# Opening-up digital platforms to accommodate patient-generated healthcare data

Dragana Paparova
University of Agder
*dragana.paparova@uia.no*

**Abstract.** This paper investigates the process of opening-up digital platforms to accommodate patient-generated healthcare data (PGHD) and argues that in data platforms, barriers arise due to the entanglement of technology and policy. The empirical setting for the study is the opening up of a Norwegian eHealth platform for PGHD captured by external vendor technologies. The possibility to accommodate PGHD opens up new innovation arenas by recombining data from multiple sources and actors. However, such process is prone to a unique set of challenges when innovation is centered around data, instead of functionalities, such as: 1) open-up the data core using boundary resources; 2) control patient data across long chains of actors; 3) establish uniform rules to co-create data value. The findings show that the proves of opening up data platforms faces legislative barriers which should be overcome in a way that technology and policy enable each-other.

## Introduction

Novel sources of patient-generated healthcare data (PGHD) captured through medical devices, sensors and smartphone apps are entering the healthcare landscape, holding the potential to transform the way patient information is "generated, collected and analyzed in healthcare practices and used in clinical decision making" (Grisot et al., 2020). PGHD can support the needs for circular interaction between patients and healthcare professionals, changing the role of patients from passive recipients to active "prosumers" (Barrett et al., 2016), as

they consume and produce data using dispersed devices and under diverse circumstances. This opens up new arenas for data innovation by interconnecting patient data across a variety of actors and from dispersed sources, which can be recombined along multiple value-creation pathways.

One way of supporting innovation across multiple actors is by organizing their interactions around digital platforms. Digital platforms are underpinned by modular architectures, interconnecting core and peripheral modules using boundary resources as standardized interfaces. The boundary resources support the use and re-use of common components and can facilitate innovation on a larger scale by transferring design capabilities to external actors (Ghazawneh and Henfridsson, 2013). However, opening up the boundary resources can also lead to greater fragmentation and loss of control over the emerging innovation network in the platform periphery (Gawer and Cusumano, 2014). Therefore, the boundary resources need to be governed in a way that balances the trade-offs between expanding the platform with complementary components, but at the same time keeping control by setting up uniform rules, standards and shared institutional logics which govern the behavior of distributed actors (Autio and Thomas, 2020; Eaton et al., 2015)

In data platforms instead, the boundary resources connect the core and periphery, but those relations are established based on digital data, rather than functionalities, tools or applications (Tempini, 2017). The common core modules are full of data, and the external actors innovate with these data in the platform periphery (Bonina and Eaton, 2020). However, data are not components and do not embody functions in the same way as components do (Alaimo et al., 2020). Rather, data are captured as events, computed into tokens and then assigned meaning when they are used by actors, due to their semantic nature. Therefore, innovation in data platforms does not necessarily follow the same recombinant logic of modular architectures, as it happens at a more granular levels than assembling together a set of components. Although data innovation does not happen independently of the components, actors recontextualize these data across their value-creation processes and assign them meaning, rather than constructing a functionality (Aaltonen et al., 2021). Thus, data innovation distinguishes from recombinant innovation with modules, as it takes place in the way data is ported and used across actors' value-creation trajectories (Alaimo et al., 2020).

Research by Grisot et al. (2020) already shows that the re-combination of components can lead to multiple alternative value pathways around PGHD by aligning the underlying digital infrastructure and work practices to tailor the diverse needs for data exchange between patients and healthcare professionals. Besides this architectural perspective, research also reveals the complexity of data-based value-creation due to their use-agonistic character. For example, Tempini (2017) shows the multidimensionality of PGHD in the value-creation trajectories of actors organized around a social media platform. Similarly, Barrett et al. (2016) follows the evolution of an online community, and discloses the tensions stakeholders face as their goals and the meaning they assign to data change over time. However, while previous research has been focusing on the

generation of patient data using approved devices in an outpatient clinic (Grisot et al., 2020), and in peer-support communities (Barrett et al., 2016; Tempini, 2017), I hereby put my focus on the process of opening up the boundary resources to support larger scale innovation centered around PGHD from devices and technologies developed by external actors.

