5220/0012460900003657
In *Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2024) - Volume 2*, pages 194-200
ISBN: 978-989-758-688-0; ISSN: 2184-4305

epilepsy and high blood pressure. These are patients that do not need to be in a clinic at all times, but their vitals would ideally be tracked at short intervals over a long period of time to detect developments in their condition or predict deterioration in advance so that proper care can be administered. To increase the connectivity of this type of patient, we have developed a proof of concept in this work for a FAIR-based data pipeline that is capable of processing Patient Generated Health Data (PGHD) in a way that makes it interoperable with the existing VODAN-Africa framework without requiring access to the internet. PGHD refers to health data that is collected by the patient at their homes or other settings outside the clinic environment towards their care (Hussein et al., 2021).

This research began with an interview conducted with a doctor, a patient and two data managers from a clinic in Nigeria. The results and discussion surrounding this interview are presented by Kawu et al. (2024). They find that PGHD is currently already in use, mostly for addressing white collar hypertension, this is when patients typically have higher blood pressure readings in a clinic then they do at home. However this is limited, purely on paper, and not integrated with the rest of the electronic health records. From the interviews, Kawu et al. (2024) define four recommendations for a system integrating PGHD with digital health systems in low resource settings. The system should be supportive of the economic conditions of users and should be intuitive and accessible to all users (I). The metadata surrounding the PGHD should clearly describe both time and location when the data is gathered (II). The data should be kept separate from the traditional clinical data to create trust in data quality (III). The PGHD should be embedded as part of the clinical workflow to ensure continued engagement with patients (IV).

Following recommendation I, we find that the devices to be used for both the collection and communication of PGHD should be kept both cheap and simple accounting for the low-resource setting of this research. In a western and more resource-rich context, wearables can fulfill the role of automatically collecting and communicating PGHD over the internet. We have briefly investigated the possibility of utilizing them here, but find these to often be both too expensive and too unintuitive as was also stated by Kawu et al. (2024). Therefore we opt for a more practical approach requiring only a digital health monitor which a patient is able to use themselves and a working cell phone. To allow the patient or their carer to communicate the digital health monitor data, we develop an Interactive Voice Response system which is accessible over any kind of mobile phone, using the services provided by Africa's Talking. We include various types of auxillary data surrounding the creation of PGHD and keep our data separate from existing data source in the VODAN-Africa network to adhere to respectively recommendations II and III from Kawu et al. (2024).

Further, Kawu et al. (2023) provides a framework which describes how PGHD can be made FAIR using a curation tool. In this work we use The Center for Expanded Data Annotation and Retrieval (CEDAR, Musen et al., 2015). In the following section, we will outline the method used to achieve patient connectivity using our IVR-developed application with the CEDAR tool.

2 IMPLEMENTATION

In this section we present the implementation details of our proof of concept for a data pipeline from data collection up to data storage and analytics. The implementation presented here has been developed specifically for a generic blood pressure monitor that can measure a patient's pulse rate, systolic- and diastolic blood pressure readings. However, this can be easily expanded to include other types of digital health monitors, or even incorporate wearable technology via the internet if this is available. As suggested by Kawu et al. (2023), some patients may be required to track multiple PGHD which will require multiple devices, hence multiple sources of PGHD should be considered in the integration with the electronic health system. It may be the case that two PGHD devices collect the same or similar data, and so, a consolidation mechanism may need to be in place.

We present an overview of the pipeline in Figure 1. It is divided into three sections, the first two of which are further described in the following sections. The first is the data collection & communication which contains the creation of Patient Generated Health Data and the communication of this data to the clinic using the procedures described in the next subsection. The second part is the data storage & processing, which contains all the operations performed on the data within the clinic to transform the raw PGHD into FAIR data objects. The third and last step is the data application & access which uses the processed data objects stored in a clinic for prediction and visualization, this will be expanded upon further in future work. The code corresponding to the implementation outlined is accessible online.[2]

