Empowering Healthcare Actors to Contribute to the Implementation of Health Data Integration Platforms: A Retrospective of the medEmotion Project

Marcel Parciak, Noëlla Pierlet, Liesbet M Peeters

Submitted to: Journal of Medical Internet Research
on: October 28, 2024

Table of Contents

Empowering Healthcare Actors to Contribute to the Implementation of Health Data Integration Platforms: A Retrospective of the medEmotion Project

Marcel Parciak[1, 2*] MSc; Noëlla Pierlet[1, 2, 3*]; Liesbet M Peeters[1, 2, 4]

[1]Biomedical Research Institute UHasselt - Hasselt University Diepenbeek BE
[2]Data Science Institute UHasselt - Hasselt University Diepenbeek BE
[3]Ziekehuis Oost-Limburg Genk BE
[4]University Multiple Sclerosis Center UHasselt - Hasselt University Diepenbeek BE
[*]these authors contributed equally

Corresponding Author:
Marcel Parciak MSc
Biomedical Research Institute
UHasselt - Hasselt University
Agoralaan
Diepenbeek
BE

Abstract

Background: Health data integration platforms are vital to drive collaborative, interdisciplinary medical research projects. Developing such a platform requires input from different stakeholders. Managing these stakeholders and steering platform development is challenging, misaligning the platform to the partners' strategies might lead to a low acceptance of the final platform.

Objective: We present the medEmotion project, a collaborative effort between seven partners from healthcare, academia and industry to develop a health data integration platform for the region of Limburg in Belgium. We focus on the development process and stakeholder engagement, aiming to give practical advice for similar future efforts based on our reflections on medEmotion.

Methods: We introduce Personas to paraphrase different roles that stakeholders take and Demonstrators that summarise personas' requirements with respect to the platform. Both the personas and the demonstrators serve two purposes. First, they are used to define technical requirements for the medEmotion platform. Second, they represent a communication vehicle that simplifies discussions between all stakeholders.

Results: Based on the personas and demonstrators, we present the medEmotion platform based on components from the Microsoft Azure cloud. The demonstrators serve as use cases that showcase the utility of the platform. We reflect on the development process of medEmotion and distil takeaway messages that will be helpful for future projects.

Conclusions: Investing in community building, stakeholder engagement, and education is vital to building an ecosystem for a health data integration platform. Instead of academic-led projects, the healthcare providers themselves ideally drive collaboration among healthcare providers. The providers are best positioned to address hospital-specific requirements, while academics take a neutral mediator role. This also includes the Ideation Phase, where it is vital to ensure the involvement of all stakeholders. Finally, balancing innovation with implementation is key to developing an innovative yet sustainable health data integration platform.

(JMIR Preprints 28/10/2024:68083)
DOI: https://doi.org/10.2196/preprints.68083

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Empowering Healthcare Actors to Contribute to the Implementation of Health Data Integration Platforms: A Retrospective of the medEmotion Project

Viewpoint

Marcel Parciak[1,2,*], Noëlla Pierlet[1,2,3,*], Liesbet M Peeters[1,2,4]
[1]Biomedical Research Institute, UHasselt - Hasselt University, Diepenbeek, Belgium
[2]Data Science Institute, UHasselt - Hasselt University, Diepenbeek, Belgium
[3]Ziekenhuis Oost-Limburg, Genk, Belgium
[4]University Multiple Sclerosis Center, UHasselt - Hasselt University, Diepenbeek, Belgium
[*] these authors contributed equally

Abstract

Background: Health data integration platforms are vital to drive collaborative, interdisciplinary medical research projects. Developing such a platform requires input from different stakeholders. Managing these stakeholders and steering platform development is challenging, misaligning the platform to the partners' strategies might lead to a low acceptance of the final platform.

Objective: We present the *medEmotion* project, a collaborative effort between seven partners from healthcare, academia and industry to develop a health data integration platform for the region of Limburg in Belgium. We focus on the development process and stakeholder engagement, aiming to give practical advice for similar future efforts based on our reflections on medEmotion.

Methods: We introduce *Personas* to paraphrase different roles that stakeholders take and *Demonstrators* that summarise personas' requirements with respect to the platform. Both the personas and the demonstrators serve two purposes. First, they are used to define technical requirements for the medEmotion platform. Second, they represent a communication vehicle that simplifies discussions between all stakeholders.

Results: Based on the personas and demonstrators, we present the medEmotion platform based on components from the Microsoft Azure cloud. The demonstrators serve as use cases that showcase the utility of the platform. We reflect on the development process of medEmotion and distil takeaway messages that will be helpful for future projects.

Conclusions: Investing in community building, stakeholder engagement, and education is vital to building an ecosystem for a health data integration platform. Instead of academic-led projects, the healthcare providers themselves ideally drive collaboration among healthcare providers. The providers are best positioned to address hospital-specific requirements, while academics take a neutral mediator role. This also includes the Ideation Phase, where it is vital to ensure the involvement of all stakeholders. Finally, balancing innovation with implementation is key to developing an innovative yet sustainable health data integration platform.

Keywords: Data Science; Health Data Integration; Health Data Platform; Real-World Evidence.

Introduction

Accurate and well-formatted data is key to delivering high-quality healthcare and fueling medical research [1–3]. All healthcare actors acquire real-world data, defined as any healthcare-related information captured from the patient [4]. The volume, velocity and variety of acquired data, however, raises challenges for data processing systems [5]. Data engineers work in interdisciplinary environments to ensure that users receive data in the right format, at the right time and place to generate real-world evidence [4]. Health data engineering is a complex and time-consuming task that cannot be managed without IT solutions tailored to environment-specific requirements [6]. Consequently, health data integration platforms are a hot topic in medical informatics research [7–11]. In this paper, we present our design and development approach to building such a platform.

