Conceptualizing open health data platforms for low- and middle income countries

Daniel Kapitan
PharmAccess Foundation, Eindhoven University of Technology
daniel@kapitan.net

Femke Heddema
PharmAccess Foundation

Julie Fleischer
PharmAccess Foundation

Chris Ihure
PharmAccess Kenya

Antragama Abbas
Delft University of Technology

Steven Wanyee
IntelliSOFT

XXX
ONA

John Grimes
Australian e-Health Research Centre

Mark van der Graaf
PharmAccess Foundation

Mark de Reuver
Delft University of Technology

Abstract This document is only a demo explaining how to use the template.

Introduction

Digital platforms vs. data platforms for healthcare

It is a widely held belief that digital technologies have an important role to play in strengthening health systems in low- and middle income countries (LMICs), as exemplified by the WHO global strategy on digital health [1]. The adoption rate of mobile phones in LMICs has been an important driver in implementing digital health solutions [2]. Yet, there are many shortcomings and challenges, including the current fragmentation of digital platforms and the lack of clear-cut pathways of scaling up digital health programmes, such that they can support sustainable and equitable change of national health systems in LMICs [2]–[4].

A commonly used perspective to scrutinize digital health is to consider it as a digital platform [5]. Digital platforms have disrupted many sectors but have just started to make inroads into highly regulated industries such as healthcare [6]. In this light, the challenges faced by LMICs in establishing national digital health platforms have a lot in common with those faced by high income countries. From a technological perspective, interoperability issues, weak integrations, siloed data repositories and overall lack of openness are often reported as key impediments [7], [8]. From a societal perspective, issues pertaining to the winner-takes-all nature of digital platforms are hotly

debated as many jurisdictions make work to ensure these platforms indeed serve the common good of achieving universal health coverage [9].

Case studies on digital platforms in healthcare point to an emerging pattern where the focus shifts from the digital platform with its defining software and hardware components, to the data as the primary object of interest in and of itself [6], [10]. This observation ties into the proposed research agenda by de Reuver et al. to consider data platforms as a phenomenon distinct from digital platforms [11]. Generally, data platforms inherit the characteristics of digital platforms. For example, the economic perspective on digital platforms stresses their multi-sidedness with developers and consumers, while data platforms are used by data owners, data consumers and third party solution providers. At the same time, data platforms differ as their main offerings revolve around data. From a market perspective, data platforms have more moderate network effects and are more susceptible to fragmentation and heterogeneity.

Particularly relevant in the context of health data platforms (HDPs), is the conceptualization of openness. The objective of openness, the ways to realize openness and the risks of opening up can be considered as elements within the ongoing debate on data spaces [12] and data solidarity [13]–[15] in the healthcare domain, which is a current issue for high income countries and LMICs alike. Openness is particularly relevant if we are to realize a solidarity-based approach to health data sharing that i) gives people a greater control over their data as active decision makers; ii) ensures that the value of data is harnessed for public good; and iii) moves society towards equity and justice by counteracting dynamics of data extraction [12].

From health information exchanges to health data platforms

This paper is motivated by the conflation of a number of developments relevant to the design and implementation of solidarity-based HDPs in LMICs. First, the OpenHIE framework [16] has been adopted by many sub-Saharan African countries [17] as the architectural blueprint for implementing nation-wide health information exchanges (HIE), including Nigeria [18], Kenya [19] and Tanzania [20]. These countries have, as a matter of course, extended the framework to include "analytics services" as an additional domain. The rationale for this addition is to facilitate primary and secondary re-use of health data for academic research, real-world evidence studies etc. which can be framed within the context of ongoing efforts towards Findable, Accessible, Interoperable and Reusable (FAIR) sharing of health data [21]. In doing so, however, we have implicitly moved from conceptualizing digital health platforms (the original OpenHIE specification) to health data platforms. This is problematic because the notion of openness, which is assumed to be essential in establishing solidarity-based approaches to data sharing, is inherently different for a data platform compared to a digital platform.

Conceptually, the OpenHIE framework constitutes a framework for an open digital platform. Openness for digital platforms refers to i) the use of open boundary resources, that is, specifications for the various healthcare specific workflows and information standards such as FHIR; and ii) the use of open source components that are available as digital public goods[22]. If we are to use the OpenHIE framework as an open data platform, we need to extend the standards, technologies and architecture to include functionality for data sharing and re-use. Distinguishing four types of data sharing (Table 1), the purpose of this paper is to investigate how new standards

and technologies that can establish openness of health data platforms can be integrated into the OpenHIE architecture framework. The lack of detailed specifications and consensus of this addition to OpenHIE currently stands in the way of development projects that aim to establish HDPs in LMICs.

Table 1: Types of data sharing and in relation to new standards and technology enablers to create openness.

