TALLINN UNIVERSITY OF TECHNOLOGY
School of Business and Governance
Ragnar Nurkse Department of Innovation and Governance

Hille Hinsberg

Data governance in digital health care: enablers and capabilities

Master's thesis
Programme HAGM09/18

Supervisor: Veiko Lember, PhD
Co-supervisor: Kerli Onno, MA

Tallinn 2022

I hereby declare that I have compiled the thesis independently and all works, important standpoints and data by other authors have been properly referenced and the same paper has not been previously presented for grading.
The document length is 14,045 words from the introduction to the end of conclusion.

Hille Hinsberg /signed digitally/
December 30, 2022
Student code: 191915HAGM
Student e-mail address: hillehinsberg@gmail.com

Supervisor: Veiko Lember, PhD:
The paper conforms to requirements in force

……
(signature, date)

Co-supervisor: Kerli Onno, MA:
The paper conforms to requirements in force
……
(signature, date)

Chairman of the Defence Committee:
Permitted to the defence
……
(name, signature, date)

TABLE OF CONTENTS

Abstract

This study centres on data management capabilities of health technology providers in Europe. The research reflects current practices in their data governance to indicate the organizations´ potential to absorb the impact of proposed data regulations. European Health Data Space initiative and other recent data-related regulations aim to enhance the use and exchange of health-related data on the European scale. Better data governance is instrumental in adopting, upscaling and transferring technological solutions to foster the development of digital and data-intensive health and care services. The study brings new knowledge to define which data management capabilities, i.e. set of skills, routines and resources that lead to enhanced data use and interoperability.

Keywords: data governance, health technology, innovation capacity, data management capabilities, data-related regulations

Introduction

The study is set to analyze how well data-intensive service providers in the health-care sector are prepared to meet the challenges in the legislative context as set forth by the European Commission (EC) in the data governance domain. Data-processing technologies and data-driven services are important facets of service innovation in Europe, including in the domain of health and care. Successful innovations in health and care are expected to bring personal, economic as well as social gains, starting from the human-centric aspect of personal well-being to considerable effects in the society (European Commission 2020c). Digital solutions supporting this vision rely on data to be collected, processed, and accessed for direct use in the provision of care, which is defined as the primary use of data. The regulations aim for health data to be increasingly re-used for scientific research, for scaling up innovation and for policy decisions to improve the public health system referred to as secondary purposes (*Ibid.*).

Innovative organizations that provide technology-based services on the market do not exist in a vacuum. They are part of an ecosystem that is influenced by national as well as supra-national level context, including regulations, policies and funding instruments to boost growth and enable scale-up on the market (Lundvall 2007). While member states in Europe make their own policy decisions on how to steer their health and care systems, technology that supports provision of health services can be used to connect disparate systems. All service providers must comply with regulative norms and common standards on technology use such as privacy protection, data management and information security. Legal acts such as General Data Protection Regulation (GDPR) have a direct effect on service providers in the health and care sector. The market actors must be ready to comply with and adopt the new regulations to their benefit.

The European Commission envisions common European Health Data Space and other acts on data governance to achieve better accessibility and efficiency of services by making data available for re-use. "Better access to health data could significantly facilitate innovation over healthcare ecosystem: in public health and prevention, population health management, health resources planning, decision-making, product and services development and innovation, education and research" (European Commission 2020c). A number of recently adopted and upcoming regulations and policies (incl Data Governance Act, Data Act and others) are expected to strengthen the innovation capacities of companies, as well as research and development organizations. The Acts should "create new opportunities to collect, share and combine data between stakeholders, sectors and member states" (European Commission 2020c). By harmonizing data governance, the goal is improved interoperability within the national ecosystem of health and care, and between them so that more value will be created for citizens, the service users.

Data-intensive health-care services are designed and delivered by a variety of actors. The health and care service providers, such as hospitals and care homes, are generally managing the care pathways in direct contact with end users and act as procurers of technology solutions that facilitate the service provision. Providers of data-intensive technology solutions develop health applications, provide data sharing platforms, conduct data analytics including AI-enabled support in decision-making for health care. Ground principles for managing data are relevant for all these actors. Each organization needs data governance function to have control over data they process, to comply with regulatory frameworks and ensure appropriate data management operations.

This thesis studies capabilities of data-intensive service provider organizations, focusing on the management and exchange of health data. Based on theory, the study shows that data governance capacity affects the overall innovation capacities of business organizations. Data management capabilities are a good indicator of maturity for innovation deployment. and ability to scale up innovative services across populations and borders (European Commission 2018b). The question is, which configuration of capabilities is needed to produce the anticipated outcomes.

As such, the aim of the thesis is to find out how can providers of data-intensive services better adapt to the changes induced by the regulatory initiatives on data governance. What challenges must be overcome to harmonize data use and exchange in Europe? What capabilities must be improved in business organizations to facilitate implementation of regulations?

The main research question as proposed by the author is the following:

- What are the critical capabilities that digital health service providers need to adopt the emerging European regulatory framework on data governance?

In order to approach the main question, the author conducts a case study to explore challenges related to data governance as experienced by health technology providers in Europe. In the empirical part, the maturity of data management capabilities is described based on a sample of providers of data-intensive services. The capabilities are assessed in the light of policies and regulations that strive to harmonize data governance in the digital health and care domain.

The thesis is structured according to the questions that were formulated for the research.
In the first chapter, the theoretical base is reviewed to describe the relationship between innovation capacities and data governance as a concept of exercising authority and control over data. Subsequently, data management functions are described and the author proposes a synthesis of data management capabilities that are needed in business operations for providing data-intensive services. Capabilities assessment framework is used to operationalize the indicators of maturity in data management practices. Chapter 2 outlines the European Commission regulations that have a direct effect on using and sharing health data in the European Union. The author presents an overview of challenges that stakeholders in the health care sector have highlighted in assessing the potential impact of the regulatory proposals. In Chapter 3, the method, process and findings of empirical reseach are presented. Chapter 4 is dedicated for discussing the research findings in the context of data policies that affect market actors. Based on the findings and analysis, the author

presents some recommendations that could facilitate policy implementation and support capabilities of organizations providing data-intensive services in the health care domain. The final chapter concludes the thesis.

1. Innovation capacities: theoretical framework

Based on the stated potential impact of proposed regulations such as European Health Data Space, the Governance Act and the Data Act, the general goal of EC policymakers is to scale up the deployment of health and care innovation across Europe (European Commission 2020c). The impact is expected to appear in a systematic diffusion and improved availability of services for end users. As defined by WHO, innovation means "health service components or practices that are new or perceived as new in a particular context" (World Health Organization 2009). Rather than a single component such as new technology, innovation refers to a set of interventions, including organizational processes to support implementation capacities *(Ibid.)*

To be able to scale up innovative solutions, there is an ongoing need to elaborate the pan-European interoperability of technological innovation (European Commission 2022b). Interoperable in this context does not mean a standardized approach on how to innovate health care provision as there cannot be a one-size-fits-all model due to differences in national health and care systems. The common and connecting aspects come from providing a framework of innovation enablers and drivers. These should be applicable independent of contextual specifics. In addition to driving data sharing and interoperable exchange, such enablers for innovation diffusion may include recognition of professions and professional conduct, standardizing conditions for reimbursement of health applications and other technology solutions; defining criteria for health technology assessment or enhancing general digital literacy to support take-up of innovative services (World Health Organization 2009).