The medEmotion project was a collaborative initiative between three hospitals (Jessa Ziekenhuis, Noorderhart, Ziekehuis Oost-Limburg), two academic institutions (Hogeschool PXL, UHasselt) and two industry partners (LRM, BioVille) from the region of Limburg, Belgium. The primary objective of medEmotion was to establish a comprehensive data integration platform designed to effectively address a wide array of health data-related inquiries and challenges encountered by various stakeholders within the healthcare ecosystem. The medEmotion platform aggregates medical data from hospitals and general practitioners, personal health information collected through wearable devices, and environmental data. It enables researchers, healthcare professionals, and entrepreneurs to test innovative data-based concepts and generate real-world evidence. We built the medEmotion platform to be secure and compliant with legal standards to be fit for purpose for real-world health data. The platform enhances decision-making and stakeholder collaboration by facilitating comprehensive analysis of integrated real-world health data.

In this paper, we outline the medEmotion project, focusing on our chosen development approach. In particular, we describe a team of personas and a set of demonstrators. Personas are fictional stakeholders whose requirements and unique perspectives we implemented into the project. Demonstrators describe real-world use cases that guided the definition of functional requirements of the medEmotion platform. Each demonstrator's requirements are affected by one or more personas.

Furthermore, we discuss our approach, combining personal reflections on our challenges during medEmotion. With these reflections, we reminisce about our experiences that we believe present common pitfalls of healthcare-related projects that handle real-world data. We aim to shed light on the tension between stakeholders from healthcare, academia, industry, and investment societies and give advice for future efforts.

The paper is structured as follows. We introduce the personas and the demonstrators of medEmotion alongside a summary of its implementation. Afterwards, we discuss our development approach and the result. We end with distilling key takeaway messages from the medEmotion project.

Personas

Stakeholder management is a core project management task that aligns the requirements and visions of stakeholders. A stakeholder may be part of the core project team or take a consultant role, advising the project team regularly. From our experience, most stakeholders combine multiple personas. Each stakeholder might take on a different "hat", sharing their view and arguments in different, situation-specific ways to drive the project forward. Therefore, we define six personas that represent the different "hats" a stakeholder may represent in the following.

- *Data Protection Officer.* The Data Protection Officer persona is concerned with keeping information up to date, private, secure and traceable. These concerns can be enforced by law, such as the general data protection regulation, or can be intrinsic in the sense that keeping data well governed makes aspects such as data security and contract management easier. People who represent this persona are data protection officers of hospitals, contract managers, or data stewards.

- *Researcher.* The Researcher persona is concerned with conducting research more efficiently or easily. In health data integration, conducting research typically means running experiments on health data. Thus, the researcher aims to get easier access to health data, larger amounts of health data, more computational resources, or access to software tools that accelerate data processing. A common example of a researcher is a PhD student.
- *Healthcare Practitioner.* The Healthcare Practitioner is concerned with delivering the best healthcare quality to the patient. Health data integration commonly refers to finding relevant and rich patient information on time with appropriate interfaces. Clinicians and nurses represent this persona, for example.
- *IT Expert.* The IT Expert persona is concerned with efficiently and reliably running the hospital information system. The IT-security aims of confidentiality, integrity, and availability are of major concern, which results in a more reluctant attitude towards new and unknown systems. The hospital's chief information officer represents this persona.
- *Hospital Board Member.* The Hospital Board Member persona is concerned with the strategic advancement of the hospital within the resources available. As such, this persona will wage strategic and long-term goals, such as participating in international research studies, against short-term available resources, such as the workforce and budget of a hospital. For example, a hospital's chief executive officer represents this persona.
- *Sponsor.* The Sponsor persona is concerned with the success of the investment. In a research project such as medEmotion, this includes ensuring that project resources are well-spent and securing potential return on investments about the project results. This persona is represented by project sponsors, such as venture capital, pharma or insurance companies.

