Proceedings of the 14th Health Informatics in Africa Conference

HELINA' 23

PART II
Full Research Papers

Effective Implementation, Meaningful Use and Sustainability of Digital Health Interventions: The Role of Health Informatics and Imaging Informatics in Africa

Editors: Nicky Mostert, Ulrich Kemloh

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License.

Published in the Journal of Health Informatics in Africa
Volume 10 (2023) Issue 2

ISSN: 2197-6902
ISSN/NLM abbreviation: J Health Inform Afr
DOI: http://dx.doi.org/10.12856/JHIA-2023-v10-i2

Publisher
Koegni eHealth, Innovation for Development e.V. Germany
Postfach 652166
D-22372 Hamburg, Germany
www.koegni-ehealth.org
E-mail: info@koegni-ehealth.org

Table of Contents

Journal of Health Informatics in Africa

Editorial to the HELINA 2023 Proceedings
JHIA Vol. 10 (2023) Issue 2

Nicky Mostert

Nelson Mandela University, Gqeberha, South Africa

The 2023 edition of the HELINA (HEaLthINformaticsinAfrica) conference was held from 01 – 03 November 2023 as an in-person conference in Cape Town, South Africa. The conference was hosted by the South African Health Informatics Association (SAHIA) in partnership with HELINA. The conference theme was focused on "Effective Implementation, Meaningful Use and Sustainability of Digital Health Interventions: The Role of Health Informatics and Imaging Informatics in Africa". The aim was to provide a platform to showcase interventions in the following formats:

- Full research papers based on mature research results;
- Work in progress papers based on work in process research; and
- Case study presentations based on case studies and experience in digital health interventions.

Review Process

A total of 61 submissions in these categories were received. A double-blind peer review process was used for evaluating each full research and work in progress paper. These submissions were anonymized before being submitted to reviewers according to their area of expertise. The Scientific Programme Committee (SPC) based their final decision on the acceptance of each full research and work in progress paper on the recommendations and comments from reviewers. Accepted submissions were then sent back to the authors for revision according to the reviewers' comments. Case study submissions were reviewed by the SPC for inclusion in the conference. This review process resulted in the following acceptance rates:

- Full research papers: 44% acceptance rate (9 received, 4 accepted)
- Work in progress papers: 66% acceptance rate (6 received, 4 accepted)
- Case studies and experience papers: 85% acceptance rate (46 received, 39 accepted)

In order to be included in the conference proceedings, authors had to submit their final camera-ready papers after incorporating reviewer comments, and accepted papers had to be presented at the conference in person.

Nicky Mostert
HELINA 2023 SPC Chair
January 24

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License.

Journal of Health Informatics in Africa

Implications of Software Platform Architecture and Documentation on Developer Productivity: A Case of the Malawi Point of Care EMR Software

Emmanuel Mkambankhani [a], *, Tiwonge Davis Manda [a]

Department of Computing, University of Malawi, Zomba, Malawi

Background and Purpose: Software platform architecture and documentation affect platform customisability, developer productivity, and third-party contributions. This study examines how the Malawi Point of Care Electronic Medical Records Software (PoC-EMRS) architecture and documentation shape the platform's customisation, developer productivity, and third-party contributions. The Malawi PoC-EMRS, as the primary case of analysis, was compared to CommCare and DHIS2 as configurable software platforms. Theoretically, we draw on Generativity and the Boundary Resource Model (BRM). Generativity evaluates the overall capacity of an artefact to produce solutions, for diverse use cases. BRM was used to evaluate how owners of the PoC-EMRS facilitate ecosystem value co-creation through documentation and exposure of APIs.

Methods: Primary and secondary data were collected through interviews, observations and document analysis. Both qualitative and quantitative data were analysed to find common themes.

Results: Malawi PoC-EMRS is less configurable than DHIS2 and CommCare, necessitating more developer effort to support a variety of use cases. Though Malawi PoC-EMRS exposes boundary resources, it lacks incentives to attract third-party developers. Lack of or limited access to detailed documentation also negatively affects internal and third-party development productivity.

Conclusions: We established that CommCare and DHIS2 have the following strengths. First, the platforms feature standardised interfaces that allow third parties to design, configure, and customise the platforms into solutions for a varied range of use cases. Second, CommCare and DHIS2 have comprehensive documentation that is accessible through online repositories, communities of practice and demo sandboxes. Applying the strengths of these platforms in the development of the Malawi PoC-EMRS is likely to boost developer productivity.

Keywords: Health Systems, Configurable Platforms, Software Architecture, Software Documentation, Software Development Productivity.

1 Introduction

Software architecture is defined by ISO/IEC/IEEE 42010:2011 as the fundamental concepts or properties of a system embodied in its components, their relationships to each other and to the environment, and the principles guiding its design and evolution [1]. Software architecture is critical to the success of software platforms, as it manifests not only business and domain requirements but also other attributes like adaptability, maintainability, flexibility, scalability, performance, modifiability, and security [2]. An architecture that is well documented improves communication among stakeholders, serves to restrict design alternatives, channels the creativity of developers, reduces design and system complexity, and acts as a foundation for training new members of a team [3]. Thus, software architecture and its documentation influence the implementation of software platforms and have an impact on software developer productivity and third-party innovation [4]. Currently, software developers are moving towards developing their platform architectures to be more configurable and generic to reach a broader range of use cases beyond the targeted use cases that their platforms were initially designed for, and also to open opportunities for third-party contributions [5].

*Corresponding author address: emmanuel.mkambankhani@gmail.com

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):1-9. DOI: 10.12856/JHIA-2023-v10-i2-383

In Malawi, the Ministry of Health and Population (MOH), in coordination with the Digital Health Division (DHD) and the e-Government department, commissions various implementing organisations to develop eHealth solutions, through donor-funded projects. Following are some of the leading organisations in developing eHealth solutions: Baobab Health Trust (BHT), Elizabeth Glaser Paediatric AIDS Foundation (EGPAF), Luke International Norway (LIN), D-Tree, and Lighthouse. Collectively, these organisations have, over the past two decades, provided digital health solutions that have improved the delivery of healthcare services to patients or recipients of care, enrolled in different health programs. Currently, the District Health Information System 2 (DHIS2), the Malawi Point of Care Electronic Medical Record Software (PoC-EMRS), and CommCare are the leading software platforms for HIS in Malawi [6] [7] [8] [9].

Since 2001, Malawi PoC-EMRS has been evolving as an open-source solution for Point of Care (PoC) systems in Malawi, with over 206 health facilities using it as a Point of Care solution and 520 health facilities using it as a retrospective data entry application. Despite its widespread use, the development of the Malawi PoC-EMRS has hardly attracted external contributions. Thus, organisations responsible for its development have had to considerably keep expanding their internal software development teams, to meet growing demands for software products. Efforts to acquire skilled and experienced software developers, as well as bring new developers on board, can be quite costly. For example, new developers may require extensive training and onboarding efforts. In contrast to the Malawi PoC-EMRS, DHIS2 and CommCare solutions development in Malawi has leveraged the contributions of smaller teams from across multiple organisations. The two platforms have a growing international reach and increasing third-party initiatives. According to extant literature, internal and external development productivity and third-party innovation can be attributed to software architecture [3] [4] [10] [11] [12].

In order to improve the Malawi PoC-EMRS architecture and documentation, we sought to investigate the implications of its software's architectural design and documentation and compare it with the configurable platform approaches of CommCare and DHIS2. Following this exercise, our aim was to inform software architectural design revisions and improvement of documentation for the Malawi PoC-EMRS, to facilitate improvements in developer productivity and third-party contributions. The study was guided by the following question: *"How does the software architecture and documentation of the Malawi PoC-EMRS affect internal and external productivity and innovation?"* The study also analysed how the strengths of DHIS2 and CommCare can be leveraged to increase the productivity of internal and third-party contributions in Malawi PoC-EMRS.

2 Materials and Methods

2.1 Case Selection

During the study, the primary unit of inquiry was the Malawi PoC-EMRS platform. However, CommCare and DHIS2 were used as comparison benchmarks. The platforms were selected for two major reasons. First, they are widely deployed in Malawi's health sector. For example, the Malawi PoC-EMRS has the point of care and ART eMastercard modules deployed in 726 sites [7]. CommCare is used for mHealth solutions for Community Health Workers (CHW) and it is the platform upon which ART Back to Care (B2C) applications and a key Logistic Management Information System (LMIS) are based [8]. DHIS2 is also widely used, forming the basis for the National Health Management Information System [9], the Integrated Community Health Information System (iCHIS), and the National Agriculture Management Information System (NAMIS). Further to this, PEPFAR and WHO use DATIM, a DHIS2-based implementation [13]. Second, the study platforms were chosen because they are open-source. In being open-source, the platforms provided installation and code analysis flexibility, allowing analysis of each platform's architecture, maintainability, customization, and possibilities for facilitating third-party contributions.

2.2 Research Paradigm and Approach

Research paradigms, often termed "philosophical worldviews," are fundamental beliefs, assumptions, and values that govern research decisions. The study employed a constructivist view. Researchers may use

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):1-9. DOI: 10.12856/JHIA-2023-v10-i2-383

constructivism to develop subjective meanings of research phenomena when interacting with participants or research objects [14] [15]. To triangulate data and control for biases, the study used multiple sources and data collection methods, including questionnaires (for structured questions), participant observations, document analysis, and artefact analysis. Collected data include perceptions of developers for the Malawi PoC-EMRS, DHIS2, and CommCare in terms of ease of mastery, adaptability, and boundary resource use.

2.3 Sampling

Participants were selected using purposive and snowball sampling. First, purposive sampling relies on the researcher's judgement to select the appropriate sources of data to meet study goals. Purposive sampling was used to acquire data from CommCare, DHIS2, and Malawi PoC-EMR platform developers. Key people from organisations and government departments that develop and implement digital health solutions using the platforms were contacted to participate in the research. Second, snowballing was utilised to locate other participants since the research population was specific and more data was needed to qualitatively address the research question. Snowballing is useful when researchers know little about a group or organisation [16]. Specialists from the following five organisations participated in the study: Baobab Health Trust (BHT), Elizabeth Glaser Paediatric AIDS Foundation (EGPAF), Last Mile Health, Luke International Norway (LIN), and the Ministry of Health's Digital Health Division (DHD). By applying the techniques mentioned, 22 individuals were identified to participate in the study but only 19 participants participated in the study. Table 1 depicts participants' positions by the platform they responded to.

Table 1: Participant’s Platform and Position

#	Position	CommCare	DHIS2	Malawi PoC-EMRS
1	Digital Health Specialist	0	1	0
2	DHIS2 Programme Manager	0	1	0
3	Informatics Specialist	1	3	0
4	Software Architect	1	0	0
5	Software Developer	2	2	7
6	Systems Analyst	0	0	1
	Total	**4**	**7**	**8**

2.4 Data Collection

The study used a self-administered questionnaire, participant observations, document analysis, and artefact analysis. A self-administered questionnaire was used in the form of a Google form, sent through email. This was done to allow participants to respond at their own pace. The questionnaire was followed by a virtual and in-person interview, to clarify participants’ responses. Second, the study used participant observations, to gain more insights beyond what participants provided through the questionnaires and interviews. Participant observations have been documented to foreground hidden issues that may be sensitive for people to reply to [17]. Participants were observed through two exercises. The first author attended daily stand-up meetings and sprint review meetings and talked to developers about the study platforms. Furthermore, three of the 19 participants that responded to the questionnaire were chosen for controlled observations, through practical tasks. The participants were first oriented on the platforms after which they were given 30 minutes to develop a simple patient registration form, aimed at collecting demographics.

Document analysis also revealed hidden meanings. The definition of documents goes beyond text data; it also includes audio-visual data (photographs, diagrams, animations, video, and sounds) and electronic data like screenshots [18]. Hence, researchers reviewed technical platform documents, attended platform courses offered as video clips, PowerPoint slides, conferences, and webinars, and generated qualitative data notes. Finally, the researchers examined platform artefacts by executing simple tasks on online platforms. The researchers then checked open-source repositories for GitHub commits, forks, and pull requests from

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):1-9. DOI: 10.12856/JHIA-2023-v10-i2-383

internal and external developers. The researchers were able to assess platform contributions and whether internal and third-party contributors followed software programming standards.