In the current study, enablers and drivers for data-driven innovation will focus on legal and policy aspects associated with data governance and data management capabilities. The challenges of complex regulation, along with funding and reimbursement issues are found by groups of experts and stakeholders as the main barriers hindering successful scale-up of digital health innovations (Schlieter *et al.* 2022). It should be noted that the alignment with regulatory provisions will not be explored in view of merely ensuring compliance – that is, adhering to legal norms. Instead, the focus is on policies as enablers for innovation.

1.1. Innovation management

Innovation management is viewed as a form of organizational capability. Organizations should invest into and nurture innovation capabilities to develop new products, processes, and services (Sultana *et al.* 2022; Lawson, Samson 2001). Innovation capabilities are regarded as a strategic asset (Sultana *et al.* 2022; Kalmuk, Acar 2015) and are often considered critical resources for success in a highly agile environment (Lyon, Ferrier 2002). The strategic performance of business is positively associated with developing internal capabilities – such as capabilities related to data management (Sultana *et al.* 2022; Kopanakis *et al*. 2016).

Many scholars have conducted research to describe information technology (IT) related capabilities as a subtype of innovation capabilities. For example, Bharadwaj (2000) defines IT capabilities as a multidimensional construct, the aim of which is "to mobilize and deploy IT-based resources in combination or co-present with other resources and capabilities". This approach depicts IT capability to explicitly include business resources (such as people, processes, technology) and emphasizes the connection of IT-related and business resources and capabilities. Scholars further argue that different types of capabilities can have different effect on organizational performance (Shuradze, Wagner 2016). Several studies demonstrate significant positive relationship between IT capability and organizational ability to deal with changes that arise unexpectedly in business environment – such as changes in the regulatory context (Shuradze, Wagner 2016; Lu, Ramamurthy 2011).

Data-driven innovations (DDI) are defined as "innovation processes that apply techniques (such as big data analytics) and technologies (such as machine learning, deep learning, AI) to extract meaningful value from data to generate innovative results" (Sultana *et al.* 2022). Organizational capabilities and technology-related capabilities should be combined to govern innovation (Cepeda, Arias-Perez 2018). Sultana and other authors show that by applying DDI in operations, research and development, new product and service development, marketing and management, companies can achieve a sustained competitive advantage (Sultana *et al.* 2022). Zolnowski *et al.* (2016) argued that data-driven innovations enable optimization of internal and external processes and thus productivity can be achieved. The managerial, technological and personnel capabilities that support DDI can be referred to as data-driven innovation capabilities (DDIC) (Sultana *et al.* 2022).

The capability to tackle the fast and often unprecedented changes within the DDI environment is known as agility (Sultana *et al.* 2022; Teece *et al*. 2016). It represents the capabilities of responding competently to market opportunities or uncertainties in the external context (Sultana *et al.* 2022; Li *et al*. 2021). Therefore, developing agile organizational capabilities allows companies to adapt to changes and build higher value within the business ecosystem (Sultana *et al.* 2022; Akhtar *et al*. 2018; Chan *et al.* 2019).

Innovation based on the adoption and usage of digital technology alone will not create an effective data governance function and vice versa – the ability to use data is not an innovation in itself. Shuradze and Wagner (2016) propose that many organizations have systems in place to collect data, however, they lack an approach to put data into use for strategic purposes. In order to achieve competitive edge by data-driven innovation, the organization must govern data to achieve its strategic goals (Sultana *et al.* 2022). It is recommended that organizations adopt a capability framework that incorporates people, processes and technology. These three elements are combined for successful data governance: to include expertise and capacity to perform data management functions; make proficient use of the technological tools and conduct data management processes (Shuradze, Wagner 2016). In the empirical part of this study, the author operationalizes this framework to conduct the empirical research.

1.2. Data governance

How to govern data as a valuable asset has become critical for business organizations (Alhassan *et al.* 2016). Legacy technologies and outdated business functions may prevent gaining value from data that the organization has acquired by considerable efforts and costs. There are data protection challenges and security risks to address. On top of internal challenges, there are factors in the external context such as regulatory requirements that the organizational data management function must comply with (Enterprise Data Management Council 2014).

Data governance is “the exercise of authority and control over the management of data” (DAMA International 2009). Addressing external factors in the organizational policies is closely connected to data governance in an organization. Data governance comprises setting standards, defining rules, establishing policy and implementing oversight to ensure adherence to best practices. This would be achieved by activities related to strategy, operations, data architecture, IT deployment,

and maintaining data quality, with a goal to support business goals. (Enterprise Data Management Council 2014)

The purpose of data governance is "to increase the value of data while minimizing data-related costs (e.g. for management, sharing, analysis) and alleviating risks such as privacy breach" (TEHDAS 2021a). The main difference between the terms 'governance' and 'management' is that governance refers to what decisions must be made and who makes them in order to ensure effective management and use of resources, whereas management involves implementing decisions (Fu *et al.* 2011; Khatri, Brown 2010). Hence, management is influenced by governance (Otto 2011). Governance also ensures that the principles of data management are backed by relevant capabilities.

1.2.1. Data management capabilities

Data management capabilities are defined as "sets of skills, routines, and resources a company needs to have in order to support business capabilities through data management" (Competence Center Corporate Data Quality CC CDQ 2022). Typical data management capabilities refer to data capture (i.e data collection from a range of sources), data standardization and harmonization, data processing, and data access policies including protection *(Ibid.)*.

Given that data governance is a dynamic area, organizations need to be agile to respond to constant change. Many change factors come from legislative context and need to be adopted in the business model and service processes that rely on data management. It is generally upheld that in the health and care service provision, "existing and emerging regulation and policies are likely to have a huge impact on data governance, affecting patients' data privacy rights, professionals delivering care and healthcare organizations" (KMPG 2018). Organizations are also bound to address the problem of inaccurate and incomplete data and set up processes to constantly monitor and ensure data quality (Abraham 2019; Kim & Cho 2018).

Categorizations of data management functions are proposed by several sources. The models described below are chosen as relevant for data-driven service providers.
The reference model by CC CDQ specifies data management in three categories: goals, enablers, and results, which are interlinked in a continuous improvement cycle. (CC CDQ 2022; see also Figure 1)

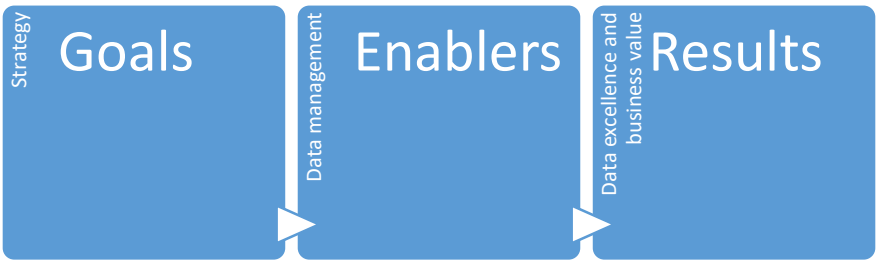

Figure 1. Structure of the Data Management Framework
Source: CC CDQ (2022)

DATA STRATEGY (goals)

"A data strategy is a broader strategic framework including corporate, digital, functional, divisional and IT strategies, all linked to the general business strategy" - such as the strategy for data-driven innovation (Davenport *et al.* 2001). It follows by logic that the clearer and more targeted is the business strategy, the easier it is to define what data and what kind of data management capabilities are required to create business value *(Ibid.)*.