The procedures drawn with solid arrows corre-

[2]https://github.com/RenVit318/ivr_implementation_bp

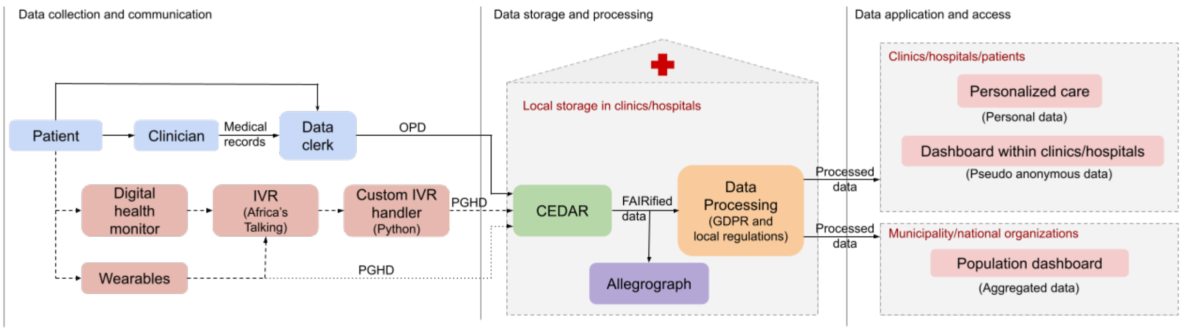

Figure 1: Overview of the complete data pipeline used for the PGHD collection.

spond to procedures present in the current VODAN implementations. The dashed arrows correspond to the new pipeline developed in this report. The dotted arrow leading from "Wearables" directly to "CEDAR" represents a direct connection between a digital wearable and CEDAR which requires a direct internet connection and is therefore not developed in this work. In the following sections, we will expand on the pipeline up to and including integration with CEDAR.

2.1 Data Collection & Communication

The first step in any data pipeline is the collection of the data to be processed. In the current version of the pipeline this collection has to be performed manually by the patient or their caregiver. They do this by either using a digital blood pressure monitor they have at home and reading off their vitals from its screen, or by simply reading these off from a wearable that continuously monitors them. In both cases the patient will have to memorize or write down their pulse rate, systolic blood pressure and diastolic blood pressure. Once they have this data, they call a cellphone number connected to our Interactive Voice Response (IVR) system, which we have set up using a service from Africa's Talking (AT).

Interactive voice response (IVR) is a technology that allows people to call and converse with a program that uses a synthetic human voice to communicate information and ask questions. The human interacting with this service can respond using the number keys on their cell phone.[3] This technology is ubiquitous in telephone systems for hospitals, where you are normally first placed inside a queue and need to interact with an IVR program first to be redirected to the correct department or physician. Our decision to utilize IVR for this implementation mainly stems from its accessibility; there is no requirement for literacy as the data requests are purely conveyed through sound and key presses. Of course a certain level of technical literacy is required to utilize both the digital blood pressure monitor and the cellphone, but that is inescapable. In addition to this, the IVR service provided by AT in Nigeria is also the cheapest option which means less financial strain on the clinics.

We have created a custom application written in Python that runs using Flask (Grinberg, 2018) and utilizes the Africa's Talking IVR service.[4] Our application communicates with this using Hypertext Transfer Protocol (HTTP) requests and provides instructions written in Extensible Markup Language (XML) interpretable by their Application Programming Interface (API) which describe what to do during a call with a patient, an example of this XML code is presented in Figure 3. The full procedure is outlined in Figure 2. It begins with the PGHD being collected using the BP monitor (step 1) and then calling our virtual phone number registered with Africa's Talking (step 2). The patient is then greeted and asked for their identification number. In future implementation work this can be used together with the incoming phone number to authenticate the patient and deduce which type of data needs to be requested. After this the medical data is requested field-by-field (steps 3-6), the patient can respond by inserting the corresponding measurements using the number pad on their phone. Then some additional auxiliary information surrounding the collection of the data is requested by providing numbered options. Currently implemented are the position of the patient during data collection (sitting, standing, laying), location during the collection (at home, or not at home), and the person collecting the data (the patient themselves, or a caregiver). The decision to use these fields is based on the results of Kawu et al. (2024) but could change in the process of future im-

[3] In some modern implementations of IVR users can also respond by speaking which is then translated into text by a machine learning algorithm. However this is naturally less reliable than manual key press of numbers on the mobile phone, and is also not currently offered as a possibility by Africa's Talking.