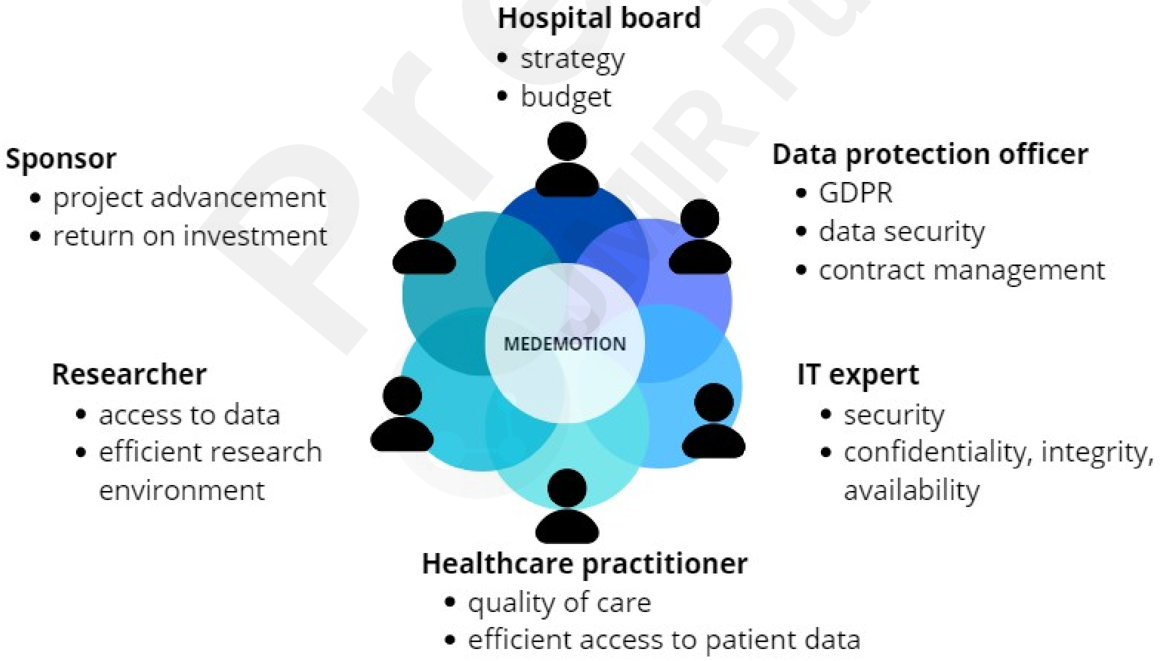


Figure 1: personas involved in the different demonstrators, with some of their focus points indicated.

During medEmotion, we found that defining personas and assigning them to our stakeholders made discussions more transparent and efficient. The set of people participating in discussions and meetings changes frequently during a project. A list of predefined personas allowed us to ensure we included all distinct perspectives to decide on the next steps. In particular, we assigned these personas to demonstrators to gain well-structured functional requirements, which we will list in the following section.

Demonstrators

We defined a series of real-world demonstrators to facilitate the requirement engineering process. We present a short description of each defined demonstrator and the personas involved.

- *Orchestration of complex pipelines.* In biomedical research projects, (para-)clinical data must be processed through complex data integration steps. Researchers often implement these steps ad-hoc using different programming languages (e.g. Python, R, Matlab), and software frameworks (e.g. Torch, Pandas, dplyr) or reuse existing implementations as black boxes [12]. The resulting software code is complex and difficult to automate, orchestrate and reproduce, pushing it out of reach for IT experts. Further, Data Protection Officers see a lack of in-depth documentation, essentially losing real-world data due to missing knowledge. With this demonstrator, we showcase the platform's capabilities to orchestrate complex data integration pipelines automatically.

- *Private research environments.* Access to real-world data greatly increases the efficiency of exploratory research tasks. With strict privacy regulations such as the General Data Protection Regulation of the European Union or the Health Insurance Portability and Accountability Act of the United States of America, data access is often achieved by sharing datasets after bilateral contracts between the data custodians (e.g. hospitals) and universities are set in place [13]. This process, however, is slow, complex, and lacks transparency, creating tension between researchers, data protection officers and IT experts. Researchers need to deal with tedious administration tasks, while data privacy officers and IT experts are concerned with keeping sensitive data secure. With this demonstrator, we showcase the platform's ability to create private research environments that allow researchers to access data, tooling and compute resources while keeping real-world health data secure.

- *OHDSI OMOP CDM.* Transnational medical research can be facilitated with internationally standardised data schemas, such as the OMOP Common Data Model (CDM). Transforming healthcare records to OMOP is a highly operational task, enabling data custodians to participate in federated, transnational research projects without sharing sensitive datasets [14]. Transforming and maintaining data structured according to the OMOP CDM is desirable for researchers and hospital board members, as it allows for simple access, privacy perseverance and increased hospital visibility in the academic landscape. With this demonstrator, we showcase the platform's ability to facilitate transforming data to and maintaining data in the OMOP CDM, including running a suite of software tools developed by the OHDSI or EHDEN communities.

- *Visualization of integrated datasets to support care.* Providing high-quality healthcare necessitates large amounts of integrated, patient-centric data that facilitate informed decision-making. Analyzing these extensive amounts of patient data and gathering insights on the evolution of the patient's medical conditions over time in a few minutes is only possible with data-rich visualizations [15]. The clinical decision-making process remains healthcare practitioner-specific as well as multi-disciplinary. Hence, visualizations must be adaptable per individual and still sharable within a multidisciplinary team to stay helpful. Privacy remains essential in the eyes of data protection officers, hence, sophisticated methods to share information according to therapeutic relationships between healthcare practitioners and patients are needed. With this demonstrator, we showcase the platform’s ability to facilitate interactive visualizations of rich datasets that can be tailored to individual needs and shared amongst multi-disciplinary teams.

- *Data ingestion from hospital data.* The primary sources of real-world data meant to generate real-world evidence are hospital information systems (HIS). Even for clinical studies that collect data with study-specific case report forms, HIS data helps provide important context information, such as demographics or comorbidities [1]. Due to the value of HIS data, one of the sponsor's main targets is integrating such data into the platform. This real-world data is highly sensitive and the prime target of data protection measures set up by data protection officers and IT experts. Researchers aiming to access the data naturally clash with data protection measures, as they tend to over-request data. With this demonstrator, we showcase the platform's ability to ingest data from multiple pseudonymized data sources.
- *Data ingestion from devices.* With eHealth solutions on the rise, data integration platforms receive more interest from sponsors to include data collected outside the traditional boundaries such as the hospital or the practice. Patients collect data themselves, either actively using health apps or passively using wearable devices [16], both using lifestyle and software-as-medical device apps [17]. Providing richer datasets for clinicians and researchers, IT experts face challenging data integration tasks of high volume and erroneous-prone data streams. Of course, these data streams are sensitive and thus are a concern of data protection officers. With this demonstrator, we showcase the platform's ability to include data streams from home monitoring devices that are not part of traditional hospital information systems.
- *Connection with high-performance computing centres.* High-performance computing resources are needed to enable researchers to use machine-learning approaches [18]. Those can ramp up costs quickly when acquired from public cloud providers such as Microsoft Azure, Google Cloud or Amazon Web Services. Many academic institutions operate high-performance computing (HPC) centres to meet the demand for HPC. To keep research projects using machine learning feasible, data transfer to and from academic HPC centres needs to be simple while retaining IT security standards set by Data Protection Officers. With this demonstrator, we showcase the platform's ability to transform data between the Vlaams Supercomputer Center, an HPC centre operated by and for the five Flemish universities.