The research question I seek to address is: *what are the barriers in opening-up digital platforms to accommodate patient-generated healthcare data?* The empirical case follows the Norwegian national eHealth platform and reveals barriers in the process of opening up the data core for PGHD generated by external actors' components. The paper is organized as follows. In the methods section, I elaborate on the methodology used, introduce the case and explain the data analysis process. Next, I introduce the main findings of the study and elaborated on them in more details in the discussion section. At last, I present the main conclusions of the conducted study.

# Research Methodology

## Case description

The empirical setting for this study is the Norwegian healthcare context. The case is conducted by including: 1) a national eHealth platform in Norway "HelseNorge", which undergoes a process of opening up its boundary resources for external development and accommodate PGHD; as well as 2) three private vendors solutions for PGHD, which are part of the public digital ecosystem, but are not integrated with the national platform. HelseNorge provides citizens access to information stored about them in several health registries in the public sector. The platform was launched June 15th 2011, driven by the need to create a single point of entry portal for citizens, instead of letting them "search Google for health" across many websites. The platform has grown considerably throughout the years, integrating with several systems across the public infrastructure, including 11 approved digital health apps which are part of its tool catalogue. Currently, HelseNorge is undergoing a process of opening up its boundary resources for external development and incorporate PGHD as part of its ecosystem (Directorate for eHealth, 2019).

The case also incorporates three digital health tools for PGHD. 1) Mobile medical record system which is built jointly for patients and healthcare professionals and shares medical patient data using secure messaging, video consultation, photos, forms and other health information, stored in the cloud. 2) Shared patient diary, a solution for information exchange between healthcare professionals, service recipients and their relatives. The app shares information in the form of text, images, and video includes a common calendar for real-time updates and combines both medical and lifestyle data about a patient. 3) Outpatient clinic tool, used for specialist care which collects structured medical data about patients,

stored systematically in hospital systems, from where healthcare professionals can extract the data and monitor patient's health status.

## Data collection and analysis

The study was explorative and conducted using a qualitative method (Sarker et al., 2018a, 2018b). The data was gathered via 10 semi-structured interviews, using a snowball approach where participants recommended potential suitable candidates further. 7 interviews were conducted with representatives from the national HelseNorge platform, and 3 interviews with private vendors. To gain more insights into the context, online information, including websites, presentations, and strategy documents were also collected. The data gathering process lasted for 5 months, June-Nov 2020, although the case was followed in retrospect dating back to 2010, when the development of the national eHealth platform started to take place. Participants were technical and managerial staff, working with the national eHealth platform and private software vendors. The background of participants varies between: software architects, software developers/medical doctors, lawyers, consultants, data scientist, providing a wider perspective over the cross-disciplinary nature of the research problem. The interviews lasted for approximately one hour and were afterwards transcribed to analyze the data. The data gathering process was guided by the research interest on capturing and interpreting informants' meanings, and their understanding of the decisions that need to take place towards opening up the national platform to incorporate PGHD (Dubois and Gadde, 2002). The interview guides included questions on the challenges for extending HelseNorge's core functionalities with external solutions, the role of policy in making decisions about incorporating PGHD within the ecosystem and opinions on what is the way forward towards integrating PGHD as part of the ecosystem.

The data was analyzed in an abductive and iterative way (Alvesson and Sköldberg, 2009; Dubois and Gadde, 2002). The data collection and analysis were informed by existing theoretical concepts in the digital platforms literature (Aanestad et al., 2017; Ghazawneh and Henfridsson, 2013), which were used to initially define and re-articulate the research problem (Dubois and Gadde, 2002). The theoretical concepts informed the data gathering process, and were used to categorize the empirical data, coordinate the findings, as well as to direct and redirect the study as new insights emerged. The analysis moved iteratively between asking questions, generating the findings, making comparisons with existing knowledge and refining this again. The data was coded to establish categories for grouping the information gathered from the empirical work (Maxwell and Miller, 2008) and analyzed from the perspective of the platform owner. The findings were first organized around 6 key decisions of opening up the platform for PGHD and grouped around 3 barriers which need to be overcome throughout this process. Although the data collection was triangulated with information from official documentation and strategy documents by the Directorate of eHealth, these documents were only used as a first step to provide a contextual understanding over the case, and to inform the data collection process.