[4] https://africastalking.com/voice

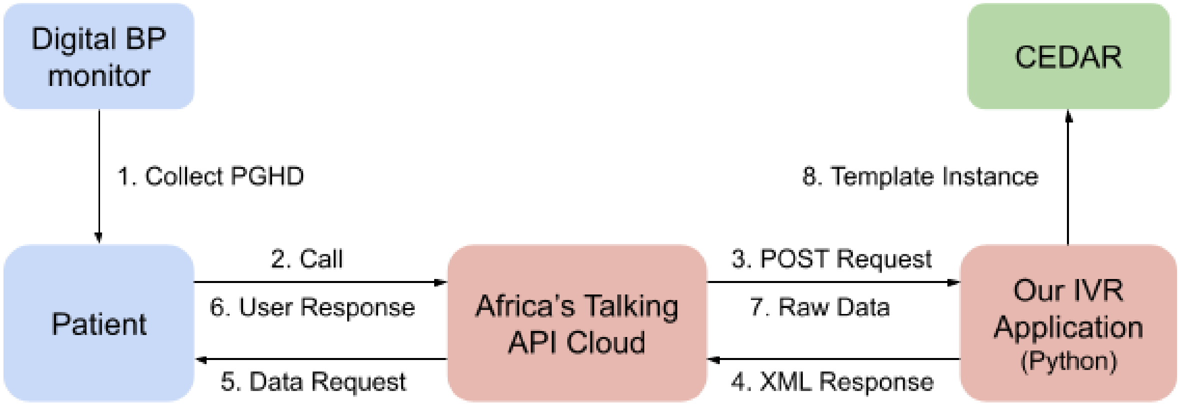

Figure 2: IVR communication protocol used in this work described in more detail in text.


```xml
<Response>
    <Say>Welcome to PGHD registration service</Say>
    <GetDigits timeout="30" finishOnKey="#" callbackUrl="https://9b1e-
80-114-159-161.eu.ngrok.io/pghd_cardio_handler">
        <Say>Press one followed by the hash sign to provide your readi
ngs from your digital heartrate monitor. Press two followed by the has
h sign to abort this call</Say>
    </GetDigits>
</Response>
```


Figure 3: XML Response given when sending a POST request to /pghd_handler.

plementation research. A field that is currently empty is some identifier of the device used for collecting data, this could be the International Mobile Equipment Identity (IMEI) number which should be known in the clinic and related to the patient ID.

The call is ended when the application has collected all data and stored it in a local JSON file. This file is sent using an HTTP POST request to CEDAR using its REST API accessed through application/json (steps 7 and 8 in Figure 2). In this step the data is converted into FAIR objects and saved as an instance of our CEDAR template which we will describe in more detail in the next subsection. Once the data is successfully sent it is immediately deleted from our IVR application and it is ready for the next caller. In a real-life implementation our application would be hosted on a web-server, but for the purposes of testing this prototype we temporarily host it online using ngrok.[5]

2.2 Data Storage & Processing

The second step in the data pipeline commences once the data communication procedure described above is completed. As described in more detail above, the collected data is sent to the CEDAR metadata center as an instance of the `Patient Generated Health Data: Blood Pressure` metadata template created for this project based on the existing outpatient data (OPD) template. Figure 4 shows a filled-in example of respectively the JSON file and the CEDAR template. The JSON file snippet (left) is how the data is stored during the data collection phase and what is sent to the CEDAR API in step 8 of Figure 2, the CEDAR template (right) shows the same information in the CEDAR webview. Each CEDAR field is described in the JSON file by its field name, e.g. `hasPulseRate`; this is not the same as the preferred label shown in the template ("The Pulse of the client") which is more verbose to improve human readability. Within the curly brackets the value (`@value`) and constraint on allowed values (`@type`) are given. For the three collection detail fields (`CollectionPosition`, `CollectionLocation`, `CollectionPerson`) the information is represented as a URI (`@id`) which is related to the ontology made for this project that will be discussed further later in this section.