The demonstrators served as user stories to guide discussions and the implementation of the medEmotion platform. Based on the personas, we could identify which stakeholders to include in discussions. Hence, we were able to have more ad-hoc discussions in smaller groups, replacing large and time-consuming meetings to discuss the entirety of the platform components. Combining personas and demonstrators allowed for a rapid implementation process, which we summarize in the following.

Implementation

We implemented the medEmotion platform together with an industry partner using components of Microsoft's Azure Cloud Platform. The platform's architecture foresees a general environment and multiple use case-specific data silos. We find modules and tools in a general environment that handle pseudonymised, cleaned and integrated data. In the hospital-specific data silos, we find modules and tools to pre-process datasets either for the general environment or for hospital-specific projects. All data is pseudonymised when entering the platform with data source-specific keys, creating a data mesh to ensure data privacy. The platform retains encrypted quasi-identifiers to allow for ad-hoc record linkage if needed.

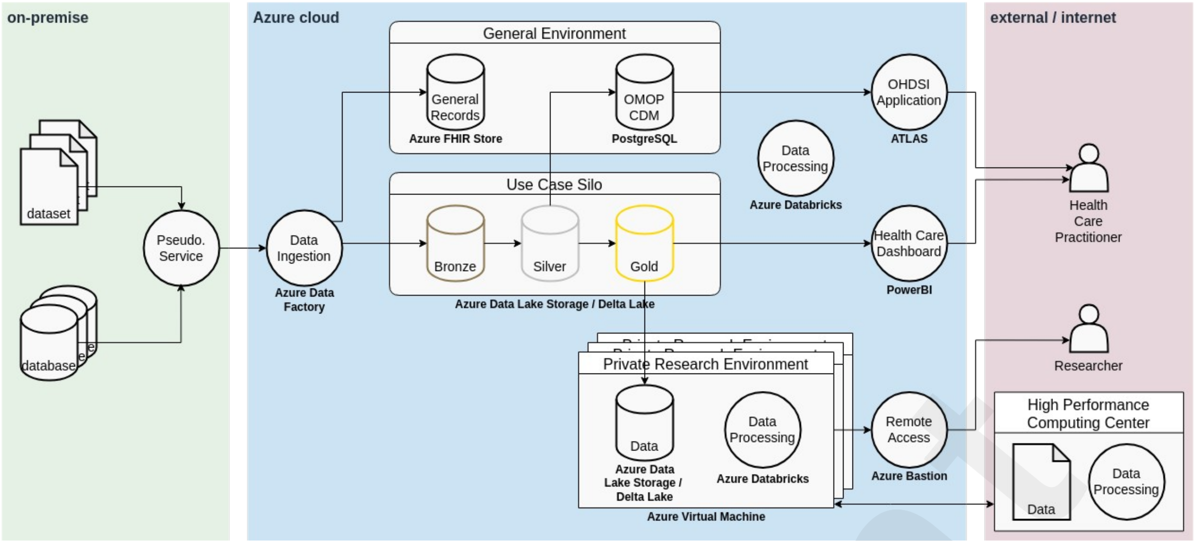


Figure 2: a high-level overview of the medEmotion platform implementation. The figure shows components from three environments: on-premise (e.g. a partner hospital), the Azure cloud and external / internet, which shows users outside the platform. The arrows indicate data flows, starting with five different data sources on the left and flowing to the data users on the right.

The demonstrators were implemented exclusively with public or mock datasets. We created a mock-up data source following the HL7 FHIR standard for the data ingestion demonstrators. The Azure FHIR Store is a natural fit for this data source. Further, an industry partner that offers wearable eHealth devices supplied us with a mock-up data stream for our eHealth device demonstrator. We generally use a decentralised lightweight pseudonymisation service that runs on-premise at a hospital. This service polls for and pulls data in scheduled intervals from the sources, pseudonymises it, and ingests it into the platform using Azure Data Factory.

To showcase the OHDSI OMOP CDM demonstrator, we generated a mock-up dataset based on an MS-specific data acquisition system, which we transformed into the OMOP CDM format. With this data, we operationalised OHDSI tools such as ATLAS to test the transformed data [14].

We used the datasets described by Yperman et al. [18] for the remaining demonstrators. We use the Yperman et al. open-access dataset [19] and an open-access magnetic resonance image dataset from the MICCAI 2016 challenge [20]. We use these datasets to showcase the orchestration of complex pipelines with Azure Databricks, as this enables our existing pre-processing scripts to be automated, orchestrated, and monitored. We encapsulate Azure Databricks in Azure Virtual Machines, where we block all connections to public networks (such as the web) to showcase private research environments. With the same tools but less stringent blocking rules (i.e. we allow connection to selected HPC centres), we enable the use of HPC resources from within the platform. Lastly, we use Microsoft’s PowerBI to create dashboards per healthcare practitioner’s requirements. In particular, we built an interactive, data-rich dashboard that respects therapeutic relationships.

We combined personas and demonstrators to effectively steer the development and implementation of the medEmotion platform, resulting in a feature-rich yet adaptable solution. The use of mock-up datasets has proven particularly useful to accelerate development, as access to patient-level data remains a time-consuming process. In the following, we review the development of the medEmotion platform.