ENABLERS

Enablers are necessary for providing the required data management capabilities, including several sub-categories (CC CDQ 2022):

- People, roles and responsibilities to ensure effective data management and consistent use of data across the organization;
- Performance management, by setting the appropriate measures to monitor and control the performance (i.e., progress and outcomes) of data management;
- Processes and methods, such as adhering to standards for managing and using data;
- Data architecture specifies data storage and describes how data flows between its applications, or organizational functions;
- Data lifecycle defines data sources and operations ranging from data acquisition and creation to data archiving), connecting data consumers and data use contexts;
- Data infrastructure designates software components and tools supporting data management activities.

The sub-categories are inter-related and the boundaries cannot be clearly defined in practice. For example, technology, including the hardware and software used within an organization supports data management operations which are conducted by certain processes and methods. In parallel, technology is also connected to the infrastructure capabilities needed to share data and provide end-user access (Davenport *et al.* 2001).

RESULTS

In the CC CDG model, the results, or outcomes of data management can be viewed from two aspects: first, data management has a direct impact on the data itself (such as its quality and compliance with security standards or privacy regulation), defined as "data excellence" in the reference model (CC CDQ 2022). Secondly, data management is an instrument of creating "business value", which relates to financial gains, improved business processes, more customers, and organizational growth *(Ibid.).* The two aspects are interrelated. For example, organizations with poor data governance spend time and human efforts reacting to data-related issues, therefore they cannot dedicate attention to improvements in other business processes (Abraham 2019; Barker 2016).

1.2.2. Data management capability assessment models

Data Management Capability Model (DCAM) defines the scope of capabilities required to establish, enable and sustain data management practices in an organization (Enterprise Data Management Council 2014). The model is organized into core capabilities that are needed for sustainable data management, including strategy, data management business case and funding, the operating model, data architecture which focuses on the core concepts of how data is defined (*Ibid.).* Technology architecture describes the relationship of data with the IT infrastructure needed for operational deployment. Data quality category describes the processes of control over data supply chain. Data operations are meant as data lifecycle process and how data content is linked to conducting organizations' functions. Each capability is complemented with a set of measurement criteria to be used in evaluating data management practices *(Ibid.).*

For the purpose of conducting analysis for this paper, the author proposes a categorization of capabilities that is drawing from and synthesizing the Data management capabilities assessment model (DCAM) provided by Enterprise Data Management Council (2014), Data Management Framework by CC CDQ (2022) and data governance functions proposed by KPMG (2018) (see Table 1). The categorization aims to cover all components that have been considered by researchers as most relevant for the data management functions.

The proposed structure will further be used to formulate indicators that mark the readiness of an organization to respond to challenges brought about by external factors such as policies and standards, including regulatory requirements proposed by the European Commission and described in Chapter 2. The data management capabilities assessment model will be used to guide

the empirical research to develop hypotheses for the case study and answer research questions, as presented in the Introduction.

Table 1. Capability categories and indicators

Data management capability area	Indicators
Data governance Governance defines the organizational structure by which the data will be managed, and ensures that the principles of data management are implemented throughout the organization and the service chain	○ How are external and internal policies, procedures and standards incorporated (and compliance ensured) in the operations? ○ How is the alignment coordinated with stakeholders, e.g. with vendors and service users?
People, roles and responsibilities How are skills and organizational structure supporting data management activities by determining roles and responsibilities as well as reporting and collaboration modes for operational functions.	○ Are the tasks defined and skills sufficient to ensure effective data management and consistent use of data across the entire organization?
Performance management A methodology for measuring data management progress, along with a set of metrics and consistent monitoring to identify the outcomes of data management activities.	○ Are the measures defined to monitor and control the performance (i.e., progress and outcome) of data management? ○ Are the key performance indicators defined and distributed across the organization?
Processes and methods Work flows and protocols, including applying relevant standards for using and sharing data.	○ Are relevant procedures and standards defined for managing and using data properly, securely and consistently?
Data architecture Data architecture identifies data domains, defines critical data elements, establishes taxonomies. All these elements ensure that the usage of data is consistent and the data content is relevant for the function it serves.	○ Is the conceptual model and policy defined to guide data flows between applications and business functions?

Data operations (data lifecycle management) Lifecycle management refers to the operations ranging from data acquisition and creation to data archiving, with an aim to improve the data quality and availability, including for re-use.	○ How are data objects and documents defined, along with reviews of data sources, operational data activities (i.e ranging from data acquisition and creation to data archiving), data consumers, and data use contexts?
Data infrastructure management The category includes choosing and operating software components, platforms and tools to support data management activities in their complexity, including enabling data sharing with users outside organizational boundaries.	○ Are appropriate software components and tools used to support data management activities?

Source: author, synthesis based on Data management capabilities assessment model (DCAM) by Enterprise Data Management Council (2014), KPMG (2018) and Data Management Framework by CC CDQ (2022)

2. Regulation on data governance in the European Union

In 2020, the European Commission presented a reform package known as European strategy for data, a set of policy measures and investments to support the data economy (European Commission 2020a). The Commission foresees a legislative framework for the governance of common European data spaces, including *inter alia* European Health Data Space regulation (EHDS), the Data Governance Act (DGA) and Data Act (DA).

The European Health Data Space initiative is specific to digital health and care sector, comprising rules, standards and practices, commonly deployed infrastructures for data processing and a governance framework for digital health. The framework is foreseen to address issues that are specific to the exchange of and access to health data and the use of digital services, including artificial intelligence in the health and care ecosystem (European Commission 2022b). Along with non-sector specific Data Governance Act and Data Act, the initiative is expected to have a strong impact on all actors whose business model is dependent on generating and exchanging health-related data and personal data (European Commission 2020c).

The EHDS regulation seeks to standardise patient health files and ensure that electronic health data is interoperable (European Commission 2020c). Requirements on security and interoperability requirements would be introduced for Electronic Health Record (EHR) systems, the software used for storage and sharing of health records. A distinction is made between using electronic health data for medical reasons which is defined as primary use, and the re-use of health data for purposes such as research, innovation or evidence-based policymaking referred as secondary use (*Ibid.)*.

The aim of proposed Data Act is to "foster business-to-government data sharing for the public interest, support business-to-business data sharing, and review the rules on the legal protection of databases, to further enhance data access and use" (European Commission 2020d). The Act also foresees improving technical standards for portability of data generated by individuals.

The EU Data Governance Act aims to support the availability of data for use by increasing trust in data intermediaries and by strengthening data-sharing mechanisms across the EU. Personal data re-use would be enabled also on altruistic grounds. (European Commission 2020e) For all data holders, this requires setting up specific processes to collect personal data, ask for

informed consent by users, grant access for re-use, establish necessary safeguards to protect the person´s privacy. The aim is to enable persons to control who accesses their data and for which uses, to restrict access, and to transmit information from one health care provider to another. While GDPR focusses on "enabling individual control for data subjects of 'their' data" (van Veen 2018), the new conditions supplement the GDPR with additional rights and mechanisms relating to personal health data.