2.5 Data Analysis

Thematic Analysis was utilised to analyse data based on the generativity and boundary resources model conceptual frameworks [19] [20] [21] [22] [23]. Thematic analysis was used to discover themes in the questionnaire, participant observations, document analysis, and artefact analysis data on developers' viewpoints and platform issues. The researchers then summarised the data to identify themes. Ultimately, the researchers linked and categorised themes according to the conceptual framework to analyse data based on platform generative capability and constrained resource availability and use [17] [24]. The primary themes derived from the conceptual framework encompassed the analysis of platform architecture's capacity for leverage, adaptability, and documentation, with their impact on the productivity of software developers and contributions from third-party entities.

3 Results

3.1 Platform Architecture on Productivity in Solutions for Different Use Cases

The investigations on how the architectural design of the Malawi PoC-EMRS shapes software developer productivity were based on the theory of Generativity [21]. The focus of this study was the generative technology and generative capacity. A generative technology is described to have five characteristics; the capacity for leverage, adaptability, ease of mastery, accessibility, and transferability [19]. Figure 1 illustrates the responses of participants on each of the study platforms' capacity for leverage. The numbers in the figure indicate a distribution of responses by the 19 participants in the study.

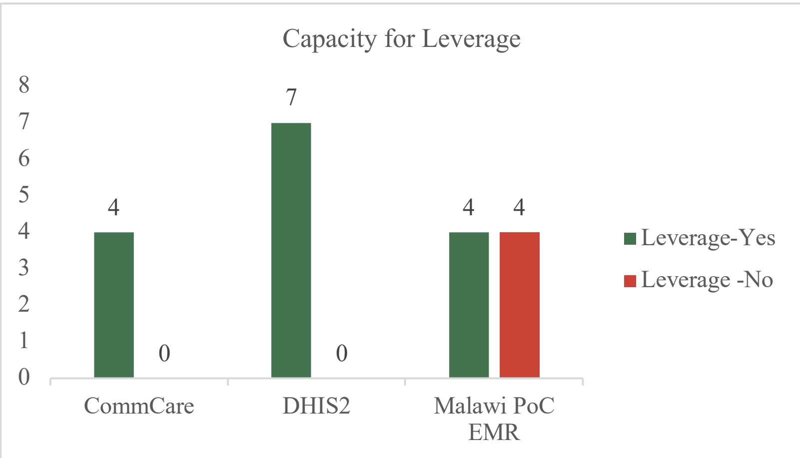


Figure 1: Number of Participants and Capacity for Leverage

From the data, it was established that the leverage of DHIS2 is attributed to its easy-to-use functionalities implemented by drag-and-drop features within the system's user interface. Thus, making it simpler to customise existing or develop new customizable data collection apps, as well as developing dashboards and analytics as one of the participants indicated below.

"DHIS2 has drag and drop functionality and it does not really need one to have the IT background to customise or manage it" (DHIS2 Programme Manager-DHD, 2021).

Similarly, in CommCare, leverage is provided by how simple it is for platform users to develop mobile data collection tools quickly. The CommCare HQ interface facilitates the implementation of data collection apps

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):1-9. DOI: 10.12856/JHIA-2023-v10-i2-383

and data management that would otherwise have to be created from scratch. This is clear from the comments that follow:

“The platform provides interfaces that scaffold the heavy lifting e.g. authentication, data storage, etc” (Software Architect EGPF, 2022).

“There are a lot of tools one can use to come up with an app easily” (Software Developer-EGPAF,2022).

In the case of Malawi PoC-EMRS, it was established that the platform does not fully have the capacity for leverage. This is so because a common response was not obtained from participants who are using the platform. The following are some of the expressions that support the preceding statement:

“It is user-friendly and has clear guides on what needs to be done” (Systems Analyst BHT, 2022).

"There are a lot of workarounds required to add specific features to the UI using the touchscreen tool kit. There's no standardisation or reusability of business logic and modularization is terrible.” (Software Developer EGPAF, 2022)

In relation to leverage, adaptability is another attribute, which is described as the potential of a platform to be flexible to change, for use in different contexts than the one it was designed for. Figure 2 illustrates the responses of participants to each platform’s adaptability.

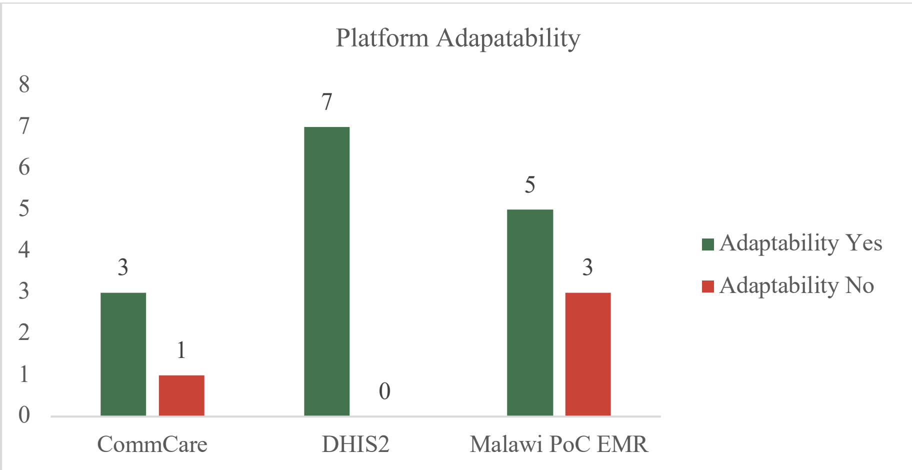


Figure 2: Adaptability of Platforms

From another perspective, adaptability can be expressed in how the platforms respond to multiple end-user devices. From the results in artefact analysis, all three platforms are capable of running on mobile and desktop devices. Front-ends for CommCare HQ, DHIS2, and Malawi PoC-EMRS are all capable of resizing and fitting on multiple device screens for desktops, tablets, and phones.

3.2 Platform Architecture Design and Third-Party Development

Boundary resources are software tools and regulations, such as application programming interfaces (APIs) and software development kits (SDKs), that act as an interface between platform owners and application developers, as well as allow the development of third-party applications. Table 2 shows the existence of boundary resources that each platform is able to expose.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):1-9. DOI: 10.12856/JHIA-2023-v10-i2-383

Table 2: Existence of Boundary Resources in Platforms

Boundary Resource	CommCare	DHIS2	Malawi PoC-EMRS
API endpoints	✓	✓	✓
Open Source code Repository	✓	✓	✓
Debugging Tools	✓	✓	✓
Reusable Libraries		✓	✓
SDK		✓	

GitHub forks are one indicator of the presence of third-party developers. Forking is the process of creating a new software repository by copying an existing one [25]. This enables third-party developers to experiment with their own local repository without affecting the original project [26]. Furthermore, forking allows third-party developers to submit pull requests to original code repositories to which they do not have access rights contribute [25] [27]. Table 3 shows the summary of repository forks that third-party developers made to CommCare, DHIS2, and Malawi PoC-EMRS.

Table 3: Summary of Platform Forks

Platform	Platform module	Number of forks
DHIS2	DHIS2 Core	263
	DHIS2 application platform	7
	DHIS2 UI	4
CommCare	CommCare HQ	196
	CommCare Mobile	22
Malawi PoC-EMRS	EMR-API	3
	Core	3

3.3 Platform Documentation and Productivity

Documentation is a form of boundary resource that gives actors the generative capacity to utilise other boundary resources on a platform [23]. As shown in Table 4, DHIS2 and CommCare platforms have a variety of information available online that enables third-party development via user manuals, online courses, and conferences.

Table 4: Participants able to access platform resources

Boundary Resource	Number of Participants		
	CommCare	DHIS2	Malawi PoC-EMRS
SDK/API documentation	3	7	5
Technical Specification Documents	4	7	3
User Manuals	4	6	3
Training material or tutorials	4	6	3
Demo site/platform for practice	0	1	0
Platform Academy	0	1	0

The availability of documentation contributed to how much time it took the participants in this study to master the platforms. The study established that in CommCare, three of four participants said it took them no more than three months to be proficient using the platform, while four of seven in DHIS2 said the same. This is attributed to the presence of online training resources for CommCare and DHIS2. In the case of Malawi PoC-EMRS, five of eight participants indicated that it took the same time, while three of eight participants indicated that it took 7 - 12 months, thus establishing that CommCare and DHIS2 were easier to master. This was also observed during the participant observation exercise. The task took an average of 16.11 minutes to complete for all participants using CommCare and an average of 22.86 minutes for people

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):1-9. DOI: 10.12856/JHIA-2023-v10-i2-383

using DHIS2. All participants were able to complete five of the nine form fields needed for the Malawi PoC-EMRS in an average of 44.6 minutes but they took an average of 82.15 minutes to complete the task.

4 Discussion and Conclusion

Malawi PoC-EMRS was compared with DHIS2 and CommCare architectures to determine the productivity of internal and third-party developers. Using concepts from the theory of generativity like generative technology and generative capacity, it was established that platform architecture affects developer productivity [21] [23] [28]. We also established that software platforms that provide incentives to developers enhance productivity enhance productivity and attract third-party contributions. By incentives, we mean platforms being configurable, flexible enough to be used for more than one use case, and cross-platform. This agrees with Msiska and Nielsen [19] that, in order to leverage a platform's capabilities, the platform needs to provide incentives to actors to use it. The Malawi PoC-EMRS can be used for cross-platform devices, but it lacks the configurability of DHIS2 and CommCare. Thus, the Malawi PoC-EMRS requires more effort by developers, in order to develop solutions for a diverse number of use cases. In this scenario, adaptability is expressed by how the system can be used for multiple business use cases therefore, the platform is not adaptable in this regard, thus concurring with other studies [11] [29] that platforms that allow customisations through the use of configurable templates attract more innovations and contributions, and that when the configurable templates fail, developers can find a workaround.

Software architecture design defines the availability of interfaces that allow internal and external software developers to interact with the platform's core functionality. This study used the Boundary Resource Model (BRM) to examine the three platforms' availability of boundary resources and how they are used. We established that a lack of third-party oriented incentives in the Malawi PoC-EMRS contributes to the absence of third-party contributors. These findings agree with Russpatrick [30] and Chirwa et al. [31] in that the availability of boundary resources alone is insufficient to attract third-party developers; external incentives are required. In addition, according to Msiska and Nielsen [19] third-party development is only possible when there is sufficient external generative capacity, regardless of how good a software ecosystem's boundary resources are. This was obvious because DHIS2 and CommCare have a greater number of third-party developers than the Malawi PoC-EMRS.

Extending from the theory of generativity and the usage of boundary resources, the construct of generative capacity has two perspectives: one that focuses on technology and another that focuses on actors. Documentation of a platform increases the generative capacity of an actor to be able to produce something, using the boundary resources of the platform, thus reducing the gap between technology and actor [32]. In this study, we established that the availability of documentation increases productivity, while its absence limits it and concurs with Duarte [4] where it was identified that effective training and the availability of quality documentation on a platform, increase the productivity of software developers. This is evident as DHIS2 and CommCare have considerable online documentation accessible to developers and third-party innovators, while Malawi PoC-EMRS has limited documentation and access to it is also limited [33], thus affecting productivity.