	Type of data sharing	**New standards and technology enablers to create openness**
1	Data at the most granular (patient) level, which is persisted and used to provide a longitudinal record.	**Bulk FHIR API** which has by now been incorporated in all major FHIR implementations the FHIR standard can be readily used to support type 1 data sharing to patient-level data across a patient population [23], [24].
2	Aggregated data, for example statistics for policy evaluation and benchmarking	**SQL-on-FHIR specification** [25]: provides a standardised approach to make FHIR work well with familiar and efficient SQL engines that are most commonly used in analytical workflows. Builds on FHIRPath [26] expressions in a logical structure to specify things like column names and unnested items. Implementations of this approach are available or forthcoming, including open source implementations such as Pathling [27] and commercial offerings like Aidbox [28].
3	Data analytics modules, that provide access to work and access the data.	**Federated learning (FL) and privacy- enhancing technologies (PETs)** [29]–[31]: new paradigms that address the problem of data governance and privacy by training algorithms collaboratively without exchanging the data itself. Models can be trained on combined datasets and made available as open source artifacts for decision support. Data analysts can use FL and PETs to work with the data in a collaborative, decentralized fashion.
4	Trained models that have been derived from the data and can be used stand-alone for decision support.	^ ^

Methods

In the paper, we present a design that extends the OpenHIE specification to include the four types of data sharing mentioned above. Using the full-STAC approach [8] we combine open standards, open technologies and open architectures into a coherent modular HDP platform that can be configured and re-used across a variety of use-cases. Subsequently, we employ the Framework for Evaluation of Design Science (FEDS) research [32] to evaluate the design. The purpose of this evaluation is formative: we aim to to help improve the outcomes of the design and the artifacts presented. We do so by taking a naturalistic approach, that is, the expected performance of the design is considered with its real environment where we draw on our experience in two use-cases, namely the Momcare programme and the ONA Canopy data platform.

As part of our design research, we have taken a narrative approach in surveying existing scientific studies on health data platforms, focusing on the seminal reports and subsequently searching forward citations. In addition, we have searched the open source repositories (most notably GitHub) and the online communities (OpenHIE community, FHIR community) to search for relevant open standards, technologies and architectures. This paper should not be considered as a proper systematic review.

Abbreviations

FAIR	Findable, Accessible, Interoperable and Reusable
FHIR	Fast Healthcare Interoperability Resources
HDP	Health data platform, explicitly differentiated from health digital platform
HIE	Health Information Exchange
FL	Federated learning
PET	Privacy-enhancing technologies
LMIC	Low- and middle income countries

Bibliography

[1] World Health Organization, “Global Strategy on Digital Health 2020-2025”, Nov. 2021.

[2] J. McCool, R. Dobson, R. Whittaker, and C. Paton, “Mobile Health (mHealth) in Low- and Middle-Income Countries”, *Annual Review of Public Health*, vol. 43, no. 1, pp. 525–539, 2022, doi: 10.1146/annurev-publhealth-052620-093850.

[3] W. H. Organization, “Recommendations on Digital Interventions for Health System Strengthening”, 2019.

[4] T. Neumark and R. J. Prince, “Digital Health in East Africa: Innovation, Experimentation and the Market”, *Global Policy*, vol. 12, no. S6, pp. 65–74, Jul. 2021, doi: 10.1111/1758-5899.12990.

[5] M. de Reuver, C. Sørensen, and R. C. Basole, “The Digital Platform: A Research Agenda”, *Journal of Information Technology*, vol. 33, no. 2, pp. 124–135, Jun. 2018, doi: 10.1057/s41265-016-0033-3.

[6] H. Ozalp, P. Ozcan, D. Dinckol, M. Zachariadis, and A. Gawer, “``Digital Colonization'' of Highly Regulated Industries: An Analysis of Big Tech Platforms' Entry into Health Care and Education”, *California Management Review*, vol. 64, no. 4, pp. 78–107, Aug. 2022, doi: 10.1177/00081256221094307.

[7] K. Malm-Nicolaisen, R. Pedersen, and A. J. Fagerlund, “Exploring the Emergence of Open Platforms in Healthcare: Design Considerations and Experiences from an Initial Case in Norwegian Primary Care”, in *56th Hawaii International Conference on System Sciences*, 2023.

[8] G. L. Mehl *et al.*, “A Full-STAC Remedy for Global Digital Health Transformation: Open Standards, Technologies, Architectures and Content”, *Oxford Open Digital Health*, vol. 1, p. oqad18, Jan. 2023, doi: 10.1093/oodh/oqad018.

[9] T. Sharon, “When Digital Health Meets Digital Capitalism, How Many Common Goods Are at Stake?”, *Big Data & Society*, vol. 5, no. 2, p. 2053951718819032, Jul. 2018, doi: 10.1177/2053951718819032.