The acts aim to increase re-use of data and exchange across organizational boundaries, including between stakeholders in a regional or national ecosystem, and across borders. The underlying ambition of the regulatory package is increased interoperability of technology platforms and data-driven services in the European market. However, most prerequisites for achieving this ambition are rooted in the data management operations that a company needs to put in place internally.

For example, companies routinely share their data with stakeholders such as vendors, other service providers or with public agencies for reporting and reimbursement purposes, etc. However, external policies and standards need to be followed to decide what data can (and cannot) be shared, what approvals permit further use of data and how data delivered to third parties will be protected, in alignment with information security standards (TEHDAS 2021a). To support these routines, the capabilities related to people, processes, or technology form the base for enabling further re-use for purposes such as data portability or ensuring data exchange across platforms (*Ibid.).*

2.1. Stakeholder challenges in data governance

Analyzing the digital health and care ecosystem in Europe, stakeholders can be described according to their functions and roles in view of data governance (Kütt *et al* 2022):

a) Users of health and care services fall into different profiles – they may be patients or health care professionals, formal or informal care givers. In the aspect of data governance, they can be in the role of data subject, the person who owns data upon themselves or data user, a person who has lawful access to certain personal or non-personal data and is authorised to use that data, according to Data Governance Act (European Commission 2020e).
b) Service providers i.e., actors who are directly involved in the service delivery for end users. They can be in either data controller or data processor roles according to GDPR definitions

acting as custodians of the data assets, ensuring the quality, accuracy and security of the data collected in their service processes (European Commission 2016).

c) Data facilitators (data intermediaries) who provide data mediation services for re-using data sets to create new value within the ecosystem. The role of data intermediaries was defined by the Data Governance Act – these are services that deploy appropriate methods to enable data access in databases, or establishment of specific infrastructure for data processing (European Commission 2020e). Examples of data intermediaries are FinData in Finland and Health Data Hub in France which are government-backed agencies with a mandate to promote and support secondary use of health and social care data.
d) Regional and national policymakers enforce norms and standards for data use and exchange within a jurisdiction, including defining access rights to data held by public registries.
e) Regulators on the EU level set forth regulatory norms and technical requirements, including proposing infrastructure frameworks to enable exchange of data and services across borders within the EU.

In light of the proposed regulations on data governance, stakeholders in the ecosystem experience diverse challenges. In a trouble-shooting analysis conducted by TEHDAS (2022b), European data users highlighted a wide range of barriers to data sharing, mostly related to legal and data management issues caused by misalignment of interpretations and implementation due to lack of semantic interoperability and differing interpretations of key terms. Regarding the real-life examples of the barriers identified by case studies, more than half were legal-related barriers, 30% were caused by data management, 13% were technical issues and 5% were trust-and-transparency-related barriers (TEHDAS 2022b). Fragmentation of the health data legal landscape is found to create significant challenges to data use and reuse, with new problems arising when scaling up innovative services across borders (EIT Health 2021).

The following overview of challenges is based on consultation documents, impact assessment reports and policy papers that reflect European stakeholders´ views and expectations towards the proposed regulatory initiatives.

i. Enablers for secure access to data

Insufficient availability of data and barriers to health data exchange have negative impact on the provision of healthcare services, referring to primary use of health data (Digital Europe 2021). The

level and mindset for digitalization varies in member states which creates a barrier for interoperability between healthcare providers. Even within the national ecosystem, stakeholders have different approaches for access to and sharing of health data. Public authorities have diverging interpretations of what can be shared as anonymized data. This sets a barrier for data sharing as data holders deploy overly risk-averse procedures. Potential risk for non-compliance with the GDPR leads to a slow adoption of new technologies and low motivation to share data business-to business (European Commission 2018b). Another challenge lies in disparate sources of health-related data needed for secondary use. Most remarkably, socio-economic data could be valuable for providing personalized services but it cannot be collected from patients' medical records that only capture clinical data (TEHDAS 2021a).

Based on the framework presented in Chapter 1, the challenges are connected to the following capabilities (see Table 1):

- technical maturity, including infrastructure and tools to provide access to data in controlled and secure way, to enable business-to-business and business-to government exchange of health data;
- capacity to establish clear safeguards and deidentification methods such as anonymization to protect the data from misuse and ensure privacy compliance in the full chain of service provision;
- data holders´ ability to ensure quality and accuracy of data, including of metadata (description of data that is collected and processed for the primary purpose) to enable findability and re-use.

ii. Users´ control over their own data

Exercising access and control over their own health data is often difficult for persons. Electronic health records (EHRs) are not yet available in many countries, meaning patients cannot easily access and use the information digitally. Where data are held across several data controllers, data portability is not enabled, meaning the patient cannot exercise their right to move data between service providers (TEHDAS 2022b). The EHDS foresees mandating the development of specific tools to facilitate citizens' access to and portability of their own health data, including citizen-generated data (European Commission 2022b). For service providers, application of relevant safeguards and maintaining processes for user consent management means significant costs.

Based on data management capabilities framework (see Table 1), the service providers would need to demonstrate readiness for:

- implementing and/or improving technical standards (such as standard APIs and interoperable data formats) to enable access to and portability of data;
- setting up appropriate business processes to meet the requests by patients (service users) to access, control and move their data between service providers. The right for personal control includes the purpose of data altruism for the benefit of health care provision as defined by Data Governance Act (European Commission 2020e).

iii. Standard-based interoperability

The regulations reinforce an understanding that data sharing including business-to-business and business-to-government enables to create new value by serving different purposes than the primary function. However, limited technical interoperability and inconsistent use of technical standards sets limits on scaling up digital health services. It is a challenge for businesses when integrating digital health care solutions in healthcare systems on a national level, and even more in cases of cross-border provision of services. The EHDS aims to "strengthen semantic and technical interoperability by developing infrastructures, and related services to facilitate cross-border storage, processing and analysis of health data" (European Commission 2020c).

Based on data management capabilities framework (see Table 1), service providers would need to demonstrate readiness for the following aspects:

- capability to connect to interoperable data access infrastructures;
- voluntary adherence to relevant standards and guidelines on technical and semantic interoperability;
- participation in communities and networks that strive to harmonize interoperability standards with a view to enable cross-border data exchange.

2.2. Capacity to adapt to the regulatory changes

In order to conduct the empirical study and analyze the potential to absorb the effects of the regulatory aspects, the author proposes a framework for combining the data management capability areas defined in Chapter 1 with the challenges addressed by the regulations as summarized in Chapter 2.1 (see Table 2). Not all data management capabilities can be directly

aligned with the enabling provisions in the upcoming EC regulations. Based on the referenced capability assessment models, categories are included in the framework for comprehensive coverage. The framework is not meant to be normative and therefore does not include scores, scales or maturity levels (e.g. from basic to advanced). Rather, it can serve as a tool to describe the relevance of data management capabilities for technology providers who would be expected to not only comply with, but apply the regulatory provisions to their benefit. The questions which allow to assess current capabilities were developed by the author to conduct empirical research. They were created based upon synthesis of challenges as presented in Chapter 2.1.