However, they were not systematically analyzed to generate the findings of the study, as the findings were bases solely on the empirical data.

# Findings

The findings are organized around three main barriers for opening up the digital platform to accommodate PGHD captured through external components, defined as: 1) open up the data core using boundary resources; 2) control patient data across long chains of actors; 3) establish uniform rules to co-create data value. The findings are elaborated in more details as follows.

## Open-up the data core using boundary resources

HelseNorge.no was launched on 15<sup>th</sup> of June 2011, as a "Citizen Portal" working as a single point of entry for Norwegian citizens, which were previously "searching Google for health" across many websites. At the beginning some team members thought that platform-thinking could be beneficial in the long-run. However, this vision was abandoned due to the pressing deadline for delivering the first version and all functionalities were developed in-house. After the initial launch, discussions have been on-going on what functionalities should be supported next, what parts should the core connect with and which components should be developed externally. So far, decisions about expanding HelseNorge were supported when the team would identify a functionality they need and then 1) build it inside; 2) get a third-party vendor to build it; 3) or integrate with an existing vendor solution on the market providing these functionalities. Such an approach has resulted with many vendor-specific APIs adapted to the requirements and functionalities of the external solutions. What is currently set in place for vendors who want to connect with HelseNorge is a requirements list for certain areas, such as: video conferencing, appointments, journal patient health record, message exchanges between healthcare professionals and patients. Therefore, HelseNorge as of now has a very limited set of APIs exposed to third-parties, very few consumers that send data packages to it and even fewer which retrieve data and are creating solutions based on access to data from the core.

The lack of published APIs has resulted with some of the vendors having to "self-resource" boundary resources themselves in order to enter the ecosystem. Others decided to step back from the integration process, due to delays in API provision, followed by bureaucracy and large documentation. *"It would be much easier if the system was ready to share APIs with the private companies like us, if there was a system and rules that are there for that and APIs. We have a feeling that we have to fight for every access we get. (...) If there were more information and advice from the government on how to do things and more APIs ready from the beginning, then we would do things differently and we would save a lot of resources, money and time." (Informant, Vendor 1).* Decisions about opening up HelseNorge are currently made around two options. 1) Contained environment, where smaller vendors who need a stable platform to support them can re-use a

lot of the core capabilities, with lower thresholds for innovation and fast innovation cycles. 2) Uncontained environment, in which external actors have a stable standalone platform, build all the functionalities themselves and connect to HelseNorge. However, the technical side of publishing APIs is considered "the easiest", as challenges arise since APIs need to validate the party which gets access to patient data, verify the user and make sure that patient data is handled in accordance to policy and law. Therefore, these APIs need to work as digitalized contracts between multiple parties which do not only guide what functionalities to be developed, but also what data elements are processed, and how to regulate the behavior of multiple parties across the ecosystem.

## Control patient data across long chains of actors

Initially, data in HelseNorge was stored in a single storage solution called the Personal Health Archive (PHA). Although there were discussions about data being stored with responsible entities interacting with each-other, versus laying in a single storage, the sense for platform architecture which should do connections instead of storage was not there from the beginning. As HelseNorge started integrating with GP ERP and hospital EHR systems, patient data was getting exchanged outside of HelseNorge's control. Therefore, decisions had to be made on how to govern patient data exchanged across multiple systems. It was decided that once citizens are handed over to another party, the respective systems are the owners of this data and take the responsibility as a data controller. Therefore, HelseNorge only provides access to view this data, and in some cases stores a copy in the PHA. *"Many people think 'Helsenorge knows a lot about me', but we are not allowed to read this data. (...) HelseNorge is saying 'we know that you have some data at the health thrust, we cannot open it, but we can help you see it', so we cannot snatch the information on the way to the user" (Informant, HelseNorge).* Currently, a lot of the data provided by HelseNorge is not owned by the platform. The platform is organized as consent-based and takes responsibility as a data processor only for the data processed inside its core components. Patient consent has to be registered in HelseNorge and the patient can stop sharing at any time. Although citizens can choose to use different digital health tools via HelseNorge, they have to accept the terms of use in the particular tool. This is due to the lack of control HelseNorge has over how patient data is handled throughout the entire chain of actors and whether this process is legal all the way.