In this study, we identified three areas where platforms show their strength and increase the productivity of internal and third-party developers. Malawi PoC-EMRS can increase the productivity of internal and third-party developers by improving the availability of boundary resources through providing configurability and standardisation of its APIs, boundary resource use, and platform management. Improving the availability of boundary resources can be achieved by enhancing the Malawi PoC-EMRS to provide configurable interfaces to enable end-user configurability. This concurs with the assessment done by Munthali et al. [34] on BHT, which was the Malawi PoC-EMRS implementer, where five key areas of improvement were identified, one of which was to improve the productivity and quality of the platform to achieve 70% reconfigurability and 30% customizability. Secondly, Malawi PoC-EMRS needs to standardise and open its API endpoints in order to give third-party innovators the capacity and flexibility to improve or customise the platform [35]. In addition, there is a need to enhance the capacity of developers thus, improving boundary resource use [10] [30] [31]. This can be achieved when Malawi PoC-EMRS improves its documentation and provision of training materials, to internal and external developers. The absence of sufficient documentation is another factor that limits productivity, as highlighted by Duarte [4]. Lastly, the

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):1-9. DOI: 10.12856/JHIA-2023-v10-i2-383

Malawi PoC-EMRS needs to establish a Community of Practice (COP) to allow platform owners and contributors to share knowledge, and establish a governance structure. COPs play a significant role in platform management and governance as they are a major component of generativity in platform ecosystems [21] [30] [31] [36].

Acknowledgements

The authors would like to thank everyone who participated in this study.

Statement on conflicts of interest

The authors declare that there is no conflict of interest

References

[1] A. A. C. da Costa, S. Misra, and M. S. Soares, "A Systematic Mapping Study on Software Architectures Description Based on ISO/IEC/IEEE 42010:2011," in *Computational Science and Its Applications – ICCSA 2019*, S. Misra, O. Gervasi, B. Murgante, E. Stankova, V. Korkhov, C. Torre, A. M. A. C. Rocha, D. Taniar, B. O. Apduhan, and E. Tarantino, Eds., in Lecture Notes in Computer Science, vol. 11623. Cham: Springer International Publishing, 2019, pp. 17–30. doi: 10.1007/978-3-030-24308-1_2.
[2] S. M. A. Khan, *Popular Software Architecture Used In Software Development*. 2023.
[3] N. A. Ernst and M. P. Robillard, "A Study of Documentation for Software Architecture." arXiv, May 26, 2023. Accessed: Aug. 16, 2023. [Online]. Available: http://arxiv.org/abs/2305.17286
[4] C. H. C. Duarte, "Software Productivity in Practice: A Systematic Mapping Study," *Software*, vol. 1, no. 2, pp. 164–214, May 2022, doi: 10.3390/software1020008.
[5] M. Engert, M. Pfaff, and H. Krcmar, "Adoption of Software Platforms: Reviewing Influencing Factors and Outlining Future Research," 2019.
[6] Baoba Health Trust, "BHT 3-Tier Software Architecture." 2018.
[7] EGPAF, "Improving Quality of Care and Health Impact Through Innovative Systems and Technologies in Malawi (Health Information Systems (HIS)), CDC - EGPAF," *Elizabeth Glaser Pediatric AIDS Foundation*, 2022. https://pedaids.org/project/improving-quality-of-care-and-health-impact-through-innovative-systems-and-technologies-in-malawi-health-information-systems-his-cdc/ (accessed Jun. 02, 2022).
[8] A. Chhetri, M. Iversen, J. Kaasbøll, and C. Kanjo, "Evaluating mHealth Apps Using Affordances: Case of CommCare Versus DHIS2 Tracker," in *Information and Communication Technologies for Development. Strengthening Southern-Driven Cooperation as a Catalyst for ICT4D*, P. Nielsen and H. C. Kimaro, Eds., in IFIP Advances in Information and Communication Technology, vol. 551. Cham: Springer International Publishing, 2019, pp. 619–632. doi: 10.1007/978-3-030-18400-1_51.
[9] Ministry of Health of Malawi, "NATIONAL DIGITAL HEALTH STRATEGY." 2020.
[10] D. Ametowobla and S. Kirchner, "The organization of digital platforms," p. 26, 2022.
[11] L. F. Byson and T. D. Manda, "Software Architecture and Software Usability: A Comparison of Data Entry Form Designing Configurable Platforms.," p. 8, 2020.
[12] S. Himayat and Dr. J. Ahmad, "Software Understandability using Software Metrics: An Exhaustive Review," *SSRN Electron. J.*, 2023, doi: 10.2139/ssrn.4447189.
[13] "DATIM | OpenHIM." http://openhim.org/docs/implementations/datim/ (accessed Aug. 17, 2023).
[14] S. Dawadi, S. Shrestha, and R. A. Giri, "Mixed-Methods Research: A Discussion on its Types, Challenges, and Criticisms," *J. Pract. Stud. Educ.*, vol. 2, no. 2, pp. 25–36, Feb. 2021, doi: 10.46809/jpse.v2i2.20.
[15] D. Mertens, "PHILOSOPHICAL ASSUMPTIONS AND PROGRAM EVALUATION," 2015.
[16] D. Makwana, P. Engineer, A. Dabhi, and H. Chudasama, "Sampling Methods in Research A Review," 2023.
[17] J. Creswell, *Research Design*, 4th ed. SAGE Publications, 2014.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):1-9. DOI: 10.12856/JHIA-2023-v10-i2-383

[18] D. M. Arya, J. L. C. Guo, and M. P. Robillard, “Information correspondence between types of documentation for APIs,” *Empir. Softw. Eng.*, vol. 25, no. 5, pp. 4069–4096, Sep. 2020, doi: 10.1007/s10664-020-09857-0.

[19] B. Msiska and P. Nielsen, “Innovation in the fringes of software ecosystems: the role of socio-technical generativity,” 2018.

[20] D. Fürstenau, A. Baiyere, K. Schewina, M. Schulte-Althoff, and H. Rothe, “Extended Generativity Theory on Digital Platforms,” *Inf. Syst. Res.*, p. isre.2023.1209, Mar. 2023, doi: 10.1287/isre.2023.1209.

[21] L. D. W. Thomas and R. Tee, “Generativity: A systematic review and conceptual framework,” *Int. J. Manag. Rev.*, vol. 24, no. 2, pp. 255–278, Apr. 2022, doi: 10.1111/ijmr.12277.

[22] N. Staub *et al.*, “Governance Mechanisms in Digital Platform Ecosystems: Addressing the Generativity-Control Tension,” *Commun. Assoc. Inf. Syst.*, vol. 51, pp. 906–939, 2022, doi: 10.17705/1CAIS.05137.

[23] T. Wulfert, “Boundary resource management in innovation ecosystems: The case of e-commerce,” *Electron. Mark.*, vol. 33, no. 1, p. 27, Dec. 2023, doi: 10.1007/s12525-023-00651-6.

[24] J. Caulfield, “How to Do Thematic Analysis | A Step-by-Step Guide & Examples,” 2019.

[25] S. Zhou, B. Vasilescu, and C. Kästner, “What the fork: a study of inefficient and efficient forking practices in social coding,” in *Proceedings of the 2019 27th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, Tallinn Estonia: ACM, Aug. 2019, pp. 350–361. doi: 10.1145/3338906.3338918.

[26] J. Businge, M. Openja, S. Nadi, and T. Berger, “Reuse and maintenance practices among divergent forks in three software ecosystems,” *Empir. Softw. Eng.*, vol. 27, no. 2, p. 54, Mar. 2022, doi: 10.1007/s10664-021-10078-2.

[27] B. B. Chua and Y. Zhang, “Applying a Systematic Literature Review and Content Analysis Method to Analyse Open Source Developers’ Forking Motivation Interpretation, Categories and Consequences,” *Australas. J. Inf. Syst.*, vol. 24, Jun. 2020, doi: 10.3127/ajis.v24i0.1714.

[28] S. D. S. Amorim, J. D. McGregor, E. S. De Almeida, and C. V. F. G. Chavez, “Software Architectural Practices: Influences on the Open Source Ecosystem Health,” *J. Softw. Eng. Res. Dev.*, Aug. 2023, doi: 10.5753/jserd.2023.967.

[29] M. Li, “Utilizing space for User Participation : Enabling and constraining factors for User Involvement when Designing a Data Entry Interface in a Generic Health Information Software in Uganda,” University of Oslo, 2018.

[30] S. Russpatrick, “Understanding Platform Ecosystems for Development: Enabling Innovation in Digital Global Public Goods Software Platforms,” in *The Future of Digital Work: The Challenge of Inequality*, R. K. Bandi, R. C. R., S. Klein, S. Madon, and E. Monteiro, Eds., in IFIP Advances in Information and Communication Technology, vol. 601. Cham: Springer International Publishing, 2020, pp. 148–162. doi: 10.1007/978-3-030-64697-4_12.

[31] E. Chirwa, P. Abbott, and J. Foster, “UNDERSTANDING THE INTERPLAY BETWEEN BOUNDARY RESOURCES AND GOVERNANCE PRACTICES IN INFLUENCING ECOSYSTEM VALUE CO-CREATION FOR DIGITAL PLATFORMS: A CASE FROM THE GLOBAL SOUTH,” 2021.

[32] C. Sadowski and T. Zimmermann, Eds., *Rethinking Productivity in Software Engineering*. Berkeley, CA: Apress, 2019. doi: 10.1007/978-1-4842-4221-6.

[33] D. B. Nkhoma and K. Munthali, “Investigation of Key Factors Affecting Quality of Patient Data from National Antiretroviral Therapy Electronic Medical Record System in Malawi,” 2021.

[34] A. Munthali, T. D. Manda, and K. Chikadza, “Strategic Plan for Baobab Health Trust (Malawi).” 2019.

[35] I. Miremadi, M. Khoshbash, and M. M. Saeedian, “Fostering generativity in platform ecosystems: How open innovation and complexity interact to influence platform adoption,” 2023, [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0048733323000653

[36] D. Paparova and M. Aanestad, “Governing Innovation in E-Health Platform Ecosystems – Key Concepts and Future Directions,” 2020.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):1-9. DOI: 10.12856/JHIA-2023-v10-i2-383

Journal of Health Informatics in Africa

Digital Health Model for South Africa's National Health Insurance: Addressing Hospital Occupancy and Emergency Care

Wesley Moonsamy [a,*], Shawren Singh [a]

[a] University of South Africa, Johannesburg, South Africa

Background and Purpose: South Africa will be unifying the current fragmented healthcare system by implementing the National Health Insurance. This poses inevitable challenges to the associated information systems. This research introduces the need for an information systems model to support the Federated Health Information Architecture proposed by the Department of Health.
Methods: The Health Normative Standards Framework documentation was studied in conjunction with the Life Esidimeni Health Ombud report to determine gaps in the proposed architecture.
Results: Five gaps were identified during this research (the view of shared Electronic Health Records (EHR) is oversimplified, key decision makers are not included in the list of stakeholders, emergency services is not adequately supported, medical aid schemes are not included, mature architectural standards need to be developed) and it was determined that a digital health systems model needs to be developed to support the current architecture to assist with resolving some of the identified gaps.
Conclusions: Data synchronisation issues are inevitable with a large project such as the NHI. To minimise mistakes, fewer assumptions regarding the interoperability of systems need to be made. The proposed architecture as it stands may not cater to the needs of the NHI. A new model that can support the NHI especially within emergency care and hospital occupancy monitoring needs to be created.

Keywords: National Health Insurance, Federated Health Systems Architecture, Digital Health Model, Electronic Health Records, Health 4.0

1 Introduction

The purpose of this paper is to introduce the need for a conceptual digital health model that is cognisant of the existing healthcare system constraints in South Africa and contextually relevant to the Department of Health's National Health Insurance (NHI) plan. The effective implementation of the NHI depends on healthcare systems which are sustainable, in other words future fit. The scenario which this research has focused on is emergency care and hospital occupancy.

2 Background

Electronic Health Records (EHRs) have become an integral part of patient care in South Africa [1]. EHRs not only benefit individual patients but can have a positive impact on public health goals [2]. Real time hospital occupancy information across hospitals in Gauteng is not yet available. As a result, paramedics could unknowingly transport a patient to an emergency facility that is at full capacity. In conjunction with addressing these types of scenarios, the Department of Health (DoH) aims to migrate South Africa from the current fragmented healthcare system to the NHI [3]. The NHI aims to unify the healthcare system to benefit all South Africans. This requires a substantial digital health transformation as existing information systems need to connect to each other to provide a seamless experience for healthcare workers and patients.