Table 2. Data management capabilities to address the impact of forthcoming regulations. Source: Author´s own creation

Data management capability area	Challenges addressed by regulations	Indicators of relevant capabilities	Questions to assess current capabilities
1.Data governance How are policies, procedures and standards incorporated (and compliance ensured) in the operations? How is the alignment coordinated with stakeholders, e.g. with vendors and partners?	Member states have different regulative approaches for access to and sharing of health data, even within the national ecosystem	Capability to establish clear safeguards and implement legal compliance in the service chain, coordinating with partners and vendors	Q1.What procedures have you established to ensure compliance with EU legislation? (e.g. in areas of privacy and data protection policy, data exchange policy, IoT generated data management or others) Q2. Do you coordinate compliance throughout the service chain, with vendors and partners?
2.People, roles and responsibilities Are the tasks defined and skills sufficient to ensure effective data management and exchange?	Digitalization and data-driven economy models in the healthcare sector will require new skills and positions to support service provision	Capabilities to conduct effective data management in operations and ensure consistent use of data across the entire organisation	Q3.Have tasks been assigned to perform the tasks relevant to data exchange functions (B2B and B2G)? Q4.Is the staff equipped with sufficient skills to perform these tasks?
3.Performance management Are the measures defined to monitor	Not regulated in the EC regulations in scope	Capability to monitor performance and achieve outcomes of data management	Q5.How you defined metrics to measure performance related

and control the performance (i.e., progress and outcome) of data management?			to enabling data exchange and re-use? Q6.Are the key performance indicators distributed across the organization?
4.Processes and methods Are the procedures and standards defined for collecting, managing and using data properly, securely and consistently?	Inconistent and fragmented use of interoperability formats for data re-use and to enable personal data portability	Ability to capture data from public databases Capability to provide non-personal as well as personal data for re-use by other actors by applying appropriate safeguards Processes and tools to ask informed consent of users to process their data for secondary use for the benefit of health care provision	Q7.Have you defined processes to capture and re-use data from public sources? Q8.Are you able to apply safeguards or privacy-preserving technologies on personal data? Q9.Do you enable data portability for users (data subjects)?
5.Data architecture The conceptual model covering what data is stored in which application, and how data flows between applications	Lack of interoperability of electronic health records, as well as semantic and technical interoperability of different types of data	Capability to connect to interoperable data access infrastructures Motivation to adhere to relevant standards and guidelines on process and semantic interoperability	Q10.What enablers (technical organizational, semantical) are in place for interoperable exchange with external stakeholders (incl national and cross-border)? Q11.What standards (such as data models or specifications) do you adhere to in data processes?
6.Data lifecycle management Operational activities (ranging from data acquisition and creation to data archiving) with an aim to improve the	Voluntary standards exist on the maintenance process of health information but their take-up is low. Standards serve as guidance to ensure quality and accuracy	Capability to ensure quality of data, including metadata (description of data that is captured and processed for the primary purpose) to enable re-use	Q12.What is your data storage policy and how does it support secure access? Q13.Is there an appropriate retention policy to erase or

data availability and quality	of data, including of metadata		review the data stored? Q14.Has data quality maintenance process been defined (along with roles and tasks assigned)?
7.Data infrastructure management Are the software components able to support data management and exchange?	Fragmentation of digital standards and limited technical interoperability	Technical capacity to provide access to data in controlled and secure way Capacity to implement and/or improving technical standards (such as standard APIs, open standards and interoperable data formats) to enable portability of data Participation in communities and networks that strive to harmonize interoperability standards to enable cross-border data exchange	Q15.Do you have an inventory of applications and software components that are involved in data processing? Q16.Do you regularly monitor their coherence with relevant standards (national and EU) Q17.Do you have resources (skills and finances) to adapt your infrastructure to the standardized requirements? Q18.Which networks do you belong to for discussing interoperability standards with peers?

The framework as summarized in Table 2 will be used to guide the empirical research in view of finding answers to the posed research question. Assessment of current capabilities of service providers reflects readiness to comply with the proposed norms and benefit from the opportunities created by the recent EC regulations.

3. Empirical research

3.1. Research methods and process

This thesis focuses on the case of the EU regulations on data governance and their potential impact in the health and care dservice provision. The objective of the case-study is to juxtapose current capabilities of data-intensive service providers, and the regulatory provisions which are expected to increase innovation capacities in the sector. The author explores which configuration of data management capabilities is needed to adopt the new opportunities in providing digital health and care services on the European market.

Based on Yin (2003), the aim of the study is compatible with what the case study method, namely:

- Investigating a contemporary phenomenon within its real-life context and acknowledging that the boundaries between phenomenon and context are not clearly evident;
- The research benefits from the prior development of theoretical propositions to guide data collection and analysis.

The unit of analysis of the empirical study is a small sample of organizations in the domain of health and care who have developed a data-intensive product or service. Such a business model relies on capturing, processing and using data to provide value to end-users of service. The end users may include different groups, such as people with needs to improve their health condition, or professionals that facilitate health and care services for end users. There may be other actors involved in the service provision, such as health care service providers as buyers and facilitators of technology-enabled products, but they are not included as direct units of analysis.

The total population of providers of health technology solutions on the EU market is not known. Solution providers are diverse in the organization type including start-ups, small and medium-sized enterprises as well as innovation teams in research institutions, or units incorporated in large health care service organizations such as hospitals or care homes. The focus of the study is on the provision of innovative data-driven services, independent of the type or size of organization, or if operated by private or public entity.

Following the analytical model developed in Chapter 2, the study relies on qualitative data gathering and analytical methods to enable in-depth case analysis. Qualitative research conducted for this study included interviewing a sample of health and care service providers to gain in-depth understanding of their current data management capabilities.

Interviews were conducted with representatives of service provider organizations who develop technology-enabled solutions in the health and care sector. Companies which have developed data-intensive services were selected. The interviewees were identified by searching networks of actors that have formed communities of practice around digital health and care services. The reason for selecting these networks is that they maintain inventories of health technology solutions that have been considered as 'good practices' according to criteria defined for entries in each inventory. Listed services have Technology Readiness Level at a minimum level of TRL5, which means the service has been validated in real-life environment with end-users.The innovative service has been brought to market in at least one EU state, and the organization aims to scale it up outside the home market.

Being listed as good practice is an opportunity to promote the solution or service in view of scaling up and make the service usable in other contexts. The networks and collections of good practices were selected as follows:

- ECHAlliance is a community of over 20,000 experts and 700+ organizations including government, health & social care providers, leading companies and start-ups, researchers, insurances, patients groups and citizens. The Digital Health Observatory (DHO) and The Digital Health Society (DHS) movements facilitate and promote the transfer of knowledge, experiences and best practices creating a community of knowledge in Digital Health. The repository is available at https://echalliance.com/membership/current-members/
- EIT Health is a knowledge and innovation community of the European Institute of Innovation and Technology (EIT). It maintains a database of startups and scaleups in the domain of digital health and care. The dashboard includes companies who have participated in accelerator programs and/or received investments facilitated by EIT Health. The repository is available at https://startups.eithealth.eu/companies.startups/f/data_type/anyof_Verified
- The list of service providers who have registered as interested stakeholders of Innovation Networks for Active and Health Aging (IN-4-AHA) of which the author is a partner, and have given permission to be engaged in the research activities.