Discussion

One of the key successes of the medEmotion project was its pioneering role in bringing together a diverse group of partners—hospitals, academic institutions, and industry stakeholders—for the first time in a concrete, collaborative effort. This collaboration allowed us to gain invaluable insights into the distinct needs and concerns of various stakeholders, paraphrased as personas as mentioned before, ranging from data protection and privacy issues to the practical demands of healthcare practitioners. Through this partnership, we also learned a great deal about the organisational challenges that remain unaddressed in multi-stakeholder health data initiatives. To give two examples, we have seen that hospital-specific IT strategies vary significantly. While one hospital works within a cloud-based IT infrastructure, another prohibits using cloud-based components in favour of an exclusively on-premise infrastructure. The same can be said for data management strategies. Where one hospital aims to keep data per use case in distinct and tailored information systems, another aims to centralise as much information as possible into a single monolithic system. Personas apply these strategies in discussions but do not necessarily make them explicit. We believe that learning these implicit factors is vital to successful initiatives. Discussions about implementing a health data integration platform gave in-depth insights into hospital-specific data flows. We outlined how patient information flows into the hospital information system, an important exercise that improved mutual understanding. Importantly, beyond the technical knowledge gained, this project fostered personal connections and mutual trust among individuals representing different areas of expertise and roles within their organisations. This familiarity and trust built over time are essential for future interdisciplinary projects.

Clarifying technical real-world use cases using our demonstrators was pivotal in making abstract concepts more tangible for stakeholders. Much like trying to build a car in a world where cars don’t yet exist, stakeholders initially struggled to articulate their needs without a clear vision of what the final product could look like. In such a setting, any implementation process will fail, as the result cannot fit the stakeholders' unarticulated vision. The demonstrators provided a concrete basis for discussions, allowing individuals without extensive technical backgrounds to actively engage and provide meaningful feedback. This practical approach significantly improved communication between the partners and helped bridge the gap between technical experts and end-users. As such, we were able to develop and implement an architecture consisting of individual modules that both satisfied our demonstrators and generated general approval from the stakeholders.

Despite the many successes of the medEmotion project, several frustrations and challenges emerged throughout its course. Looking back, there were decisions and approaches that, with the benefit of hindsight, could have been handled differently to overcome obstacles more effectively. In this section, these key frustrations are summarised.

- *Effective communication and collaboration in a fluid and diverse team proved extremely challenging.* With members frequently joining and leaving, maintaining continuity of knowledge was difficult, leading to disruptions in collaboration. The large and diverse group of individuals, each bringing distinct backgrounds, interests, and expertise, made aligning perspectives and ensuring consistent communication even more complex. This constant flux created coordination barriers, making it hard to sustain productive collaboration and slowing down decision-making processes.

- *Conflicting priorities amongst key stakeholders created significant challenges in aligning goals and decision-making.* Some stakeholders were primarily focused on return on investment and long-term strategic gains, while others were more concerned with its innovative potential and practical usability. This disconnect led to frequent misalignments, where decisions made often overlooked the technical needs or the full potential of the platform, or led to less effective outcomes as critical insights were missing in the decision-making process. As a result, in some cases, the project appeared to be driven more by academic interests rather than the practical needs of the healthcare partners. While the platform seemed promising on paper, there was a disconnect between the project’s goals and the real-world priorities of the hospitals. For some hospitals, the platform was unsolicited and not aligned with their internal strategic plans, leading to doubts about its relevance.

Additionally, one of the partner hospitals already had a similar platform in place, reducing the added value of the new system.

- *The mismatch between the data demands of researchers and the hospitals' reluctance to share sensitive data for secondary use complicated the project.* From the hospital's perspective, the process of preparing data—cleaning, anonymizing, and transforming it—represents a costly and time-consuming task. There is limited immediate benefit while they expose themselves to significant risks concerning the General Data Protection Regulation. Real-world data is highly sensitive and requires highly secured IT environments, which common academic research projects fail to consider. These conflicting needs slowed progress and exposed gaps in the project's ability to balance the expectations of different stakeholders, especially regarding the complexities of data sharing.

- *The project faced significant hurdles related to funding and intellectual property (IP).* The traditional 'project-to-project' funding model limited the potential for long-term commitment. Without a clear financial roadmap beyond the initial phase, maintaining stakeholder engagement became increasingly difficult. At the same time, IP issues arose as partners contributed in different ways—intellectually, financially, or through data use cases—making it complicated to define ownership. The absence of early, transparent IP agreements led to terms of usage that were unacceptable for some partners, limiting the potential for certain project components to be utilised effectively post-project. This further hindered the scalability and long-term success of the initiative.

- *Balancing the openness of collaboration with the security and operational needs of healthcare organizations remains a critical challenge.* Investing in open science can offer significant advantages in collaborative projects. Open-source software fosters transparency and inclusivity by allowing all partners to access, modify, and contribute to the project. This transparency builds trust among stakeholders as everyone can see how their contributions are utilized and built upon. Open science also reduces barriers to entry, enabling partners who may lack financial resources to still contribute valuable intellectual input, levelling the playing field and encouraging broader collaboration. Open-source environments can also drive faster innovation, as contributors freely share ideas, experiment, and improve each other's work. This collaborative approach often accelerates development and produces more robust, well-tested solutions. Furthermore, open-source projects benefit from collective maintenance and support by the community, making them more sustainable in the long term, even after the original funding or key contributors move on. However, there are pitfalls to consider. Open-source solutions, while valuable in academic settings, often conflict with hospital policies that prioritize stringent security measures and professional support. Hospitals may require more controlled environments to meet compliance and data privacy regulations, which may not always align with open-source principles.