One possible approach towards controlling the chain is the aspiration of the HelseNorge team for having a "dedicated HelseNorge law", which states that the platform is the official national provider of healthcare services. In the absence of law which might provide an independent basis for data processing, they have to find complex solutions within the restrictions given by legal regulations. The lack of control has made it challenging to keep consent valid at all time and make sure that it is still within the scope of the time patients provided it. *"What I have experienced lately is that you should not completely rely on the definition of the terms in the GDPR, with regards to who is the controller, but you have to look at the whole chain to be able to see what is there, what is in factual circumstances that the working part has expressed. You have to see who is in the factual*

*situation the closest to take the control or responsibility in a complex chain."*
*(Informant, HelseNorge).* Therefore, it is challenging to provide clarity of responsibilities when data is exchanged across many actors who often fail to understand what their legal responsibility is and what the terminology means.

As data is re-written and re-copied across multiple systems, the data also gets stored in decomposed solutions across these long chains. Although the team is assertive that "the time has passed for single storage solutions", all parties wish access to data, but they all avoid the responsibility to store it due to the strict legislation. Smaller vendors either keep data in the cloud to enhance scalability or store it on a third-party server and public actors do not want to use data kept in a storage they cannot trust. *"If the patients think that this is critical data and upload it in a storage, they think that someone will look into this. But that is not the case if you do not have an agreement in it, because no doctor will look into it before there is an appointment, or a reason for looking at it. So, it is a dangerous misconception if the patient thinks that 'I gave this to the healthcare service in Norway, someone should react if something is wrong' (Informant, HelseNorge).* Therefore, discussions are on-going on whether HelseNorge should provide a storage solution for vendors who do not want to store data themselves, and at the same time allow flexibility for larger vendors who want to store and keep data in their own storage solution.

## Establish uniform rules to co-create data value

Decisions about opening up HelseNorge also need to encompass establishing criteria on which PGHD captured by external vendors are allowed into the platform ecosystem. Such decisions need to be informed by setting universal rules and standards which reflect patient data privacy and security in the digital health tools developed by external vendors. At the time HelseNorge was launched, most of the standards which are in use today were either non-existing or premature. From 2016, the team started working with HL7 FHIR with the aim to standardize data exchange by not bringing in too much data, but keep it as small as possible, yet still within the clinical context. However, such standards are not mandatory for all vendors in the public infrastructure, leading to lack of understanding on what standards do different solutions for PGHD have to comply with. *"They {the municipality} said you need to have high security level, because it is sensitive data, and when we did and lost all of our users, I remember I called {the person} and I said: 'you asked us to do this, and now you have to tell your workers that it is safe to log in with bankID, because we are losing everyone'." (Informant, Vendor 2).*

One way of dealing with this fragmented portfolio is for the platform to establish universal criteria for evaluating digital health tools for PGHD. In 2015, the Directorate of eHealth started working on a framework for assessing digital health tools to make sure they bring benefit to the healthcare service and are safe to use in clinical practice, but such attempts have been dropped. What is currently set in place as a "screening procedure" to assess the external tools are testing and approvement queries to verify that the tool is compliant with GDPR as well as

follows the Code of conduct for information security in the health sector in Norway. The tools also need to get a confirmation from the Directorate of eHealth that the content is clinically responsible. After that, they sign an agreement for third-party data processing. HelseNorge also has a publicly published tool catalogue as a library of digital health tools which can be prescribed to the citizens. However, the aim is for privacy and security to be embedded across all digital health tools in the ecosystem, instead of assessing individual cases. One-way forward is to provide self-declarations issued by an authenticated governmental body which approves the use of the PGHD tool and finds it to be trustworthy for patient treatment. However, as of yet, such process is still not set in motion.