The author collected the list of adressees according to the available description of the service to ensure that the service is collecting, processing and using data in the business model. The interviewees were contacted by snowball method, recruiting respondents by direct requests by the author and by recommendation of other interviewees. There were 12 interviews with representatives from 12 service providers that took place in the period April 13 to May 3, 2022 by means of conference calls by using either Zoom or Teams channel. The intervieweed persons were de-identified and given codes based on the sequence when interview was held (see Appendix 2).

The aim of interviews was to collect information on present capabilities and identify gaps (see Appendix 1 for the interview structure). The questions were based on the data capabilities and indicators framework presented in Chapter 2 (see Table 2). The findings that emerged from collected information cannot be used to make a generalized assessment on data management capabilities in the digital health and care sector. The study did not aim to rate or compare the maturity of data management practices of individual organizations. However, the collected insights allow to analyze current capabilities as well as the organizations´ potential to absorb the impact of proposed data regulations.

The interview was structured according to the proposed indicators and questions proposed by the author (see Appendix 1), but the semi-structured format allowed to elaborate discussion on topics that were prioritized by the interviewee. The interviews began by explaining the goals of the study and a request to reflect on the potential challenges stemming from upcoming EC regulations that would affect scaling up their innovation in the European market. The interviewee was then asked to give a brief introduction of the company and describe the data-intensive product that the company is providing in the domain of health and care.

During interviews, the author also requested the respondents to rate (defined on a Likert scale) how important is the use of data from external sources (such as government, research institutions, other businesses). Using external data sources indicate that the business model relies on the capture of data from outside the customer base which requires ability to exchange data and therefore higher level of maturity in data management function. In their responses, all interviewees rated the importance of external sources as high or very high.

3.2. Findings

The following findings are presented according to categories of data management capabilities as defined by the author in developing the concept of data management capabilities assessment model (see Table 2).

Data Governance

In this category, questions allowed to define the procedures which the company has established to ensure compliance with EU legislation. It was also requested to specify the regulation (acts) of most relevance for the business. To indicate readiness for interoperability, it was explored if the company aims to coordinate compliance throughout the service chain, with vendors and partners.

When respondents were asked about which data management functions have been clearly defined (documented) along with roles and responsibilities determined to operate these functions, the priority was put to handling and protecting personal data such as customers´ health-related data. Mostly, it was the data and privacy protection regulation GDPR that was mentioned as the most important act to be complied with. Stringent data protection regulations were also brought as the main reason for limited or missing access to data from external sources. The companies who have certified their products as medical devices, highlighted also the requirements posed by Medical Device Regulations and In Vitro Diagnostic Medical Devices (MDR/IVDR), which scrutinize technology components' accessibility and security (Int-ID8, Int-ID12).

It was stated that the service providers should not be disturbed or challenged by the regulations but supported to conduct their work and achieve outcomes for the benefit of end users. When prompted on the need to collaborate throughout the service chain to ensure compliance with data-related regulations, the interviewees did not see themselves as initiatior for compliance procedures. Rather, institutions who procure digital solutions for the end users were referred (such as hospitals or government entities). The responsibility for compliance is commonly defined by the contract between the procurerer and the service provider (Int-ID1, Int-ID4, Int-ID9, Int-ID11).
According to one interviewee, standard process of procurement is applied in most cases, and confirmed with the buyer based on what kind of information and parameters of data they need (Int-ID1). In another case, the company is proactively building new features in their software for ensuring privacy and anonymization, and have conducted audits to ensure compliance with privacy protection regulations (Int-ID8).

In two cases, interviewees confirmed that relevant software is already integrated to the service to ensure that all partners follow the same standards and requirements for data handling (Int-ID3, Int-ID11).

People, roles, processes

The author asked to define the roles and responsibilities connected to data management in the organization. In more detail, it was explored if tasks have been assigned for data management functions, incl functions related to sharing data with other businesses or with the government agencies. The interviewees were asked to assess if their team is equipped with sufficient skills to perform these tasks.

Interviewees struggled to separate data management roles from other roles in operations. In some cases, the respondents stressed the priority to define the customer journey as the base for setting up business processes (Int-ID2, Int-ID4, Int-ID9). The general opinion was that service process should be complented by data flow along with applying appropriate standards and regulations. There was a suggestion to use business modelling software to visualize what kind of professional fills the necessary tasks in the service process including data handling functions. (Int-ID4). This suggestion would be useful to trace the data flow, describe the processes, roles and data-related tasks, so it becomes possible to excercise control and improve data management.

In one case, the interviewee noted that with public aagencies as clients, it is customary to define roles and responsibilities in great detail in service contracts (Int-ID1). Contracting process is quite long to confirm all specific aspects, as the public agencies are aiming to alleviate all possible risks. However, internal data processing in the organiztion is generally not documented which means that in case of deviances or discrepancies, tracing can be complicated.

There was a distinct difference between smaller organizations and larger, more mature ones, the latter having a dedicated Data Protection Officer in-house or contracted as external consultant. Start-up phase respondents acknowledged the need for legal expertise but have yet not filled the role with appropriate staff (Int-ID5, Int-ID7).

In a self-critical note, the legal compliance skills were believed not yet to be prioritized, there seems to be need for additional training for staff to address data protection and privacy concerns (Int-ID5, Int-ID7, Int-ID10). Technical skills for managing data were thought to be very good in the respondent companies and generally among the service providers on the market. It was a

general belief that alignment with technical standards and requirements is prioritized over legal compliance procedures.

In small teams, staff has multiple roles and an expert dealing with software development might also be taking care of business model improvement. It was considered to be an organic process to share expertise and educate each other in complementing functions (Int-ID3). However, some interviewees state that more support is necessary in terms of consultancy, coaching, good practices or training programmes, to get knowledge on different aspects of data management, regulatory compliance and how to enable data sharing for secondary use (Int-ID1, Int-ID5, Int-ID7, Int-ID10).

Performance and monitoring

The questions explored if the company is using metrics (KPIs) to analyse the value created by data management, including data sharing and re-use, and if these KPIs are shared through the service chain. The general opinion is that metrics are needed to assess the purpose for which the data is used, i.e. health outcomes. However, systematic monitoring is not yet in place. One respondent said that the value created by data should be measured in terms of economic benefit, such as costs savings or business gains created for the customer – depending if it's a public agency or a B2B contract (Int-ID7). They intend to use this evidence for sales and for scaling up the service provision. An interviewee said that KPIs are established internally by an adopted standard operation procedure (Int-ID8). For a comprehensive evaluation of outcomes, however, it is necessary to take into account all contributors along the customer journey.

In a case where the provider of data exchange platform acts as a facilitator for other businesses and end users, they monitor very closely how data is used and validated in the process (Int-ID3, Int-ID11). These providers have defined sources from which data was collected or generated, the processes and tools for handling and sharing data, highlighting the importance of data transparency.

Processes and methods

In this category, the questions concerned protocols and other documented work flows to capture and use data internally and if these procedures are distinguishing between data sources.

In order to exercise control over data processing, some methods should be used to apply safeguards and enable further data re-use by other actors in the value chain, even in the case of personal data

(the methods can be pseudonymization and anonymization, differential privacy, generalization, or suppression and randomization). It was also explored if data portability is enabled for service users (data subjects). Safeguards are mostly applied at the point of capture of data. The technology solution providers prefer to handle data that is already anonymized, aggregated or otherwise protected by institutions in the data controller role (such as hospitals, municipalities or research institutions).