*Corresponding author address: wesleym312@gmail.com
© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):10-14. DOI: 10.12856/JHIA-2023-v10-i2-395

3 Materials and methods

This research has considered the Federated Health Information Architecture (FHIA) proposed by the DoH [4] as well as literature published in other developing countries. Since it is important for research and associated scenarios to be contextually relevant [5], The Life Esidimeni tragedy of 2016 which has close links with patient care and transfers was also considered when assessing the future readiness of the FHIA.

4 Literature Review

Globally, healthcare systems are constrained by factors such as aging populations with chronic conditions [6] however, South Africa has an increased burden on emergency services arising from violent crimes [7]. Health 4.0 poses opportunities for the optimisation of healthcare delivery by incorporating familiar technologies such as Cloud Computing (CC), Fog Computing (FC), Internet of Things (IoT) as well as the Internet of Medical Things (IoMT) [8]. To realise the benefits of Health 4.0, an all-encompassing health information systems architecture needs to be designed. South Africa is preparing to migrate the fragmented healthcare system (public and private sectors) to a unified healthcare system. This has necessitated the need for an implementation plan as well as the designing of a future state architecture.

4.1 NHI Implementation Phases

The NHI will be implemented in six phases [4]: (1) Identify national interoperability use cases; (2) National interoperability use cases are evaluated, validated, refined, and prioritised; (3) Identify gaps in the Health Normative Standards Framework (HNSF) and address them; (4) Articulation of related national interoperability specifications; (5) Implementation and informal assessment of national interoperability specifications; (6) Formal conformance assessment and conformance certification. Currently in phase 2 (2022 to 2026) of the implementation plan, the healthcare delivery mechanisms are being reinforced with more resources and the integration of selected private healthcare services [9]. The future state architecture with a focus on the NHI has been designed and is based on the Federated Health Information Architecture (FHIA). This is described in the next section.

4.2 Federated Health Information Architecture

As per step 3 of the implementation plan, the proposed FHIA was assessed, and potential gaps have been highlighted. The "Magnifying the problem" approach was then used to identify architectural gaps [10]. Health Information Architecture (HIA) encourages interactions between users (healthcare workers and patients) and health systems. Federated Architectures (FA) promote interoperability and synchronisation of data between different systems [11]. The blending of the principles of an FA with an existing HIA,

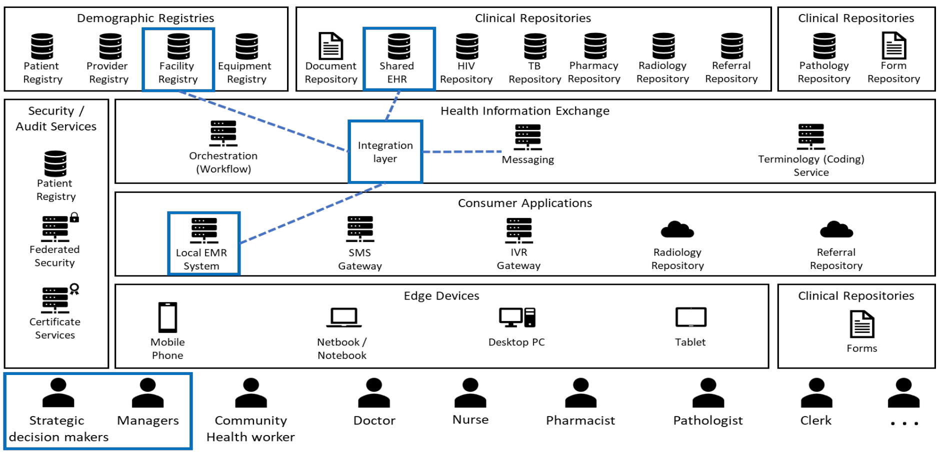


Figure 1: FHIA with highlighted focus areas based on this research. Source (adapted): [4]

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):10-14. DOI: 10.12856/JHIA-2023-v10-i2-395

results in a FHIA which affords the interoperability and data synchronisation capabilities onto an architecture that already supports interaction between users and systems. The FHIA designed for the NHI is presented in Figure 1. The FHIA was designed by the Department of Health with an objective of ensuring interoperability. Interoperability which supports the synchronisation of data is preferred over integration which focusses on the translation of data between heterogenous systems [12]. The HIE in this architecture is responsible for: Orchestration – management of workflows; Messaging – data flow of shared information; Terminology Service – Data transformation. It can be noted that minimum emphasis has been placed on the synchronisation of data between clinical systems and emergency services.

Medical records are depicted as a shared repository within the clinical repository. This layout therefore implies a consolidation of records [13]. Based on this understanding, the patient records, and by extension the data subsets, describing the overall pulse of the healthcare system will reside physically in a central location. Though there are merits to having a central repository, such as a single version of the data, this would require all facilities to change to the same HIS to be fully interoperable. Within the South African private healthcare system alone, there are no less than five hospital groups: Life Healthcare, Mediclinic, Netcare, Joint Medical Holdings and Lenmed [14]. This excludes the thousands of other smaller private medical facilities. It is reasonable to assume that hospitals in the private and public healthcare systems do not run on the same HIS or at the very least have customised their systems to an extent such that it would require significant effort to consolidate them into a single platform.

Private medical aids, emergency services and healthcare facilities will still feature in some aspects of the NHI. These components however are not adequately catered to within the FHIA. When considering the inevitable heterogenous landscape, an integration layer as suggested in Figure 1 could support the messaging layer. This would be useful in scenarios such as emergency medical care, patient transportation and inter facility patient transfers.

The Life Esidimeni tragedy of 2016 in which more than 100 mentally ill patients died as a result of being transferred to various ill-equipped healthcare facilities [15] has highlighted the following factors (Table 1).

Table 1: Information related factors that contributed towards the Life Esidimeni tragedy of 2016

Contributing factors relating to information processing	**Comments**
Lack of medical authorisation for patient transfer	Insufficient digital records existed for many of the patient transfers between the origin and destination healthcare facilities.
Lack of clinical assessment evidence	There were insufficient digital records of a clinical assessment being performed on the patients prior to being transferred. The follow-on effect resulted in a lack of future or chronic medical records being available.
Lack of patient monitoring	Insufficient digital records existed for the monitoring of patients' conditions during the transfer process.
Receiving staff were unable to interpret the incomplete medical records in cases where they were received	Receiving staff were not trained or did not know how to interpret the medical records received and were thus unable to provide the necessary medical care.
Insufficient handover and control of medical supplies that were received	Insufficient digital records were available to support the control of medical supplies that were received thus leading to deficient patient care.

Source (adapted): [16]

Considering the factors mentioned above, it can be noted that the absence of information (paper-based or digital) can have detrimental consequences. At this point, it might be useful to consider the "poka yoke[1]"

[1] Japanese term meaning to avoid unexpected surprises, poka-yoke is a safeguard that prevents a process from proceeding to the next step until the proper conditions have been satisfied and accepted.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):10-14. DOI: 10.12856/JHIA-2023-v10-i2-395

principle of ensuring that systems and processes are mistake proof [17]. The recommendations in the next section have therefore been created to support the existing architecture and to minimise information inconsistencies.

5 Recommendations

The identification of gaps is a key element of Enterprise Architecture. Step 3 of the NHI implementation plan refers to the identification of gaps in the HNSF. The gaps and recommendations in Table 2 refers to the FHIA proposed by the DoH.

Table 2: Gaps and recommendations relating to the FHIA

Possible gap	Recommendation
Shared EHR is depicted as a single, shared repository. This is an oversimplified view which assumes that the existing and future systems will be interoperable.	An integration layer needs to be considered to ensure that heterogenous systems can share and transmit patient records to healthcare providers at the point of patient care which is an aim of the DoH [18].
The stakeholder list in the FHIA has not included key decision makers.	The stakeholder list should include strategic decision makers who require selected data subsets to make important decisions. Mangers, who require other levels of information and processing should also be included. These additional stakeholder types will assist in revealing related use cases as per steps 1 and 2 of the implementation plan.
Emergency care and patient routing is not adequately depicted.	The HIE should include integration to translate messages between heterogenous EMR systems and healthcare facilities (private and public).
Medical aid schemes are not represented in the architecture	Medical aid schemes will still exist after the NHI has been implemented. These organisations require access to certain EHRs and should therefore be included in the architecture.
Architectural standards not yet as mature as the UK's National Health Service (NHS).	The NHS Architecture Principles should be consulted when creating the frameworks for the NHI [19].

6 Conclusion

It has been highlighted that inefficient health information can have negative consequences particularly during patient transfer, emergency care and facility management. During this research, the lessons learned from the recent Life Esidimeni tragedy were considered when assessing the FHIA proposed by the DoH. This led to recommendations that could ease some of the challenges that are inevitable during a large health information systems project. Future research will focus on creating a health information systems model that can exist on a standalone basis or support the FHIA. This will be achieved through conducting interviews with key stakeholders in the public healthcare system to understand their thoughts relating to current and proposed systems.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):10-14. DOI: 10.12856/JHIA-2023-v10-i2-395

References

[1] Moonsamy W, Singh S. Digital Vaccination Records: Exploring Stakeholder Perceptions in Gauteng, South Africa. The African Journal of Information and Communication (AJIC) [Internet]. 2022 May;(29):1–26. Available from: https://ajic.wits.ac.za/article/view/13756

[2] Diño MJ, Catajan MW, Patricio C, Vital JC, Gotinga TJ, Crisostomo ML, et al. Understanding healthcare providers' electronic health record (EHR) interface preferences via conjoint analysis. Int J Med Inform [Internet]. 2023 Jun 1 [cited 2023 May 23];174:105060. Available from: https://pubmed.ncbi.nlm.nih.gov/37018897/

[3] Department of Health. National Health Insurance (NHI) [Internet]. 2021 [cited 2023 May 23]. Available from: https://www.gov.za/about-government/government-programmes/national-health-insurance-0

[4] Department of Health. National Health ACT, 2003 (ACT NO 61 OF 2003) National 2021 normative standards framework for interoperability in digital health [Internet]. 2022 [cited 2023 May 23]. Available from: https://www.health.gov.za/nhi/

[5] Davison RM, Martinsons MG. Context is king! Considering particularism in research design and reporting. Journal of Information Technology [Internet]. 2016 May;31(3):241–9. Available from: http://journals.sagepub.com/doi/10.1057/jit.2015.19

[6] Gibbings R, Wickramasinghe N. Social determinants of health in the U.S.: A framework to support superior care co-ordination and leverage digital health solutions. Health Policy Technology [Internet]. 2021;10(2):100523. Available from: https://www.sciencedirect.com/science/article/pii/S2211883721000460

[7] Martin C, Thiart G, McCollum G, Roche S, Maqungo S. The burden of gunshot injuries on orthopaedic healthcare resources in south Africa. South African Medical Journal [Internet]. 2017 [cited 2023 May 23];107(7):626–30. Available from: https://pubmed.ncbi.nlm.nih.gov/29025455/

[8] Karatas M, Eriskin L, Deveci M, Pamucar D, Garg H. Big Data for Healthcare Industry 4.0: Applications, challenges and future perspectives. Expert Systems with Applications [Internet]. 2022 Aug;200:116912. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0957417422003499

[9] Department of Health. National Health Insurance Bill [Internet]. 2019. Available from: https://www.health.gov.za/wp-content/uploads/2020/11/bill-national-health-insurance-2019_compressed.pdf

[10] Luse A, Mennecke BE, Townsend AM. Selecting a Research Topic: A Framework for Doctoral Students [Internet]. Available from: http://lib.dr.iastate.edu/scm_pubs

[11] Damián Segrelles Quilis J, López-Huguet S, Lozano P, Blanquer I. A federated cloud architecture for processing of cancer images on a distributed storage. Future Generation Computer Systems [Internet]. 2023;139:38–52. Available from: https://www.sciencedirect.com/science/article/pii/S0167739X2200303X

[12] Mudaly T, Moodley D, Pillay A, Seebregts CJ. Architectural frameworks for developing national health information systems in low and middle income countries. In: Proceedings of the First International Conference on Enterprise Systems: ES 2013 [Internet]. IEEE; 2013. p. 1–9. Available from: http://ieeexplore.ieee.org/document/6690083/