Here the PGHD recorded by the patient is stored along with the date and time of collection, patient identification number and a new field indicating whether IVR was used to communicate the data to CEDAR. This final field is included for possible meta analyses in future work. The date of collection of the PGHD is currently assumed to be the same as the date of the phone call. However in reality this might not always be the case. Nevertheless due to possible issues that might arise in entering a long form value such as date using IVR we leave the exploration of this problem to future work.

The fields in the CEDAR template related to the measurement of blood pressure data are linked through Uniform Resource Identifiers (URI) to the Semantic Mining of Activity, Social, and Health data (SMASH) ontology (Phan et al., 2016).[6] This ontology describes concepts related to activity, social and health data. In this work we only utilize the pulse rate,

[5]https://ngrok.com/

[6]https://bioportal.bioontology.org/ontologies/SMASH

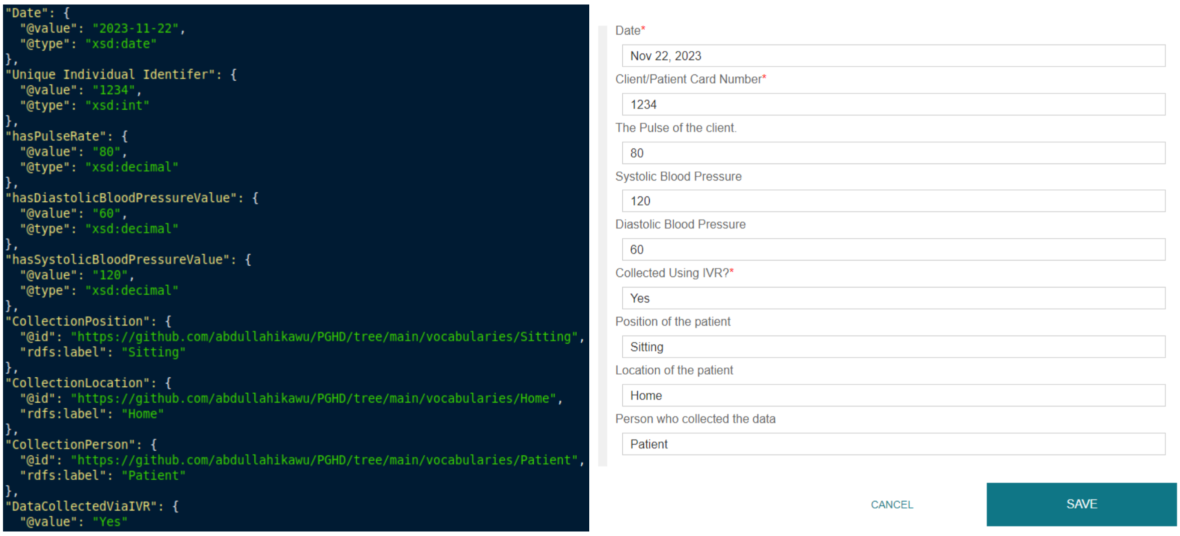

Figure 4: *left:* Snippet of the JSON package sent to CEDAR. *right:* Filled-In CEDAR template for collecting Patient Generated Health Data from a digital blood pressure monitor as used for this project.