- *Rapid scale-up and professionalisation impacted the balance between innovation and implementation.* A small group of researchers, who had initially focused on exploring and developing innovative solutions, had to quickly shift from this exploratory work to handling operational tasks as the project grew. This sudden change required them to manage the day-to-day implementation and ensure compliance with legal and technical requirements, leaving less time for creative research and experimentation. As a result, the researchers experienced a drop in motivation, finding themselves in presentations that felt more like sales pitches than discussions of research results. With no room for experimental approaches, the rapid scale-up and professionalisation of the project brought innovation to a halt—a frustrating experience for those with a research mindset, who thrive on flexibility and exploration.

We believe we are neither the first nor the last researchers to experience these or similar frustrations. Therefore, we must share our experiences as takeaway messages in the following section.

Takeaway messages

Based on our experience, we aim to offer practical advice that can help guide future multi-stakeholder collaborative initiatives in healthcare and data integration.

- *Invest in community building, stakeholder engagement and education.* Creating opportunities for stakeholders to get to know each other—both at the organisational level and on a personal level—is crucial for building trust and understanding. From an organisational perspective, this helps clarify the distinct needs of each stakeholder group and ensures that strategic goals are aligned towards a shared objective where value can be realised for everyone, even if the "return on investment" may look different for each party. On a more personal level, trust is not just an institutional concept; it is a feeling that grows between individuals and thus takes time. Providing opportunities for stakeholders to form personal connections through trustworthy relationships can significantly enhance collaboration. Organising networking events, workshops, or joint educational sessions can help foster these connections and build a strong foundation of mutual trust and respect, which will pay dividends throughout the project.

- *Collaboration among healthcare providers is ideally driven by the healthcare providers themselves.* The providers are in the best position to proactively share expertise and requests for assistance based on their unique needs. Academic partners can play a key role in facilitating partnerships as neutral mediators, helping to bridge gaps between institutions. While commercial partners bring valuable expertise, maintaining a focus on shared goals and minimising potential conflicts of interest can be achieved by ensuring that collaboration remains centred on the needs and priorities of healthcare providers. However, when it comes to implementation, there are additional challenges that need to be addressed, particularly within healthcare settings. Healthcare providers must establish secure data transfer protocols and ensure that patients are informed about how their data will be shared.

- *Ensure all stakeholders are involved in the 'Ideation Phase'.* It's essential to spend adequate time understanding the problem and the real needs and expectations of all stakeholders and future users. One effective approach is to employ qualitative research methodologies, such as semi-structured interviews and focus groups, before moving into the IT development phase. For example, methodologies from the newly launched educational program "System and Process Innovation in Healthcare" at Hasselt University could be leveraged to guide this process. In addition, involving legal departments early—particularly those within healthcare organizations—is vital for ensuring compliance and managing risks effectively. It's also critical to align on IP ownership and a long-term funding strategy at the outset. Establishing clear IP agreements and a roadmap for continued financial support ensures the sustainability of the platform beyond the initial project phase. Finally, it's equally important to sanity-check technical plans with experts to guarantee feasibility, especially given the usual constraints of a fixed budget. Early input from all relevant parties helps prevent misalignments and avoids costly revisions later in the project.

- *Balance innovation with implementation.* To ensure long-term success, it's important to strike a balance between pilot projects that foster innovation and the implementation of existing, proven solutions. These are fundamentally different activities, requiring distinct profiles and skill sets. Innovation thrives when motivated partners are eager to try new methods and solutions. This kind of exploratory work benefits from a small, focused group of individuals with an innovation and/or research mindset who can freely experiment and iterate based on the real-world needs of partners. In this context, both parties gain from the trial-and-error process that allows them to explore potential breakthroughs in a mutually beneficial environment. Implementing novel methods and solutions requires a different approach—one focused on operational efficiency and compliance, rather than experimentation. It’s also crucial to recognize that different people excel at different stages of the process. Some individuals thrive in the innovation phase, where flexibility and trial-and-error approaches are key. Others are better suited to perform structured, detailed work implementing established solutions. Attempting to blend these mindsets can lead to frustrations as the focus shifts from creative exploration to a more rigid, professionalised implementation. Understanding these distinctions and assigning the right people to the right phases can help maintain momentum and preserve the power of innovation while ensuring successful implementation.

Conclusions

Multi-stakeholder initiatives like the medEmotion project highlight the complexity of managing diverse roles, expectations, and goals. Stakeholders are rarely confined to a single role, and roles themselves are often shared among multiple individuals, which can complicate communication and decision-making. To navigate this, it is essential to communicate through full end-to-end use cases, ensuring that all parties understand how the platform functions within the broader ecosystem. The project highlighted the importance of building upon existing technical advancements and demonstrators rather than reinventing the wheel. Despite the availability of proven solutions, there was often a reluctance to fully embrace these innovations. This hesitation, driven by a desire to retain ownership of specific platforms or initiatives, leads to fragmented efforts and ultimately diminishes the overall impact. Collaboration, when done openly and transparently, benefits everyone involved—teamwork is crucial to achieving shared success. While technical sandboxes are useful for facilitating communication and experimentation, translating these ideas into real-world applications can be a slow and frustrating experience. The journey from concept to implementation requires patience and ongoing commitment. Ultimately, the lessons learned from medEmotion—building on existing successes, fostering collaboration, and maintaining a clear vision—lay the groundwork for more effective and sustainable initiatives in the future.