# Discussion

Research already shows that innovation facilitated by digital platforms can result with "unpredictable innovative contributions by large, uncoordinated audiences" (Autio and Thomas, 2020), but this paper goes further by showing the added complexity of opening up when the re-combinative innovation is centered around data, instead of functionalities. Innovation in data platforms starts with boundary resources, but the way core patient data are used is more difficult to control due to their semantic nature. This shifts the focus of data innovation from the modular architectures in which they are generated, towards the ways actors use them across their value-creation trajectories. The findings show that the combined effects of innovating with functionalities and data (Tempini, 2017) challenge the process of opening up the boundary resources, as data also bring in the emergent role of policy and legislation. In data platforms, the opening up of boundary resources does not solely mean transferring design capabilities to external actors (Ghazawneh and Henfridsson, 2013), but they also need to manifest themselves as "invisible data rights" in the background. This creates difficulties in regulating access over data and controlling how data is used and re-used by external actors in the platform periphery.

Once data leave the core, they can be re-copied across multiple actors which assign them meaning. Due to data's use agonistic nature, it is hard to track the data interactions across all actors, as the owners of the components are not necessarily the owners of data. In actors' value trajectories, data decouple from the components that carry them and can be re-copied and assigned meaning on top of those components (Aaltonen et al., 2021; Alaimo et al., 2020). Thus, data can be re-used across long chains of actors where the partitioning of data rights is not always as straightforward once data leave their natural source. Instead, there is a need to look at the actual chains to determine who holds responsibility for what and how data rights are dispersed. Current laws such as GDPR do provide certain clarity on the roles of data controllers and processors, but when opening up for innovation on a large scale, data do not always flow from one actor to another, but across long chains of actors. This brings a new set of challenges on how to regulate the patient data use and keep the chains legal all the way.

The inability to control these long chains works as a barrier to support external data innovation. The lack of control rises as it is hard to determine a-priori how data will be used and track the interactions of data across actors' value creation trajectories. Data are not components and do not embody functions, but are recombinant resources which acquire meaning as they are collected, stored and used by actors (Alaimo et al., 2020). The way data are used, recombined and aggregated can trigger a new set of interactions which are not reflected in the existing law. Therefore, instead of thinking about complying with legislation when opening up the boundary resources in a linear way, the actors also need to make sure that the process of data use and re-use is legal across the chains in an on-going manner. Such a turbulent environment requires that new laws are created to provide more clarity, as well as old laws are constantly revised to reflect the changed circumstances. Therefore, the lack of clarity on how to reflect legislation in complex environments where data is exchanged across multiple actors, works as a constraint to data innovation.

Ensuring compliance with legislation and controlling the long chains can also be facilitated by setting up upfront criteria on which data bring value to the platform and its ecosystem. Such criteria can orchestrate actors by making sure that verified technologies, provided by legitimate actors catch relevant data which is of value to the healthcare service. However, the lack of universal criteria and legislative instruments to regulate the ecosystem actors in such an automated way, also suppresses the opportunities for patient data innovation on a larger scale. This paper contributes to the literatures on PGHD and digital platforms by showing that although platforms' modular architectures can facilitate the process of accommodating PGHD in digital platforms, innovations centered around data entangle technology and policy. Although the paper shows that accommodating PGHD in eHealth data platforms may be premature at the stage, the potential is promising towards that direction.

## Conclusion

This paper investigates the process of opening-up an eHealth platform to accommodate PGHD captured through external vendor technologies and shows that barriers arise due to the entanglement of technology and policy. The data-intensive environment brings in an increasing complexity in regulating data use and re-use across the long-chains of actors in the periphery, which suppress the innovation potential with patient data in the platform periphery. This research also has certain limitations, as the empirical study is based on a limited number of interviews and the case chosen is one where PGHD from external technologies is still not accommodated in the digital platform. Further research can go beyond by showing how to tackle the interplay of technology and policy and do that in a way that enables innovation with PGHD in data-intensive environments.