systolic- and diastolic blood pressure field, however future work could incorporate fields such as those related to whether or not a patient does physical activity. The patient number is directly linked to the Unique Individual Identifier from the VODANTERMS ontology for interoperability with existing OPD records that contain information such as the age, gender, diagnoses and other physical information of the patient which could be used for additional monitoring (Van Reisen et al., 2022).[7] The fields related to the auxiliary information surrounding the measurement of blood pressure data are linked to a custom ontology hosted on BioPortal named PGHD-BP.[8] We have opted to create our own ontology for this to allow for more flexibility in describing the measurements based on current thoughts and later feedback by patients and clinicians using our system. The PGHD-BP ontology is specifically scoped to describe additional information surrounding BP data collected outside a clinical setting. It currently consists of two broad classes. The first is `CommunicationDetail` describing the ancillary information regarding the data communication which currently only describes whether or not the data was collected via the IVR service. The second is `CollectionDetail` which contains the auxiliary information surrounding data collection described in subsection 2.1.

[7]https://bioportal.bioontology.org/ontologies/VODANA-GENERAL

[8]https://bioportal.bioontology.org/ontologies/PGHD-BP

3 DISCUSSION

We have presented a prototype of an implementation for Patient Generated Health Data with the VODAN-Africa architecture through CEDAR using Interactive Voice Response. The implementation presented here fully adheres to recommendations I (any PGHD integration system should be conscious of the economic conditions of its users) and III (PGHD data should be kept separate from clinical data) by Kawu et al. (2024). It also partially adheres to recommendation II (PGHD should be accompanied with contextual metadata) insofar as there are rough descriptions surrounding the data collection, and a date is included with the communicated data. As mentioned in subsection 2.2 this date might not be the same as the real date of data collection, which is an issue to be tackled in future work. Recommendation IV (PGHD should be embedded in the clinical workflow) was not taken into consideration in this work and is also left for future implementation research.

It is important to note that the work presented in section 2 provides a complete workflow or pipeline for the implementation of a PGHD integration system, and a fully functional and tested prototype from data collection up to and including the storage on (a local version of) CEDAR. However integration with the VODAN-Africa architecture is not yet implemented, and would require further real-world testing. However once this integration is deployed, PGHD data could be included in the existing VODAN-Africa dashboard, both at an in-clinic level where clinicians can track the vitals of their patients in 'real-time' and aggregated on a national/global level to track the usage of PGHD across clinics.

Another crucial aspect that is still lacking in this pipeline is a proper method of access control. Allowing patients to send data to clinics remotely can introduce a range of issues from both data mismanagement and impersonation perspectives. To tackle these issues, secure authentication mechanisms would have to be implemented (Jati et al., 2022). These might make use of some combination of incoming phone number, patient ID and date of birth to make sure the reported PGHD are assigned to the correct patients.

Despite some of these shortcomings, the pipeline presented in this work provides a strong basis for the integration of PGHD in low-resource settings which has not been developed before to the authors knowledge and can serve as a great tool in further improving the capacities of healthcare in these regions that require it most.

4 CONCLUSIONS

We have developed a pipeline for the collection and communication of Patient Generated Health Data (PGHD) that is functional from the point of data collection to the storage on a (local) CEDAR installation. The communication of PGHD is performed using the Interactive Voice Response service provided by Africa's Talking in which an automated voice prompts the user to insert the readings of vitals that they have gathered themselves along with auxiliary information surrounding the measurements. The implementation presented here is specifically made for an at-home blood pressure monitor that is able to measure a patient's pulse rate and both systolic- and diastolic blood pressure. However expansion to more types of measurement devices is relatively simple by design. Upon data collection in CEDAR the provided PGHD is automatically FAIRified and enriched with metadata among which fields describing the date when the data was submitted and a flag indicating that the data was communicated using the IVR service.

While the prototype outlined here is not yet ready for direct implementation with the VODAN-Africa framework, once expanded, the inclusion of PGHD with existing Outpatient Data objects could form a cornerstone in further care for patients suffering from long term health problems in low resource settings such as the one considered in this work. We therefore emphasize the importance of future work to evaluate the acceptability and usability of this system with the patients that need it most.

ACKNOWLEDGEMENTS

We would like to thank the students in Fieldlab 5 from the 2023 Leiden University course 'Data Science in Practice' taught by Prof.dr.Mirjam van Reisen for their hard work and valuable insights in developing the pipeline presented in this work.