Acknowledgements

CRediT author statement: *Marcel Parciak:* Conceptualization, Methodology, Investigation, Writing - Original Draft, Writing - Review & Editing. *Noëlla Pierlet:* Conceptualization, Methodology, Investigation, Writing - Original Draft, Writing - Review & Editing. *Liesbet M Peeters:* Conceptualization, Methodology, Writing - Review & Editing, Supervision, Funding acquisition.
We thank our three partner hospitals Jessa Ziekenhuis, Noorderhart and Ziekenhuis Oost-Limburg for their contributions towards the medEmotion project. Further, we thank the Limburg Clinical Research Center for sharing their expertise in clinical research projects. The software development of the medEmotion platform was funded by LRM, with the support of the European Regional Development Fund (EFRO-1308). This research received funding from the Flemish Government under the “Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen” programme.

Conflicts of Interest

None declared.

Abbreviations

CDM: common data model
EHDEN: european health data evidence network
FHIR: fast healthcare interoperability resources
HIS: hospital information system
HL7: health level seven
HPC: high-performance computing
IP: intellectual property
IT: information technology
MS: multiple sclerosis
OHDSI: observational health data sciences and informatics
OMOP: observational medical outcomes partnership

References

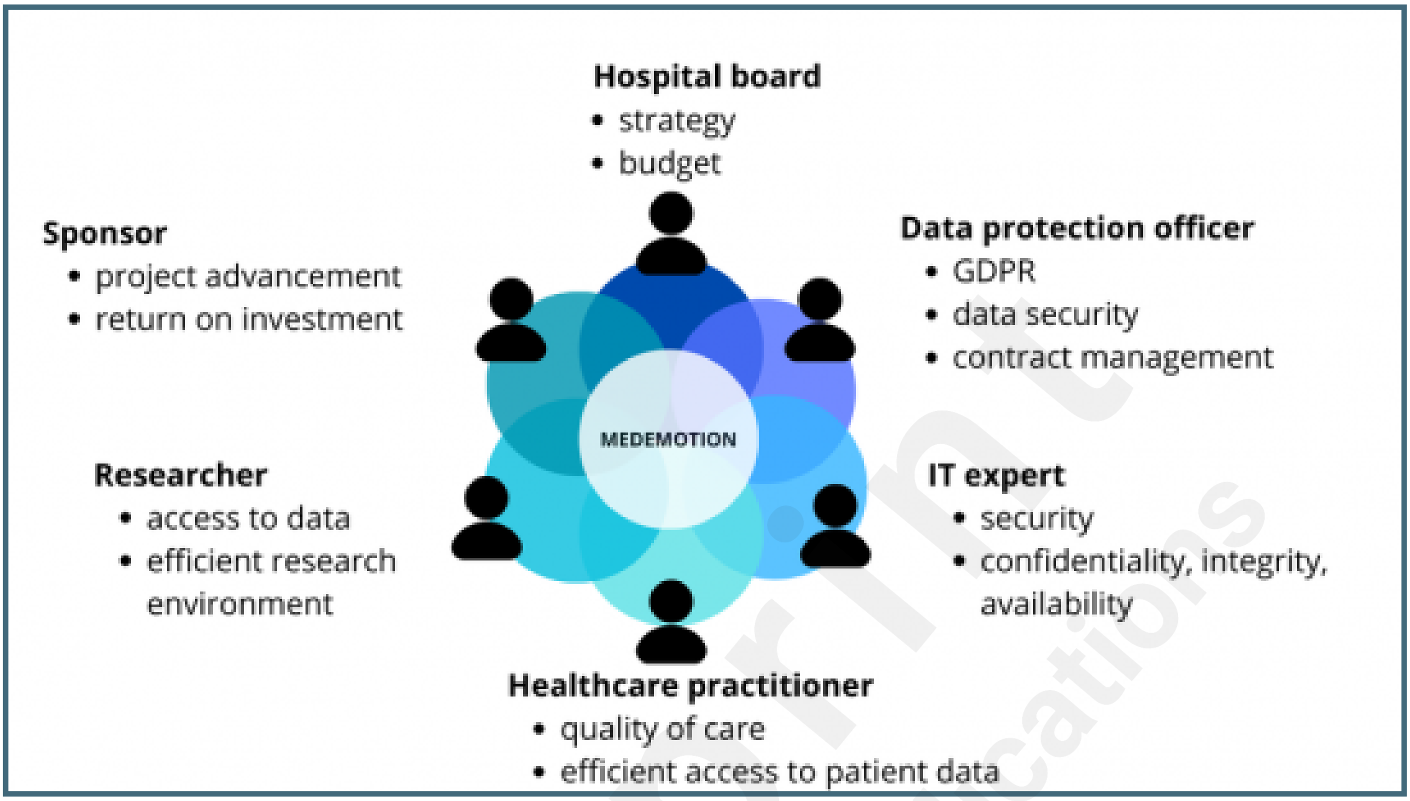
1. Martin-Sanchez FJ, Aguiar-Pulido V, Lopez-Campos GH, Peek N, Sacchi L. Secondary Use and Analysis of Big Data Collected for Patient Care. Yearb Med Inform Georg Thieme Verlag KG; 2017 Aug;26(01):28–37. doi: 10.15265/IY-2017-008
2. Peng C, Goswami P, Bai G. A literature review of current technologies on health data integration for patient-centered health management. Health Informatics J 2020 Sep;26(3):1926–1951. doi: 10.1177/1460458219892387
3. Zhang H, Lyu T, Yin P, Bost S, He X, Guo Y, Prosperi M, Hogan WR, Bian J. A scoping review of semantic integration of health data and information. Int J Med Inf 2022 Sep;165:104834. doi: 10.1016/j.ijmedinf.2022.104834
4. U.S. Food and Drug Administration. Framework for FDA's Real-World Evidence Program. U.S. Food and Drug Administration; 2018. Available from: https://www.fda.gov/media/120060/download?attachment [accessed Jan 9, 2024]
5. Ranchal R, Bastide P, Wang X, Gkoulalas-Divanis A, Mehra M, Bakthavachalam S, Lei H, Mohindra A. Disrupting Healthcare Silos: Addressing Data Volume, Velocity and Variety With a Cloud-Native Healthcare Data Ingestion Service. IEEE J Biomed Health Inform 2020 Nov;24(11):3182–3188. doi: 10.1109/JBHI.2020.3001518
6. Doan A, Halevy A, Ives ZG. Principles of data integration. Waltham, MA: Morgan Kaufmann; 2012. ISBN:978-0-12-416044-6
7. Ortega-Calvo AS, Morcillo-Jimenez R, Fernandez-Basso C, Gutiérrez-Batista K, Vila M-A, Martin-Bautista MJ. AIMDP: An Artificial Intelligence Modern Data Platform. Use case for Spanish national health service data silo. Future Gener Comput Syst 2023 Jun 1;143:248–264. doi: 10.1016/j.future.2023.02.002
8. Wang M, Li S, Zheng T, Li N, Shi Q, Zhuo X, Ding R, Huang Y. Big Data Health Care Platform With Multisource Heterogeneous Data Integration and Massive High-Dimensional Data Governance for Large Hospitals: Design, Development, and Application. JMIR Med Inform 2022 Apr 13;10(4):e36481. doi: 10.2196/36481
9. Poba-Nzaou P, Uwizeyemungu S, Dakouo M, Tchibozo A, Mboup B. Patterns of health information exchange strategies underlying health information technologies capabilities building. Health Syst 2022 Jul 3;11(3):211–231. doi: 10.1080/20476965.2021.1952113
10. Gagalova KK, Leon Elizalde MA, Portales-Casamar E, Görges M. What You Need to Know Before Implementing a Clinical Research Data Warehouse: Comparative Review of Integrated Data Repositories in Health Care Institutions. JMIR Form Res 2020 Aug 27;4(8):e17687. doi: 10.2196/17687
11. Cuggia M, Combes S. The French Health Data Hub and the German Medical Informatics Initiatives: Two National Projects to Promote Data Sharing in Healthcare. Yearb Med Inform