[13] Boczar A, Boczar J, Mi X, Schmidt LE, Tolbert J. A case study on migrating large digital collections into a consolidated repository. The Journal of Academic Librarianship [Internet]. 2023 May 1 [cited 2023 May 23];49(3):102685. Available from: https://www.sciencedirect.com/science/article/pii/S0099133323000241

[14] Hospital Association South Africa. Our members [Internet]. 2023 [cited 2023 May 23]. Available from: https://hasa.co.za/our-members/

[15] Section27. Life Esidimeni [Internet]. 2023. Available from: https://section27.org.za/life-esidimeni/#ombud

[16] Health Ombud. The report into the circumstances surrounding the deaths of mentally ill patients: Gauteng Province [Internet]. 2022 [cited 2023 May 22]. Available from: https://section27.org.za/wp-content/uploads/2016/04/Life-Esidimeni-FINALREPORT.pdf

[17] Saurin TA, Ribeiro JLD, Vidor G. A framework for assessing poka-yoke devices. J Manuf Syst [Internet]. 2012;31(3):358–66. Available from: https://www.sciencedirect.com/science/article/pii/S0278612512000313

[18] Moonsamy W, Singh S. An investigation into electronic vaccination cards for minors in Gauteng, South Africa. In: IEEE International Multidisciplinary Information Technology and Engineering Conference (IMITEC) 2019 [Internet]. 2019 [cited 2023 May 22]. p. 452–5. Available from: https://ieeexplore.ieee.org/document/9015866

[19] NHS75. NHS architecture principles [Internet]. 2020 [cited 2023 May 23]. Available from: https://digital.nhs.uk/developer/nhs-digital-architecture/principles

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):10-14. DOI: 10.12856/JHIA-2023-v10-i2-395

Journal of Health Informatics in Africa

Re-visiting Design and Development of a Low-Cost Computer on Wheels to support healthcare delivery for Low-Resource Settings

Mtheto M. Sinjani[1*], Mwabi D. Lungu[1], Timothy M. Mtonga[1], Rashid A. Deula[1], Gerald P. Douglas[1,2]

1 - Global Health Informatics Institute, Lilongwe, Malawi
2 - Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, USA

Background and Purpose: Unlike implementations in high-income country settings which span the entire breadth of a health facility, electronic health record (EHRs) implementations in low- and middle-income countries (LMICs) have mostly been disease specific. Since most disease-specific clinics are ambulatory in nature, the inpatient setting has been largely ignored with regards to EHR implementations in LMICs. Computers on wheels (CoWs) has improved access in high-income country hospital settings, but may not be financially feasible in LMIC settings. Here we describe the design and development of a low-cost CoW that has been done in a low-income country setting.
Methods: We developed a set of functional requirements for a computer on wheels and a design approach for the development of a functional prototype from concept to finished product. We conducted a laboratory function evaluation to assess how the CoW would work.
Results: We designed and developed a CoW comprising a cart, computing platform and a docking station for charging. Our computing platform is based on a Raspberry Pi single board computer with a 7-Inch touchscreen display, thermal label printer and 2-dimensional barcode scanner. The unit is powered by a rechargeable battery providing a runtime of roughly 16 hours between charges.
Conclusions: We have demonstrated that fit-for-purpose solutions that may enhance clinical care in an in-patient setting can be designed and developed in an LMIC setting. This approach can reduce barriers to entry for EMR systems in hospitals by making more affordable and locally supportable solutions available.

Keywords: LMIC, Computer on Wheels, EMR Integration

1 Introduction

The adoption and use of Electronic health records systems (EHRs) promises several benefits and improvements to the quality of care. Since the early 2000s, several publications have described implementations of EHRs in low- and middle-income countries (LMICs). One distinguishing characteristic of EHRs implementations in LMICs has been the disease specific focus of most implementations.[1] This is unlike the implementations in high-income country settings where EHR implementations often span the entire breadth of the health system or facility.

One downside of the disease-specific focus of EHR implementation in LMICs is the ignoring of other aspects of healthcare. Since most disease-specific clinics are ambulatory in nature, the inpatient setting has been largely ignored with regards to EHR implementations in LMICs. An argument can be made that focusing on the ambulatory setting for EHR implementations in LMICs addresses most of the information needs since most patients are treated as outpatients. However, some patients treated in the ambulatory setting are admitted to hospital wards, thus requiring continuity of care among the clinical team. These are often the sickest patients whose care can benefit from well documented medical records such as those offered by EHRs in these settings. The presence of such records can facilitate delivery of healthcare and

*Corresponding author address: Global Health Informatics Institute, PO Box X111, Crossroads, Lilongwe, Malawi. Email: mtheto.sinjani@ghii.org, Tel: +265-994-633-112
© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):15-22. DOI: 10.12856/JHIA-2023-v10-i2-432

help reduce incidence of adverse events, which the World Health Organization (WHO) estimates occurs in one out of every ten hospitalised patients and are fatal in 30% of all cases.[2]

One approach that has been taken in high-income country settings is the use of computers on wheels (CoWs). These have typically taken the form of placing a desktop or laptop computer on a commercially-available cart, augmented with some form of battery power solution that enables the computer to be powered during clinical rounds. These solutions have been estimated to cost more than $3,000.[3] Jen et al. describe the development of their Very-Efficient Agile Laptop (VEAL) as a lower-cost solution for use in emergency departments in high-income country settings.[3] Their solution comprised a laptop computer, commercially available cart, additional battery to extend the runtime of the laptop, and an ID Badge reader, totalling $2,721 in cost. Whether this cost includes shipping is not stated in their description, but must be considered as a component of the cost, particularly in the context of LMIC settings. One additional observation from their description is the use of a mouse shown in pictures of their CoW, despite the fact that their laptops are equipped with trackpads, suggesting that the addition of the mouse provides an enhanced user experience.

In 2003, we developed a CoW using a mechanics rolling toolbox as the base with a modified Internet appliance mounted on top. In 2015 we developed a second iteration of the CoW that was based on a commercial cart and an android tablet.[4] In this paper, we describe the third iteration of a built-for-purpose CoW for use at the bedside along with initial findings from a pilot implementation. While our development has been done in an LMIC setting the solution also has potential applications in high-income country settings.

2 Methods

We have previously published a manifesto describing an LMIC-first approach to developing EMR systems for low-resource settings.[5] The development of our CoW follows this manifesto with five of the six themes exemplified in the prototype: 1) designing solutions optimised for use at the point of care; 2) taking a process-centric approach; 3) emphasis on low-power consumption; 4) emphasis on low cost; 5) a focus on touchscreen user interfaces to maximise usability and efficiency. The remaining LMIC-first theme focused on software development processes, which is not discussed here.

2.1 Setting

The development of the CoW has been undertaken at the Global Health Informatics Institute (GHII) training centre in Lilongwe, Malawi. GHII works at the intersection of science, engineering and global health to address problems of global health importance. GHII has a fully equipped electronics laboratory and a mechanical engineering workshop capable of doing rapid prototype development.

2.2 Functional requirements

We came up with the following list of functional requirements based on our experience working in an LMIC inpatient setting for the past several years:

1. Locally manufactured so as to minimise shipping costs and maximise maintainability.
2. Create a cart solution that provides additional value beyond just the transportation for the computer, creating a stronger value proposition for the user.
3. The footprint should be sufficiently small to navigate through doorways and between patient beds on a congested inpatient ward.
4. The cart should be sufficiently light in weight and agile to allow it to move easily in tight spaces.
5. The computer should be able to run autonomously for at least 10 hours.
6. Can be recharged easily without dependence on support staff.
7. A 2D barcode scanner to facilitate staff and patient identification.
8. A thermal label printer to allow for the generation of stickers to support different applications e.g. labelling of laboratory specimens.
9. Should provide a touchscreen user interface maximise usability and eliminate the extra space required for an external mouse.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):15-22. DOI: 10.12856/JHIA-2023-v10-i2-432

10. The computer should be firmly attached to the cart to prevent theft and the design should minimise the use of external peripherals that could be easily stolen.
11. All external USB ports should be enclosed to avoid users plugging their phones for charging that could drain the batteries, which could significantly reduce the runtime of the system.

Many of these features are core in our appliance model hardware philosophy previously described. [6]

2.3 Prototyping

We focused initially on the development of a cart starting with sketches of a proposed design followed by a mock-up made from thick cardboard. Following that, each piece was individually modelled in AutoDesk Fusion 360 computer aided design software and a tool path created to cut out each piece on a computer numerical control (CNC) router. As we expected to have several iterative improvements on the design, we used 10mm plywood for the initial construction. Once the design reached a level of maturity, we switched to PVC plastic for the final version.

Our process for developing the computing platform was heavily influenced by other internal projects at GHII that unitized single-board computers and touchscreen displays. We considered display sizes of 10-Inch, 7-Inch and 5-Inch. We selected the 7-inch display as it was more easily integrated with the Raspberry Pi, and roughly half the cost of the 10-inch display, while still providing adequate size for performing simple tasks. We chose to mount the display in portrait configuration to minimise the footprint.

Version 2 of the CoW utilised a thermal label printer that would accommodate auto a 4-inch wide label. We believe that the smaller footprint and lower power consumption of a 2-Inch label printer would be preferable and opted for the narrower printer.

In Version 2 of our CoW we used a handheld barcode scanner with a USB connection to the tablet. In our current version we switched to an embedded scanner that was internally connected. This had the additional benefit that it could be used hands-free.

Based on the dimensions of the components a cardboard mock-up of an enclosure was created, moving to plywood and finally PVC as described above. Finally, we needed to design a way of charging the carts.

2.4 Usability Testing

Our goal was to make a product that the users find value in it and have the willingness to use the product day by day. When we created the first version of the cart, we tested it at Daeyang Luke Mission Hospital. We left one cart with the matron in each of the two different departments for two weeks. After two weeks had elapsed, we went back at the hospital to gather feedback.

2.5 Laboratory Function Evaluation

To assess how long the CoW would work, we conducted a laboratory function evaluation. We developed a simple script that printed a label with a timestamp every 10 minutes. This gave us an objective way of comparing runtime of different battery technologies and power management strategies.

2.6 Ethical Considerations

The development of this manuscript and prototype followed all ethical standards for research without direct contact with human or animal subjects.

3 Results

Our Initial prototype from 2003 (original image captured with a low-res camera) along with our 2016 version and our current version are shown below in Figure 1 for comparison purposes.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):15-22. DOI: 10.12856/JHIA-2023-v10-i2-432

Figure 1 - The Evolution of our Computer on Wheels 2003, 2016 and 2023 respectively

3.1 Features of the CoW Prototype

3.1.1 The Cart

After the cart was used by the healthcare workers at Daeyang Luke Mission Hospital, they provided us with the following feedback: The cart was easy to drive around due to its compactness. The height of the cart suited different nurses, and no one complained. However, there was one comment that was raised to say the top half of the cart was wobbling so much and we addressed this problem by adding a frame with a diagonal section at the back of the cart.

While the cart is intended to support the mobility of the computer, we observed that there were other potential applications that could be concurrently addressed. In particular we noticed that clinicians often move between the bedside and the nursing station to collect medical supplies during ward rounds. To reduce the back and forth between the bedside and the nursing station, we designed the cart to facilitate the transportation of frequently needed medical supplies. In addition, waste management was also observed as a challenge. Moving between the bedside and a sharps box with a used syringe increased the chances of a needle stick. To address this, we designed a section of the cart to accommodate a sharp box. Lastly, we added a space for a waste disposal bin. To maximise mobility and ensure that the cart can navigate tight spaces, we used fully-articulated wheels on all four corners of the cart. Figure 2 shows the cart at various stages of development.