When asked specifically if the organization enables personal data to be re-used by individual´s consent, some respondents said that in their practice, personal data is not shared at all, or shared only occasionally and then the decision is made on case-basis (Int-ID5, Int-ID7, Int-ID10). Very few respondents affirmed to have consent management protocol (Int-ID1, Int-ID3, Int-ID11). For non-personal data, generally there are protocols for data sharing and re-use by other stakeholders. An interviewee highlighted the case of genetics data which cannot be anonymized but need dedicated methods to handle the information i.e. personal genetic locker (Int-ID11).

Mostly, there has not been siginificant need to issue data to a data subject, enabling data portability and patient mobility between multiple data controllers and processors. According to one interviewee, current customers are not yet interested in enabling revocation and dynamic consent processes concerning personal patient summaries: It was deemed valuable for the end user, but not visible in business processes (Int-ID1).

The exception was an innovative service that provides data portability and data altruism in a controlled and secure manner (Int-ID3). In this case, the business is designed for ensuring privacy by design in the technology solution. The product enables tracking data provenance throughout processing, for the service provider and the customers to feel safe about liability and risk minimization by using a digital consent platform. The respondent stressed that functions for giving consent and revoking it i.e. dynamic consent are paramount for re-using patient data for any secondary purposes.

Data architecture

This category of capabilities includes readiness to connect to interoperable data access infrastructures, including by organizational processes and technical standards to allow for integration It was enquired which platforms or specifications are most commonly known and used for data sharing purposes.

While several networks (such as eHealth network, TEHDAS initiative) are known to dedicate efforts towards taking interoperability into practice, the interviewees highlighted the need to agree on which standards should be commonly used. HL-7 standards were used by some respondents as framework (and related standards) for the exchange, integration, sharing, and retrieval of electronic health information (Int-ID8, Int-ID11).

In semantical standards, HL 7 FHIR was mentioned as the preferred framework with requirements that address health solutions. FHIR or Fast Healthcare Interoperability Resource is a data standard (or 'language') to connect different discrete elements. It is a base set of resources that, either by themselves or when combined, can be used for the majority of common use cases in health technology. (Health Level Seven International 2022)

The general view favoured some standard to be chosen as the starting point for implementing interoperability. It was felt that any given country has failed to deploy a common layer for enabling data exchange. However, it was recommended to refrain from creating the one and only technical interoperability layer but opt for a distributed approach and rather to prefer OpenEHR (Open electronic health record) (Int-ID11, Int-ID12). It is a technology for e-health, consisting of open specifications, clinical models and software that can be used to create standards, and build information and interoperability solutions for healthcare (OpenEHR Foundation 2021).

Data lifecycle management

The questions in this category concerned data hosting policy, retention policy to trace, review or archive the data, ways of ensuring findability of data for further access and data quality control in the service provider´s business operations.

An interviewee stressed to address lifecycle management through contracts with vendors and buyers. The agreement clauses are very specific and define where data is and for how long it is kept at any data points, according to data use - if organisations interact with the data one-off or permanently (Int-ID3).

Hosting is mainly outsourced from third parties, meaning service providers do not keep data in their own servers. Experts affirmed that there is need for very careful storage of sensitive data so it cannot be destroyed even in force majeure circumstances (Int-ID3, Int-ID4). It was generally understood that in order to be compliant, only Europe-based servers and cloud services should be used.

Maintenance of data quality was seen as lagging behind in the health technology sector. SMEs that provide digital solutions are often small and they have low capability to deploy metadata standard and generally pay attention to data findability.

In just a few cases, there was a clear vision on the service provider´s role in ensuring metadata quality (Int-ID3, Int-ID11, Int-ID12). Some interviewees even stated that if data is not in described in coherence with semantic standards, the data does not give value (Int-ID4).

SNOMED CT as a widely used semantics framework was mentioned as an international clinical reference terminology designed for use in electronic health. However, interviewees admitted that they experience problems in maintaining these standards in the field (Int-ID1, Int-ID7). Some interviewees experienced glossaries as overly granular and therefore difficult to implement (Int-ID1). The WHO Family of International Classifications was also suggested for the purpose of classifications.

Infrastructure

The questions explored what software components are used, if companies keep an inventory of applications and which national or international standards and certifications are commonly used on the technical platforms.

The commonly expressed opinion was that in-house engineers have control over the technology that they use or create for themselves and should be able to choose among the most appropriate components of infrastructure. Technology providers should be ready to comply with whatever choice is prevalent on the market (Int-ID1, Int-ID8). In some cases, the company prioritizes renewing installations regularly and have automated version management of software (updates) to support their products´ usability and security (Int-ID1, Int-ID8).

Some companies highlighted the need for data warehouses for storage. It was also recommended for the EC Commission to push for standards for distributed ledger technology (Int-ID3). Relating Internet of Things, ENISA certification in connection with data security was mentioned which needs to be regularly updated. ENISA (European Union Agency for Cybersecurity) has defined guidelines for securing the supply chain for Internet of Things (Skouloudi 2020).

An interviewee stated that their company follows ISO 13485:2016 on software (Int-ID8). This standard specifies requirements for a quality management system where an organization needs to

demonstrate its ability to provide medical devices and related services that consistently meet customer and applicable regulatory requirements (International Organization for Standardization 2016). The company also applies ISO standard ISO14971:2019 which specifies terminology, principles and a process for risk management of medical devices, including software as a medical device and in vitro diagnostic medical devices (International Organization for Standardization 2019). Besides medical devices which need to undergo certification, there are other types of data-intensive solutions and services on the market. These do not need to demonstrate compliance with specific standards but would gain from adhering to the same principles as these represent good practice and give reassurance for investors and end-users.

On the subject of infrastructure standards, the interviewees admitted lack of practical expertise in understanding which normative requirements they need to prioritize and which can be gradually added as business gains maturity (Int-ID5, Int-ID7). The competence is available on the market but inaccessible for start-ups due to high price level. There are similar barriers for information security standards and certification that may hamper scaling up on the market, and prevent entering public procurement schemes. Alignment with standards is ensured with assistance by qualified consultants at a considerable cost (Int-ID6).

Collaboration and networks as instruments for organizational interoperability

Respondents were asked to list networks they have joined and describe the benefit for discussing data exchange practices (including interoperability standards) with peers. Very few respondents said that they have initiated interoperability discussions with partners and stakeholders (Int-ID3, Int-ID4), Int-ID9). Some interviewees were however part of community to share best practices in data sharing and re-use. TEHDAS initiative was mentioned several times, being most directly concerned in preparing the ecosystem of stakeholders for the changes brought be the European Health Data Space. One interviewee recommended TEHDAS to take a firmer hand toward achieving interoperable ecosystem and dictate general frameworks for technical communities (Int-ID4).

eHealth Network was mentioned as consolidating experts on interoperability policies, and European Open science cloud was referred as an expert group working on a specific topic (Int-ID11). An interviewee was leading a regional Gaia-X hub and stressed its role in achieving agreements on common data architecture models that the Gaia-X principles represent (Int-ID7).

Standards association IEE and the HL7 network were mentioned as recognized promoters of semantics standards and data models.