# References

Aaltonen, A., Alaimo, C., Kallinikos, J., 2021. The Making of Data Commodities: Data Analytics as an Embedded Process. Journal of Management Information Systems 47. https://doi.org/10.1080/07421222.2021.1912928

Aanestad, M., Grisot, M., Hanseth, O., Vassilakopoulou, P., 2017. Information Infrastructures and the Challenge of the Installed Base, in: Aanestad, M., Grisot, M., Hanseth, O., Vassilakopoulou, P. (Eds.), Information Infrastructures within European Health Care, Health Informatics. Springer International Publishing, Cham, pp. 25–33. https://doi.org/10.1007/978-3-319-51020-0_3

Alaimo, C., Kallinikos, J., Aaltonen, A., 2020. Data and value. Handbook of Digital Innovation.

Alvesson, M., Sköldberg, K., 2009. Reflexive Methodology: New Vistas for Qualitative Research, Second edition. ed. SAGE Publications Ltd, Los Angeles ; London.

Autio, E., Thomas, L.D.W., 2020. Value co-creation in ecosystems: insights and research promise from three disciplinary perspectives. Handbook of Digital Innovation.

Baldwin, C.Y., Woodard, C.J., 2009. The Architecture of Platforms: A Unified View. Platforms, markets and innovation.

Barrett, M., Oborn, E., Orlikowski, W., 2016. Creating Value in Online Communities: The Sociomaterial Configuring of Strategy, Platform, and Stakeholder Engagement. Information Systems Research 27, 704–723. https://doi.org/10.1287/isre.2016.0648

Bonina, C., Eaton, B., 2020. Cultivating open government data platform ecosystems through governance: Lessons from Buenos Aires, Mexico City and Montevideo. Government Information Quarterly 37, 101479. https://doi.org/10.1016/j.giq.2020.101479

Directorate for eHealth, 2019. Sharing health data with the patient.

Dubois, A., Gadde, L.-E., 2002. Systematic combining: an abductive approach to case research. Journal of Business Research 55, 553–560. https://doi.org/10.1016/S0148-2963(00)00195-8

Eaton, B., Elaluf-Calderwood, S., Sørensen, C., Yoo, Y., 2015. Distributed Tuning of Boundary Resources: The Case of Apple's iOS Service System. MIS Quarterly 39, 217–244.

Gawer, A., Cusumano, M.A., 2014. Industry Platforms and Ecosystem Innovation: Platforms and Innovation. J Prod Innov Manag 31, 417–433. https://doi.org/10.1111/jpim.12105

Ghazawneh, A., Henfridsson, O., 2013. Balancing platform control and external contribution in third-party development: the boundary resources model: Control and contribution in third-party development. Information Systems Journal 23, 173–192. https://doi.org/10.1111/j.1365-2575.2012.00406.x

Grisot, M., Lindroth, T., Islind, A., 2020. Digital Infrastructures for Patient Centered Care: Examining Two Strategies for Recombinability. pp. 289–300. https://doi.org/10.1007/978-3-030-23665-6_21

Maxwell, J.A., Miller, B.A., 2008. Categorizing and connecting strategies in qualitative data analysis, in: Handbook of Emergent Methods.

Sarker, S., Xiao, X., Beaulieu, T., Lee, A.S., 2018a. Learning from First-Generation Qualitative Approaches in the IS Discipline: An Evolutionary View and Some Implications for Authors and Evaluators (PART 2/2). JAIS 19, 909–923. https://doi.org/10.17705/1jais.00512

Sarker, S., Xiao, X., Beaulieu, T., Lee, A.S., 2018b. Learning from First-Generation Qualitative Approaches in the IS Discipline: An Evolutionary View and Some Implications for Authors and Evaluators (PART 1/2). JAIS 19, 752–774. https://doi.org/10.17705/1jais.00508

Tempini, N., 2017. Till data do us part: Understanding data-based value creation in data-intensive infrastructures. Information and Organization 27, 191–210. https://doi.org/10.1016/j.infoandorg.2017.08.001