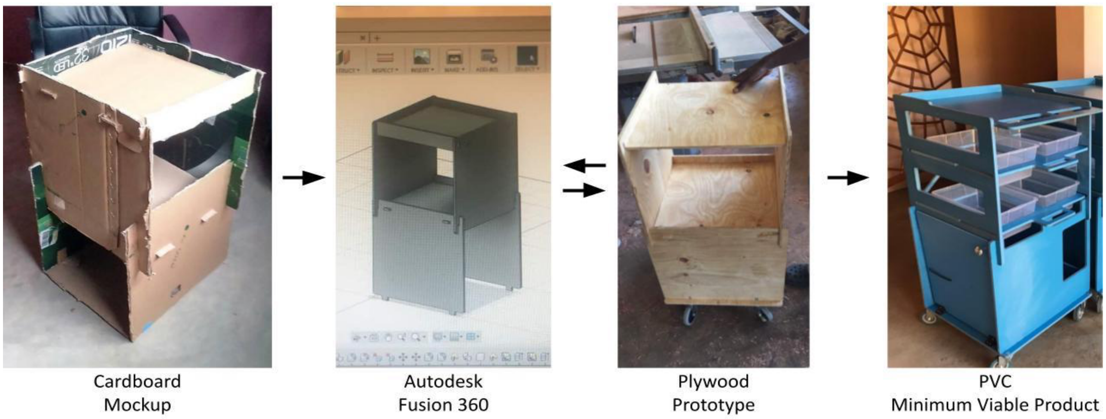


Figure 2 - The Cart at various stages of development

Our design utilised individual pieces that interlock for strength. This approach allowed us to keep the individual pieces relatively small allowing for compact shipping and assembly on-site, similar to the design

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):15-22. DOI: 10.12856/JHIA-2023-v10-i2-432

of Ikea furniture.[7] We decided to construct the cart out of PVC as it has the advantage that it is strong, water-resistant, can be easily sanitised, and does not require painting. Assembly of the pieces takes less than five minutes for two people, requiring only eight screws to hold all the pieces together.

3.1.2 The Computing Platform

We settled on using a Raspberry Pi 3B+ single board computer as the computing platform, connected to a 7-inch touchscreen display. We integrated an embedded barcode scanner and a Zebra thermal label printer with a maximum label width of 2-inches. Figure 3 shows the computing platform at various stages of development.

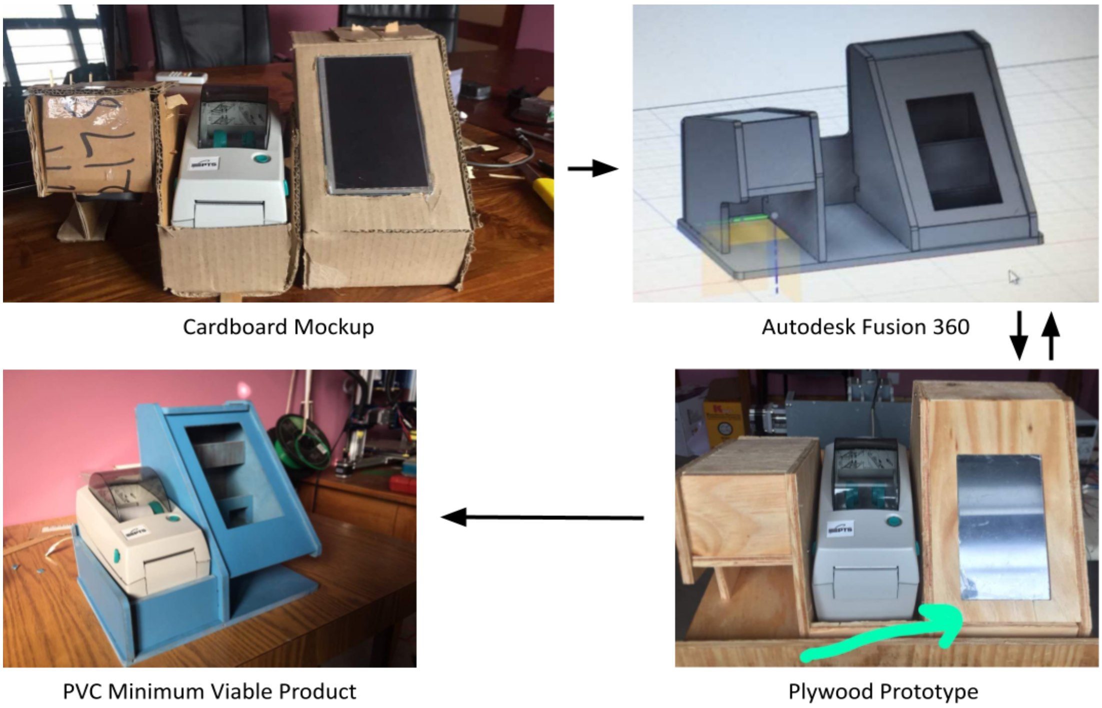


Figure 3 - The Computing platform at various stages of development

3.1.3 Component Integration

As there was a significant amount of electronics integration required, we developed a printed circuit board (PCB), which connected to the Raspberry Pi using the 40-pin GPIO header. The battery charging circuit, real-time clock, and DC/DC converters for the Raspberry Pi and label printer were all built into the PCB. The entire system is powered using a 12-volt, 8Ah Lithium Iron Phosphate (LiFePo4) battery with onboard battery management system.

3.1.4 Power management

To maximise runtime on a single charge, we implemented a number of power saving features. While we wanted to support a feature of printing adhesive labels, we knew that it may be used infrequently. However, while the printer is turned on, it drains a small amount of power even when not printing, amounting to roughly 10 percent of the overall power consumption of the CoW. To accommodate for this, we implemented a power control relay on the PCB that allowed us to turn the printer power on and off in software. With this, we were able to turn the printer on for a few seconds when we needed to print a label, and turn it off immediately after printing. To further reduce power consumption, we made changes to the raspberry Pi configuration file to disable certain unnecessary features such as Bluetooth. In addition, we adjusted the display brightness to an intensity that was still easily visible but sufficiently low to produce power savings. To ensure that the battery is as fully charged as possible at the time a user removes it from the charging station the system is entirely run from the charger leaving the battery at full capacity.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):15-22. DOI: 10.12856/JHIA-2023-v10-i2-432

3.1.5 Connectivity to Existing Clinical Systems

CoWs do not operate independently of other systems. In the majority of settings, CoWs utilise WiFi connections to interface with existing electronic medical record systems. In our experience working in LMIC hospital settings, it has been difficult to ensure complete WiFi coverage in the clinical setting. Due to the structural design of most hospitals, there are often multiple dead-zones in WiFi coverage. We used two strategies to address this information exchange issue. Our first strategy was the encoding of information (e.g., laboratory test order) into 2-dimensional (2D) barcodes that can be printed on labels. With this model, scanning the barcode at the destination (e.g., hospital laboratory) allows the information to be imported into another information system. This essentially eliminated the necessity for connectivity when information is being sent from the CoW to a destination information system. This might have traditionally been done by packaging the information in an HL7 message. We have previously described this approach as a means of health information exchange in low-resource settings.[8]

Our second strategy to minimise the need for real-time connectivity is to push information to the CoW so that no remote database query is needed at the time the user wants to access it. Provided the CoW has periodic access to WiFi it will get synchronised. In this way, it doesn't matter if the CoW happens to be in a deadzone at the time.

3.1.6 Charging Dock

We wanted a solution that was easy to connect to, could be done by the users of the cart and did not depend on a third-party to connect them for charging. Rather than having a cable connected to the cart for charging, we chose to build wall mounted "docking stations". We utilised the magnetic charger concept used on certain models of Apple laptops. Using strong magnets, we were able both hold the cart in place as well as provide the electrical connection for charging. The use of magnetically controlled switches (Reed switches) allowed us to only provide power on the charging terminals when the cart is actually docked. Rather than building separate docking stations for each cart we decided to build a station that could accommodate three carts concurrently. A picture of this "SuperDock" can be seen in Figure 4.

Figure 4 - The SuperDock CoW Charging Station with three CoWs Docked

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):15-22. DOI: 10.12856/JHIA-2023-v10-i2-432

3.2 Cost

The cost of materials to manufacture one cart is roughly $160.

The combined cost of all electronics hardware and the PVC enclosure for the computing platform is roughly $650 including the thermal printer and embedded 2D barcode scanner.

The Super-Dock station costs roughly $400 including the charging electronics capable of accommodating up to three CoWs concurrently.

3.3 Initial Findings

Below we describe some of the challenges we faced and discuss how we addressed them.

- One of the first problems we faced was a rapid reduction in battery life. We had initially designed the CoWs using a sealed lead acid battery of the type commonly used in uninterruptible power supplies for computers. When the battery was initially purchased it was able to power the CoW for roughly 8 hours. However, as they were deeply discharged every day, the battery life quickly reduced to the point where we could only achieve 20 minutes of runtime after about 4 months. To address this, we initially had two sets of batteries and swapped that at mid-day. This was followed by an iteration with Lithium-Ion cells building batteries by combining eight individual cells to make a battery. To improve performance, we added a battery management system (BMS). The final iteration used Lithium Iron Phosphate (LiFePO4) batteries with an integrated BMS. With these new batteries we are able to run a CoW for roughly 16 hours and there has been no reduction in runtime in the 18 months since we installed them.
- We observed that under poor lighting conditions the barcode scanner failed to be triggered by motion. To address this issue, we installed LEDs around the scanning area, and turned them on in software when we needed to scan.
- The first iteration of the docking station had charging connections on the side. This design made it difficult to use the cart while docked. In the redesign we moved the charging to the back.
- Version one of the carts did not have a clear indicator that it was charging once docked. This was later resolved by adding a visual indication both on the screen and with a separated LED (Red for charging green for charged).
- The cart was heavier than initially anticipated. To address this, we upgraded the diameter of the wheels to improve mobility.
- We noticed that from time-to-time carts that had been placed on charge at the docking station had become disconnected. We observed that if bumped they easily lost their charging connection. To address this we added a secondary support with an additional set of magnets.

4 Discussion

The "buy or build" question is always difficult. A solution that provides 80 percent of the benefits at 20 percent of the cost is always appealing. However, if closing the remaining gap has significant benefit then building is always the best option. As the current "state of the art" for CoWs is still based on Commercial off-the-shelf (COTS) integration of a cart and computing device, we felt this had sufficient limitations, particularly in terms of cost, we chose to opt for building a solution. While we have created a design that tightly integrates the cart and the computer, the design allows for the cart to be used without the computer where appropriate, and vice versa.

We noted that in the design of the VEAL CoW by Jen et al. they did not incorporate a printer. Since we have identified several use-cases for label printing, we believed that this could be a valuable addition to a CoW used in the clinical setting. One potential barrier to printing was accommodating the power requirements. The thermal label printer used in our design requires an external power supply similar to that used on a laptop. While the printer required 19 Volts DC to operate and the battery powering the CoW is 12-Volts, we were able to step up the voltage using a DC/DC converter. This innovation is generalizable to other CoW designs.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):15-22. DOI: 10.12856/JHIA-2023-v10-i2-432

5 Limitations and Future Work

The development of the custom CoW described in this paper does not include any formal field user effect and problem impact evaluation. This is a major limitation of this work as until the users interact with this CoW, we cannot tell how well it will work. Nonetheless, the main goal of this paper is to share the design of a low-cost CoW that can be easily assembled and maintained in low-resource settings. We believe that others can benefit from having access to this open-design and can further improve upon it. In the future, we plan to deploy the CoW in a clinical setting and conduct formal user effect and problem impact studies using our initial use-case around improving the management of laboratory test order entry and results review as has been described earlier.[4].

6 Conclusion

We have demonstrated that fit-for-purpose solutions that may enhance clinical care in an in-patient setting can be designed and developed in an LMIC setting. This approach can reduce barriers to entry for EMR systems in hospitals by making more affordable and locally supportable solutions available. Additionally the manufacturing and maintenance of the equipment bolsters the local economy, creating jobs and generally strengthening the private sector.

References

[1] M. O. Akanbi *et al.*, "Use of Electronic Health Records in sub-Saharan Africa: Progress and challenges," *J. Med. Trop.*, vol. 14, no. 1, pp. 1–6, 2012.

[2] "WHO | 10 facts on patient safety," *WHO*. http://www.who.int/features/factfiles/patient_safety/en/ (accessed Sep. 14, 2018).