This work was conducted with the financial support of the Science Foundation Ireland Centre for Research Training in Digitally-Enhanced Reality (d-real) under Grant No. 18/CRT/6224. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission

REFERENCES

Grinberg, M. (2018). *Flask web development: developing web applications with python*. " O'Reilly Media, Inc.".

Hussein, R., Crutzen, R., Gutenberg, J., Kulnik, S., Sareban, M., and Niebauer, J. (2021). *Patient-Generated Health Data (PGHD) Interoperability: An Integrative Perspective*, volume 281, page 228–232. Studies in health technology and informatics.

Jati, P. H. P., van Reisen, M., Flikkenschild, E., Oladipo, F., Meerman, B., Plug, R., and Nodehi, S. (2022). Data Access, Control, and Privacy Protection in the VODAN-Africa Architecture. *Data Intelligence*, 4(4):938–954.

Kawu, A. A., Kievit, R., Abubakar, A., van Reisen, M., O'Sullivan, D., and Hederman, L. (2024). Exploring the integration of a patient generated health data in a fair digital health system in low-resourced settings: A user-centered approach. In *Proceedings of the 4th African Human Computer Interaction Conference*, AfriCHI '23, page 215–220, New York, NY, USA. Association for Computing Machinery.

Kawu, A. A., O'Sullivan, D., Hederman, L., and Reisen, M. (2023). Fair4pghd: A framework for fair implementation over pghd. *FAIR Connect*, 1:35–40.

Musen, M. A., Bean, C. A., Cheung, K.-H., Dumontier, M., Durante, K. A., Gevaert, O., Gonzalez-Beltran, A., Khatri, P., Kleinstein, S. H., O'Connor, M. J., Pouliot, Y., Rocca-Serra, P., Sansone, S.-A., Wiser, J. A., , and the CEDAR team (2015). The center for expanded data annotation and retrieval. *Journal of the American Medical Informatics Association*, 22(6):1148–1152.

Phan, N. H., Dou, D., Wang, H., Kil, D., and Piniewski, B. (2016). Ontology-based deep learning for human behavior prediction with explanations in health social networks (information sciences - if: 4.832). *Information Sciences*, 384.

Plug, R., Liang, Y., Basajja, M., Aktau, A., Jati, P., Amare, S., Taye, G., Mpezamihigo, M., Oladipo, F., and van Reisen, M. (2022). Fair and gdpr compliant population health data generation, processing and analytics.

van Reisen, M., Oladipo, F., Stokmans, M., Mpezamihgo, M., Folorunso, S., Schultes, E., Basajja, M., Aktau, A., Amare, S. Y., Taye, G. T., Purnama Jati, P. H., Chindoza, K., Wirtz, M., Ghardallou, M., van Stam, G., Ayele, W., Nalugala, R., Abdullahi, I., Osigwe, O., Graybeal, J., Medhanyie, A. A., Kawu, A. A., Liu, F., Wolstencroft, K., Flikkenschild, E., Lin, Y., Stocker, J., and Musen, M. A. (2021). Design of a fair digital data health infrastructure in africa for covid-19 reporting and research. *Advanced Genetics*, 2(2):e10050.

Van Reisen, M., Oladipo, F. O., Mpezamihigo, M., Plug, R., Basajja, M., Aktau, A., Jati, P. H. P., Nalugala, R., Folorunso, S., Amare, S. Y., Abdulahi, I., Afolabi, O. O., Mwesigwa, E., Taye, G. T., Kawu, A., Ghardallou, M., Liang, Y., Osigwe, O., Medhanyie, A. A., and Mawere, M. (2022). Incomplete COVID-19 Data: The Curation of Medical Health Data by the Virus Outbreak Data Network-Africa. *Data Intelligence*, 4(4):673–697.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., and et al. (2016). The fair guiding principles for scientific data management and stewardship. *Scientific Data*, 3(1).