[10] C. Alaimo and J. Kallinikos, “Organizations Decentered: Data Objects, Technology and Knowledge”, *Organization Science*, vol. 33, no. 1, pp. 19–37, Jan. 2022, doi: 10.1287/orsc.2021.1552.

[11] M. de Reuver, H. Ofe, W. Agahari, A. E. Abbas, and A. Zuiderwijk, “The Openness of Data Platforms: A Research Agenda”, in *Proceedings of the 1st International Workshop on Data Economy*, in DE '22. New York, NY, USA: Association for Computing Machinery, Dec. 2022, pp. 34–41. doi: 10.1145/3565011.3569056.

[12] B. Otto, M. Ten Hompel, and S. Wrobel, Eds., *Designing Data Spaces: The Ecosystem Approach to Competitive Advantage*. Cham: Springer International Publishing, 2022. doi: 10.1007/978-3-030-93975-5.

[13] “The Lancet and Financial Times Commission on Governing Health Futures 2030: Growing up in a Digital World - The Lancet”.

[14] B. Prainsack, S. El-Sayed, N. Forgó, Ł. Szoszkiewicz, and P. Baumer, “Data Solidarity: A Blueprint for Governing Health Futures”, *The Lancet Digital Health*, vol. 4, no. 11, p. e773-e774, Nov. 2022, doi: 10.1016/S2589-7500(22)00189-3.

[15] B. Prainsack and S. El-Sayed, “Beyond Individual Rights: How Data Solidarity Gives People Meaningful Control over Data”, *The American Journal of Bioethics*, vol. 23, no. 11, pp. 36–39, Nov. 2023, doi: 10.1080/15265161.2023.2256267.

[16] “OpenHIE Framework”.

[17] A. L. Mamuye *et al.*, "Health Information Exchange Policy and Standards for Digital Health Systems in Africa: A Systematic Review", *PLOS Digital Health*, vol. 1, no. 10, p. e118, Oct. 2022, doi: 10.1371/journal.pdig.0000118.

[18] "Nigeria Data Exchange Architecture for the National Data Repository".

[19] "Kenya Health Information Systems Interoperability Framework".

[20] "Tanzania Health Enterprise Architecture". 2020.

[21] P. Guillot, M. Bøgsted, and C. Vesteghem, "FAIR Sharing of Health Data: A Systematic Review of Applicable Solutions", *Health and Technology*, vol. 13, no. 6, pp. 869–882, Nov. 2023, doi: 10.1007/s12553-023-00789-5.

[22] "Digital Public Goods Alliance".

[23] K. D. Mandl *et al.*, "Push Button Population Health: The SMART/HL7 FHIR Bulk Data Access Application Programming Interface", *npj Digital Medicine*, vol. 3, no. 1, pp. 1–9, Nov. 2020, doi: 10.1038/s41746-020-00358-4.

[24] J. Jones *et al.*, "A Landscape Survey of Planned SMART/HL7 Bulk FHIR Data Access API Implementations and Tools", *Journal of the American Medical Informatics Association*, vol. 28, no. 6, pp. 1284–1287, Jun. 2021, doi: 10.1093/jamia/ocab028.

[25] "SQL on FHIR Speciification v0.0.1-Pre".

[26] "FHIRPath (Normative Release)".

[27] J. Grimes, P. Szul, A. Metke-Jimenez, M. Lawley, and K. Loi, "Pathling: Analytics on FHIR", *Journal of Biomedical Semantics*, vol. 13, no. 1, p. 23, Sep. 2022, doi: 10.1186/s13326-022-00277-1.

[28] "SQL on FHIR".

[29] N. Rieke *et al.*, "The Future of Digital Health with Federated Learning", *npj Digital Medicine*, vol. 3, no. 1, pp. 1–7, Sep. 2020, doi: 10.1038/s41746-020-00323-1.

[30] J. Scheibner *et al.*, "Revolutionizing Medical Data Sharing Using Advanced Privacy-Enhancing Technologies: Technical, Legal, and Ethical Synthesis", *Journal of Medical Internet Research*, vol. 23, no. 2, p. e25120, Feb. 2021, doi: 10.2196/25120.

[31] S. Jordan, C. Fontaine, and R. Hendricks-Sturrup, "Selecting Privacy-Enhancing Technologies for Managing Health Data Use", *Frontiers in Public Health*, vol. 10, p. 814163, Mar. 2022, doi: 10.3389/fpubh.2022.814163.

[32] J. Venable, J. Pries-Heje, and R. Baskerville, "FEDS: A Framework for Evaluation in Design Science Research", *European Journal of Information Systems*, vol. 25, no. 1, pp. 77–89, Jan. 2016, doi: 10.1057/ejis.2014.36.