[3] M. Jen, T. Cho, S. Rudkin, A. Wong, N. Almassi, and E. Barton, "Mobile CoWs (Computer on Wheels): Hamburger or VEAL?," *West. J. Emerg. Med.*, vol. 17, no. 5, pp. 527–530, Sep. 2016, doi: 10.5811/westjem.2016.6.30118.

[4] T. M. Mtonga *et al.*, "Design and implementation of a clinical laboratory information system in a low-resource setting," *Afr. J. Lab. Med.*, vol. 8, no. 1, p. 7, Oct. 2019, doi: 10.4102/ajlm.v8i1.841.

[5] C. Neumann, E. L. Dunbar, J. U. Espino, T. M. Mtonga, and G. P. Douglas, "A LMIC-First Manifesto to Developing Electronic Medical Record Systems," *J. Health Inform. Afr.*, vol. 7, no. 2, pp. 1–3, Jun. 2021, doi: 10.12856/JHIA-2020-v7-i2-300.

[6] G. P. Douglas *et al.*, "Using Touchscreen Electronic Medical Record Systems to Support and Monitor National Scale-Up of Antiretroviral Therapy in Malawi," *PLoS Med.*, vol. 7, no. 8, p. e1000319, Aug. 2010, doi: 10.1371/journal.pmed.1000319.

[7] D. Pereira, "Ikea Business Model," Aug. 10, 2022. https://businessmodelanalyst.com/ikea-business-model/ (accessed May 29, 2023).

[8] O. J. Gadabu, T. M. Mtonga, and G. P. Douglas, "Revisiting Health Information Exchange in Low Resource Settings: Modes of Health Record Transmission," presented at the HELINA 2021, Kampala, Uganda: JHIA, 2021.

[9] K. A. Grépin, "HIV Donor Funding Has Both Boosted And Curbed The Delivery Of Different Non-HIV Health Services In Sub-Saharan Africa," *Health Aff. (Millwood)*, vol. 31, no. 7, pp. 1406–1414, Jul. 2012, doi: 10.1377/hlthaff.2012.0279.

[10] H. Mbah *et al.*, "Implementing and Measuring the Level of Laboratory Service Integration in a Program Setting in Nigeria," *PLOS ONE*, vol. 9, no. 9, p. e107277, Sep. 2014, doi: 10.1371/journal.pone.0107277.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):15-22. DOI: 10.12856/JHIA-2023-v10-i2-432

Journal of Health Informatics in Africa

Role of Metrics in Medical Image Analysis based on Unsupervised Machine Learning

Eustache Muteba A.

Correspondent of International Medical Informatics Association, DR Congo

Background and Purpose: Our study addresses one of the complex aspects of machine learning techniques, namely the unsupervised machine learning applied to medical image analysis. While unsupervised learning has many benefits, there exists a main challenge that due to the fact that the result obtaining might be less accurate as input data is not labelled, and algorithms do not know the exact output in advance. Consequently, in unsupervised learning problems, several aspects must be taken into account for the validation of the algorithm results. These were the integration of measures in unsupervised learning algorithms is important. Metrics play multiple roles such as in image analysis, and in the judgment of the algorithm's performance. On the one hand, the paper describes a methodology for building an unsupervised machine learning with an application related to the analysis and diagnosis of medical images. On the other hand, it highlights the role of metrics and their mathematical properties in solving problems through unsupervised machine learning.

Methods: Considering the objectives of the medical image analysis such as to detect patterns on the image and to guide the diagnosis, we focus our interests on the unsupervised machine learning, specifically the clustering approach based the centroid and density models. These two models allow the image analysis and classification. Furthermore, we considered that classification can be done by attribution criteria such as object semantic and similarity criteria.

Results: We have demonstrated the application of unsupervised learning for medical image analysis and diagnosis. Our model was tested on 40 different images of samples. The accuracy in our unsupervised machine learning means detection and correct classification of all necessary objects in a single image.

Conclusions: Asking the machine to make a suitable grouping of the objects of an image in classes without human intervention by algorithms, this is the goal of unsupervised machine learning. To achieve this challenge, metrics play an important role. We built the formal model based on a set of rules and functions that can analyse and classifies objects' image in classes. Our model as the benefit to be used for semi-supervised machine learning co-clustering applications. This can make it easier to label a large volume of data.

Keywords: Unsupervised machine learning, classification, clustering, metrics, image analysis.

1 Introduction

Since a long time, without scientific instruments and materials the discovery and understanding of microorganisms and organisms would have been utopian. Microorganisms, in the world, are characterized by beings invisible to the human eye and which cause many health concerns for the latter, such as microbes. Seeing organisms in the body, finding any abnormality is of a great importance for human health and life expectancy.

A biomedical image is the materialization in the form of images of anatomical or functional information in vivo of parts (organs, tissues, cells) of the human body, as well as the data extracted or derived from these images. But it can also be an image obtained ex vivo (such as for example microscopy images of samples obtained by biopsy for pathological anatomy or by thick film for parasitological analysis).

The medical image corresponds to the localized measurement of a physical signal of an object or objects in space, generally in two dimensions (2D) or in three dimensions (3D). In radiologic imaging, "*the*

*Corresponding author address: Independent Research Center, P.O. Box 14769, Kinshasa, DR Congo. E-mail: emuteba@vchf.net
© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):23-28. DOI: 10.12856/JHIA-2023-v10-i2-478

intervention of the radiologist is to explore a field of view which can be either the whole object or a part of the object" [1]. As the same in microscopic analysis, the lab tech and the biologist observes different parasite forms. Doing this job manually with an *"explosion of medical data can lead to an upheaval in the workload of radiologists, limiting the time spent with the patient and increasing the rate of error in interpretation"* [2].

Thus, under the conditions of precision and real-time medicine in accordance with [3], "*the potential of medical image analysis by artificial intelligence (AI) is immense to provide faster and more reliable diagnosis to patients"*. Furthermore, "*most of the concepts used in AI come from or are inspired by neuroscience research. The concept of neural network, or layers of neurons, comes from the understanding of the organization of the cerebral cortex, particularly the visual cortex of the cat. Indeed vision is not a holistic phenomenon where visual integration would be done globally, but a hierarchical, sequential system, within which each visual region is responsible for a segmental analysis of the characteristics of the image, edge, intensity, pattern, colors, and movement*" [4].

The collaboration of imaging professionals with artificial intelligence and machine learning technologies is becoming more and more frequent, even if it is not yet very present in some health systems in southern Saharan countries. And it is quite obvious that imagists are concerned about the quality of the automatic image analysis and diagnosis provided.

As stated in [5], "*creating strong evidence for the usefulness of ML models in clinical settings is an involved process. It requires a thorough understanding of the properties of the model itself and its performance*."

Our study addresses one of the complex aspects of machine learning techniques, namely the unsupervised machine learning applied to medical image analysis. While unsupervised learning has many benefits, there exists a main challenge that due to the fact that the result obtaining might be less accurate as input data is not labelled, and algorithms do not know the exact output in advance. Consequently, in unsupervised learning problems, several aspects must be taken into account for the validation of the algorithm results. These were the integration of measures in unsupervised learning algorithms is important. Metrics play multiple roles such as in image analysis, and in the judgment of the algorithm's performance.

On the one hand, this article describes a methodology for building an unsupervised machine learning with an application related to the analysis and diagnosis of medical images. On the other hand, the article highlights the role of metrics and their mathematical properties in solving problems through unsupervised machine learning.

2 Materials and methods

There are different types of machine learning: supervised, unsupervised, and reinforcement. In particular our study focused on unsupervised learning that derive insights directly from the data itself and have capability to uncover patterns in an unlabelled dataset.

Since it is about the analysis of images, specifically of 2 dimensions, it is therefore assumed a geometric surface and a domain containing objects. Objects are characterized by their shape or structure, by their position in the surface and also by appearance. Moreover, the rules for assembling objects in the domain are based on notions such as inclusion and proximity. This is why having an image analyzed by a machine, in general, requires the integration of geometric measurements.

As mentioned earlier in the introduction, the following is a methodology of building an unsupervised machine learning for an application a medical image analysis and an automatic diagnosis. Behind, there are descriptions of metrics to be used which demonstrate their roles. The work published in the following articles, [6] [7] [8] [9], constitutes a basic part of our research.

And in the aspect of testing the performance of our system, we resorted to other metrics such as those of the Kolmogorov-Smirnov test [10] and normalized cross-correlation [11].

2.1 Unsupervised Machine Learning

The unsupervised machine learning model [12] has often three steps namely: data collection, building and training model, and at least the evaluation. In general, unsupervised machine learning has two model: clustering algorithm and association algorithm.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):23-28. DOI: 10.12856/JHIA-2023-v10-i2-478

The present study refers mainly to clustering. The clustering [13] [14] technique aims to discover meaningful structure or to identify patterns in the input data and then put it in a specific class/cluster.

2.1.1 Clustering Model

The clustering analysis of an image aims is to identify which of a set of categories a pattern belongs to. Moreover, a clustering algorithm maps input data to a category which can be a disease. The use of maths or metrics in clustering algorithms is mandatory to perform the clustering problem.

The theories presented in [15] on clustering models fixed our issues. Considering the objectives of the analysis of a medical image such as to detect patterns on the image and to guide the diagnosis, we pay attention to two models namely: the centroid and density models.

a) Centroid Model

K-means algorithm is one of the centroid based clustering algorithms. As stated in [16], "*when the data space X is R^D and using Euclidean distance, we can represent each cluster by the point in data space that is the average of the data assigned to it. Since each cluster is represented by an average, this approach is called K-Means*".

We illustrate the problem in the following way.

As the image space is X. Let consider X be R^D. Moreover, the data set is $\{x_n\}_{n=1}^{N}$. In this set of data belonging to the space X, the algorithm searches for points or elements having a particular interest based on a given invariant or threshold. These different points of interest are considered as centroids.

Let data (x_i), $i=1, 2, \ldots, n,$ measured on n independent observations. It comes that $d_{ii'}$, denotes the distance between observations i and i'. Ultimately, the Euclidean distance d is:

$$d = \sqrt{\sum_{i=1}^{n}(x_i - x_{i'})^2} \quad (1)$$

The algorithm should also, in the meantime, calculate an area of interest around the central point using density model. This should make it easier to calculate the distance around the centroid.

From this point of view and in the medical context, the designer of the machine learning is supposed to define beforehand the clusters which correspond, for example, to a pathology. K represents the number of clusters C. Thus it is possible to determine n-K clusters as prototypes. But due to letting the machine has a self-learning, we formalize K in the following open interval (n-K to m-K); m-K is a discovered prototype.

$$K = [\![1..n \implies K = \begin{cases} c_{1=k_1} \\ \ldots \\ c_m = k_m \end{cases}$$

b) Density Model

Intuitively, the density of a given surface is understood as the fraction of the total number of elements over the size of the surface. As stated in [17] [18] [19], the usual density measurement model will have to involve two ingredients, namely a local density estimate at each point and a connection between objects. From the above, we formulate:

$$\rho_i = \sum_j X(d_{ij} - d_c)$$
$$X(x) = \begin{cases} 1, x < 0 \\ 0, x \geq 0 \end{cases} \quad (2)$$

where ρ_i is a local density, d_{ij} is the Euclidean distance between data point i and data point j, and d_c is the cutoff distance; χ is the weighting function for estimating the influence of a neighboring point.

2.1.2 Classification

Image classification is a very common task in healthcare computer vision problems. And Machine learning are frequently employed for classification.

Let a class be defined by a set of properties which are both necessary and sufficient for membership in the class. Starting from the principle that the descriptions of the objects to be classified, whatever the initial format, are translated into geometric representations where each object is associated with characteristics. Thus, we ask ourselves the question: knowing, for a given object o, these characteristics $f(o_i)$, can we determine the class c_i to which this o belongs?

Knowing that unsupervised learning problems the labels in the input data are typically unknown.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):23-28. DOI: 10.12856/JHIA-2023-v10-i2-478

Two methods presented above, can help to solve the problem, since they are used to cluster the data together and to provide information about the structure of the data. This is the class identification.