Georg Thieme Verlag KG; 2019 Aug;28(1):195–202. doi: 10.1055/s-0039-1677917
12. Raj A, Bosch J, Olsson HH, Wang TJ. Modelling Data Pipelines. 2020 46th Euromicro Conf Softw Eng Adv Appl SEAA Portoroz, Slovenia: IEEE; 2020. p. 13–20. doi: 10.1109/SEAA51224.2020.00014
13. Geys L, Peeters L. Strategic Oversight Across Real-World Health Data Initiatives in a Complex Health Data Space: A Call for Collective Responsibility: Proc 17th Int Jt Conf Biomed Eng Syst Technol Rome, Italy: SCITEPRESS - Science and Technology Publications; 2024. p. 577–584. doi: 10.5220/0012417700003657
14. Observational Health Data Sciences and Informatics. The book of OHDSI: Observational Health Data Sciences and Informatics. San Bernardino, CA: OHDSI; 2019. Available from: https://ohdsi.github.io/TheBookOfOhdsi/ ISBN:978-1-08-885519-5
15. Park S, Bekemeier B, Flaxman AD. Understanding data use and preference of data visualization for public health professionals: A qualitative study. Public Health Nurs 2021 Jul;38(4):531–541. doi: 10.1111/phn.12863
16. Dinh-Le C, Chuang R, Chokshi S, Mann D. Wearable Health Technology and Electronic Health Record Integration: Scoping Review and Future Directions. JMIR MHealth UHealth 2019 Sep 11;7(9):e12861. doi: 10.2196/12861
17. Carroll N, Richardson I. Software-as-a-Medical Device: demystifying Connected Health regulations. J Syst Inf Technol 2016 May 9;18(2):186–215. doi: 10.1108/JSIT-07-2015-0061
18. Yperman J, Becker T, Valkenborg D, Popescu V, Hellings N, Wijmeersch BV, Peeters LM. Machine learning analysis of motor evoked potential time series to predict disability progression in multiple sclerosis. BMC Neurol 2020 Dec;20(1):105. doi: 10.1186/s12883-020-01672-w
19. Yperman J, Popescu V, Van Wijmeersch B, Becker T, Peeters LM. Motor evoked potentials for multiple sclerosis, a multiyear follow-up dataset. Sci Data 2022 May 16;9(1):207. doi: 10.1038/s41597-022-01335-0
20. Commowick O, Kain M, Casey R, Ameli R, Ferré J-C, Kerbrat A, Tourdias T, Cervenansky F, Camarasu-Pop S, Glatard T, Vukusic S, Edan G, Barillot C, Dojat M, Cotton F. Multiple sclerosis lesions segmentation from multiple experts: The MICCAI 2016 challenge dataset. NeuroImage 2021 Dec;244:118589. doi: 10.1016/j.neuroimage.2021.118589

Supplementary Files

Figures

Personas involved in the different demonstrators, with some of their focus points indicated.



A high-level overview of the medEmotion platform implementation. The figure shows components from three environments: on-premise (e.g. a partner hospital), the Azure cloud and external / internet, which shows users outside the platform. The arrows indicate data flows, starting with five different data sources on the left and flowing to the data users on the right.