Furthermore, classification can be done by attribution criteria such as object semantic and similarity criteria.

a) Object Semantic

The first criteria is the domain-object relationship. An object corresponds to a series of domains, with more or less features or characteristics defined as an object semantic.

$$f(o_i) = \{f_i : f_i \in D_j^{k+1}\} \tag{3}$$

b) Similarity

The 2nd criteria is a similarity. A classification always involves much more arbitrariness than does a representation of data in a geometric continuum, with, as coordinates, the factors extracted by a suitable analysis. Despite the fact that once the machine has discovered the semantic values of the object, it can proceed to a comparison with the theoretical models and prototypes of the intrinsic classes of the machine.

In our model uses a numerical value called the discrimination threshold. This threshold determines attribution of an object to a class. Let formalize below the rule of attribution:

$$S = \begin{cases} if\ 0 \rightarrow \text{"Similar"} \\ elseif\ [1..10] \rightarrow \text{"Bit similar"} \\ elseif > 10 \rightarrow \text{"Different"} \end{cases} \tag{4}$$

where *S* is the percentage of similarity; If two images are similar, their similarity percentage will equal to 90-100%, that is <10, but if different usually >30, that is [0-70%].

3 Results

We have demonstrated the application of unsupervised learning for image analysis for medical diagnosis. This application was produced as part of a project “Virtual Community of Healthcare Facilities” and relates to a medical decision support system for the management of malaria [20] [21].

3.1 Experimentation of Unsupervised Machine Learning for Automatic Diagnosis of Plasmodium

3.1.1 Data Collection

The data were collected from the Division of Parasitic Diseases and Malaria (DPDM) of the Centers of Disease Control and Prevention (CDC). There were two modes of acquisition namely: downloaded directly via the CDC website and in a zip format via email sent by CDC. It is constituted of 40 blood image samples that are anonymous and under agreement of use.

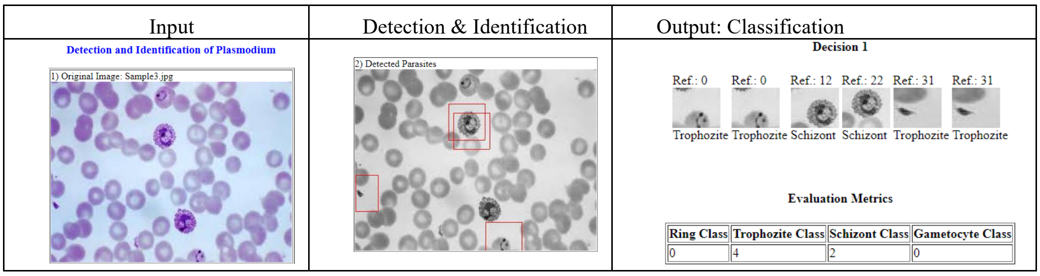
3.1.2 Pre-analytic Phase: Data Characteristics

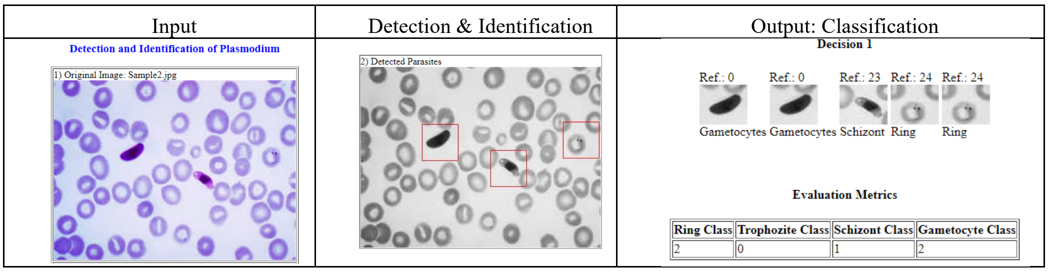
In laboratory image analysis, geometrical models of particle play an important role in the diagnosis of diseases. Theoretically, we can distinguish different geometric shapes based on some characteristic of the shapes. The experimentation concerns the specie Plasmodium Falciparum. Geometrically, the different morphologies take by the Plasmodium falciparum can be the following: dots, ring, circle and crescent. These the different morphologies are belonging to all staged of Plasmodium namely: ring, trophozoite, Schizont and gametocyte. Moreover, they are designated as clusters (prototypes).

3.1.3 Analytic Phase: Experimentation of the Model

Based on our formal model, we implemented our machine learning algorithms. We present in the figures below the execution of our machine learning and the results of the tests carried out on two images (https://www.vchf.net/vchf/simula/vlab/simulaBrowser.php).

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):23-28. DOI: 10.12856/JHIA-2023-v10-i2-478





3.1.4 Evaluation Metrics

The table below shows the indicators of performance and demonstrates the learnability of the machine learning algorithms. The accuracy in our unsupervised machine learning means detection and correct classification of all necessary objects in a single image.

Table 1: Indicators of Performance

Image Quantity	40 jpeg images (2D)		
Preparation technique	Giemsa-stained thick blood sample		
Quality	Some images have been affected by chemical added substances or others impurities		
Objects to find	Ring, trophozoite, schizont, gametocyte		
Processing time	<1 min by image		
Results	*Accuracy*	*Specificity and sensitivity (Confusions)*	*Inaccuracy*
	30%	***67,5%***	***2,5%***

4 Discussion

Within the scope of our study, we have formalized and implemented an unsupervised machine learning. The use of machine learning is often justified by the complexity of the problem such as that of decision support and in particular the analysis of images for a diagnosis.

Asking the machine to make a suitable grouping of the objects of an image in classes without human intervention by algorithms, this is the goal of unsupervised machine learning. To achieve this challenge, metrics play an important role.

The formal model presented is based on a set of rules and functions that analyse and classify objects' image in classes.

The model was tested on 40 different images of thick film samples. All these images present different objects with different configurations. The evaluation on these 40 images gives an accuracy of 30%, 67.5% with a difference in precision (mixture of true positives and negative confusions) and 2.5% completely inaccurate. We have considered that the accuracy of unsupervised machine learning is ability to detect and classify correctly of all necessary objects in a single image.

In view of these results, one might be tempted to say that it is always very complex to carry out an unsupervised machine learning. However, it is rather a question of model and therefore of the metrics used that must be refined. It is ultimately important to have knowledge of the mathematical properties of a metric to determine its relevance for a given task.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):23-28. DOI: 10.12856/JHIA-2023-v10-i2-478

It is often difficult to evaluate the performance of a model in unsupervised learning, since the true labels are not known. Furthermore, there are extrinsic factors to the model, namely the quality of the image, which often depends on the acquisition chain.

Taking into account these factors, our model borrowed the principles mentioned in [22] [23], on the Probably Approximately Correct (PAC) Learnability and constructive function approximation.

And if we already validate our model, it may well be useful for semi-supervised machine learning co-clustering applications. This can make it easier to label a large volume of data.

References

[1] Durand E, Blondiaux E. Caractéristiques d'une image médicale. Elsevier Masson SAS; 2017.

[2] Haidar Ahmad H, Hamid Salih MA, Luangrath A, Sede T, Zugaj L. Apports et limites de l'Intelligence Artificielle pour l'Imagerie Médicale. 2022. DOI: https://doi.org/10.34746/kbkc-cd39.

[3] Sun R, Deutsch E, Laure Fournier L. Intelligence artificielle et imagerie médicale. Bulletin du Cancer. 2022;vol(109-1):83-88. DOI: https://doi.org/10.1016/j.bulcan.2021.09.009.

[4] Brunelle F, Brunelle P. Intelligence artificielle et imagerie médicale: définition, état des lieux et perspectives. Bull Acad Natl Med. 2019; vol(203): 683 - 687.

[5] Hicks SA, Strümke I, Thambawita V, Hammou M, Riegler MA, Halvorsen P, Parasa S. On evaluation metrics for medical applications of artificial intelligence. medRxiv. 2021. DOI: https://doi.org/10.1101/2021.04.07.21254975.

[6] Palacio-Nino J-O. Evaluation Metrics for Unsupervised Learning Algorithms. arXiv. 2019. DOI: https://doi.org/10.48550/arXiv.1905.05667.

[7] Aziz SR, Khan TA, Nadeem A. Inheritance metrics feats in unsupervised learning to classify unlabeled datasets and clusters in fault prediction. PeerJ Comput Sci. 2021. DOI: https://doi.org/10.7717/peerj-cs.722.

[8] Müller D, Soto-Rey I, Kramer F. Towards a guideline for evaluation metrics in medical image segmentation. BMC Research Notes. 2022. DOI: https://doi.org/10.1186/s13104-022-06096-y.

[9] Reinke A, Maier-Hein L. Common limitations of performance metrics in biomedical image analysis. Medical Imaging with Deep Learning. ARXIV. 2021. https://arxiv.org/abs/2104.05642.

[10] Demidenko E. Kolmogorov-Smirnov Test for Image Comparison. In: Computational Science and Its Applications; ICCSA 2004 International Conference; Assisi, Italy. 2004. DOI: https://doi.org/10.1007/978-3-540-24768-5_100

[11] Handayan HH. Digital Image Matching Method using Normalized Cross-Correlation. GEOID. 2010;Vol(6-1). Available from: https://iptek.its.ac.id/index.php/geoid/article/download/7346/4886.

[12] Javatpoint. An Introduction to Unsupervised Learning. Available from: https://www.javatpoint.com/unsupervised-machine-learning

[13] Joshi S. What is Clustering in Machine Learning: Types and Methods. 2022. Available from: https://www.analytixlabs.co.in/blog/types-of-clustering-algorithms/

[14] Wikipedia. Cluster analysis. Wikipedia. 2023. Available from: https://en.wikipedia.org/wiki/Cluster_analysis

[15] Gavita Regunath G. 10 Incredibly Useful Clustering Algorithms You Need To Know. 2022. Available from : https://www.advancinganalytics.co.uk/blog/2022/6/13/10-incredibly-useful-clustering-algorithms-you-need-to-know

[16] Ryan P. Adams RP. K -Means Clustering and Related Algorithms. Elements of Machine Learning. Princeton University. 2018. Available from: https://www.cs.princeton.edu/courses/archive/fall18/cos324/files/kmeans.pdf

[17] Campello R, Kröger P, Sander J, Zimek A. Density-based clustering. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2020;vol(10-2). DOI: https://doi.org/10.1002/widm.1343

[18] Sun Z, Qi M, Lian J, Jia W, Zou W, He Y, Liu H, Zheng Y. Image Segmentation by Searching for Image Feature Density Peaks. Appl. Sci. 2018;vol(8-969). DIO: https://doi.org/10.3390/app8060969

[19] Heredia LCC, Mora AR. Density-based clustering methods for unsupervised separation of partial discharge sources. Electrical Power and Energy Systems. 2019; vol(107):224–230.

[20] Muteba AE. Towards Networked eHealth: OMaT Project (Phase I). Journal of Health Informatics in Africa. 2014;vol(2-1).

Muteba AE. Toward a Virtual Community of Healthcare Facilities: Virtual Lab. Proceedings of the EFMI Special Topic Conference 2016, IOS Press. 2016; 221.

Shalev-Shwartz S, Ben-David S. Understanding Machine Learning: From Theory to Algorithms. USA: Cambridge University Press;2014.

[21] Naisse JP. L'Approximation Analytique – Vers Une Théorie Empirique Constructive et Finie. Bruxelles: Université de Bruxelles;1992.

© 2023 JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. J Health Inform Afr. 2023;10(2):23-28. DOI: 10.12856/JHIA-2023-v10-i2-478

Journal of Health Informatics in Africa

© 2023 JHIA

ISSN: 2197-6902
DOI: http://dx.doi.org/10.12856/JHIA-2023-v10-i2

Publisher
Koegni eHealth, Innovation for Development e.V. Germany
Postfach 652166
D-22372 Hamburg, Germany
www.koegni-ehealth.org
E-mail: info@koegni-ehealth.org

HELINA is the Africa Region of the International Medical Informatics Association (IMIA)
www.helina-online.org
www.imia-medinfo.org
www.jhia-online.org