There are many consortia and networks who are working to promote secondary use of data, which gives plenty of opportunities for peer-to-peer exchange. A group was mentioned that engages to improve findability of data assets by creating adequate and interoperable catalogues, according to FAIR principles (Int-ID11). Information security aspects seem not be discussed through formal groups but rather practice-oriented communities. This topic was said to be directed by consultancies who provide certification and auditing for technology providers (Int-ID7).

PlatformUptake.eu is a network of stakeholders in the active and healthy domain, promoting the data sharing and data regulations in the field of open platforms and digital technology.

Patient organizations were mentioned as an important representation of end users´ needs to be considered when improving data sharing practices. The representative organizations are also consulted when preparing legislation by policy makers, or when stakeholders work out recommendations addressed to policymakers (Int-ID4, Int-ID6, Int-ID9).

4. Discussion

The study aims to highlight policy and regulation related challenges that have a major effect on data management practices in service providing organizations. The regulatory package proposed by the European Commission defines policy goals at ecosystem level, such as to stimulate free flow of data and improve access to high-quality data for re-use.
All stakeholders, including providers of data-intensive services need to align their strategy and tactics accordingly. Firstly, they need to ensure compliance with the norms in order to keep operating in the EU market. Secondly, new provisions will create new opportunities to boost innovation and generate business growth. It is however necessary for business organizations to enhance their capabilities to capture the benefits.

From the aspect of compliance obligations for companies, the current legal framework in Europe applicable to the use of health data is largely defined by the GDPR, with the new regulations to be rolled out within the next years. The regulatory package is seen to change norms and define new roles in using data for business purposes (European Commission 2022b). The focus is on increasing interoperability of data-intensive service provision. While regulatory drivers and enablers are defined by external actors, the organization in which data are collected, processed, stored and used is at the core of data governance. For business organizations, the implications of adopting regulatory reforms in their operations are complex. Transition may require new technical solutions (such as platforms to enable data portability), assuming more accountability for data quality, and setting up data sharing support within a company, including hiring or re-training staff and aligning software to commonly agreed standards (Bräutigam *et al*. 2022). The investment, in terms of time and budget, is significant but necessary if technology providers want to keep bringing digital, data-driven innovation to the benefit of patients and health professionals.

Researching the stakeholder ecosystem in the digital health and care domain, it became evident that most of challenges described by service providers are foreseen to be addressed by different policy instruments. There are several institutions and networks active on promoting data governance aspects such as adherence to common technical and semantics standards. However, as the landscape of actors is diverse, the path for policy adoption has not been adequately defined, and there is little knowledge which capabilities would need to be developed to cope with the changes.

Preparing to implement elements from the most recent (proposed between 2020 to 2022) legal package in pursuit of transforming the health and care domain in Europe, interviewed technology providers imply that either the national authorities or the Commission should support the implementation in order for the impact of regulations to be revealed.

A broad approach in implementating interoperability would enable scaling up services within one ecosystem and across borders. If interoperability is not achieved, data remain locked only for the purpose for which it was originally collected, which is the primary purpose of delivering care – and data valorisation by further re-use remains limited for all stakeholders (EIT Health 2021).
Legal interoperability is about "ensuring that individuals and organisations, including public or private organisations, operating under different legal and regulatory frameworks, procurement rules, policies and strategies, can work together" (European Commission 2013).
The capabilities connected to enabling legal interoperability mostly fall in the category of data governance that guides how policies, procedures and standards are incorporated and compliance ensured in the operations and throughout the service chain. Information shared by the interviewed service providers confirms that the importance of compliance with the data-related regulations is acknowledged but their complexity and varied interpretation remains a challenge. Most business organizations make it a priority to ensure privacy and data protection. Stakeholders also highlight the need to meet competition rules and follow safety and quality requirements such as the ones set by MDR/IVDR acts (ECHAction 2021). Complexity increases as no single legal provision such as data protection, intellectual property or competition law can be followed on its own. Instead, a comprehensive approach is required to align the business processes with all requirements.

Among the challenges pointed out by stakeholders, compliance with privacy protection remains a main barrier for data sharing, mostly due to interpretation of how 'personal data' is defined is on national scale, and moreover by different member states. The GDPR requires that all processing of data has a legal basis and that appropriate safeguards are in place to preserve the person´s privacy (European Commission 2016). De-identification methods such as anonymization or pseudonymization can be used to render data non-personal. To alleviate risks that pseudonymization can be reversed, data should be kept securely and only used for legally acceptable secondary purposes such as research (Digital Europe 2021).
Although anonymizing clinical data sets a barrier to provide personalized services, deidentification can be a valid method for purposes of research as well as innovation process such as testing a prototype. The interviewees stressed that anonymization methods should be applied at the point

where data is collected, before it is used for further processing in the technology solution. Even if there are standards on the anonymization of personal health data, they are not sufficiently mainstreamed to support re-use. Stakeholders expect the EHDS to propose harmonized rules on how to anonymize and pseudonymize data. Data intermediary agencies such as Findata in Finland are currently providing this service.

The interviewees feel that initiative for alignment with norms and standards should come from public or private payers who procure the technology. Procurement criteria and contract clauses are the instrument to describe the roles and define responsibilities for ensuring compliance amd standardization (ECHAction 2021). In this case, the more standardized are the procurement rules, the more interoperability is achieved within the national ecosystem and in cross-border service provision. Guidelines on procurement standards have been proposed by the Joint Action supporting the eHEALTH Network, including recommendations to articulate clear functional interoperability requirements in bid proposals, purchases, and contracts *(Ibid.).* It should be noted that these guidelines apply mostly for public procurement purposes. In case technology solutions and services are acquired by private companies directly, the service providers´ responsibility increases. The EHDS makes it clear that for the purposes of re-using data, there will be no distinction between public or private reusers or data holders (European Commission 2022b).

Organisational interoperability refers to the way in which organisations "align their processes, responsibilities and expectations to exchange relevant information" (European Commission 2013). Furthermore, active involvement of the user community can be part of the interoperability component. An example of user-driven needs for interoperability is data portability. Currently, there is little readiness for allowing patients to transfer their data from one service provider to another.

The capabilities connected to enabling organisational interoperability include people, roles and responsibilities to ensure that the tasks are defined and skills are sufficient to ensure effective data management and consistent use of data across the entire organization and along the service chain. As data moves through its lifecycle of value creation, it requires different competences. Sharing data assumes that relevant competences are available in the organization, to carry out tasks such as developing an API, deploying deidentification methods for safeguarding data flow, applying relevant standards for the exchange etc. These are different tasks from basic tasks such as collecting data from customers by a user interface, or operating internal information systems for

back-office functions. The need for various tasks and roles (both legally mandated and as required by the business processes) changes along with the changing needs for competences. An organization, from start-up phase to established business needs new competences that correspond with the growing business complexity, so the number and profile of people who focus on data management may change at each stage (EIT Health 2021). Lack of skilled staff in data management functions was also highlighted in stakeholder consultations on proposed EC acts. The existing competences may not be sufficient to support the organization´s strategy for innovation and scaling up on new markets.

The interviews with service providers showed that data management roles are addressed among other tasks. There is seldom a dedicated position in-house to ensure legal compliance in general or specifically, be responsible for data protection. External consultants are employed to audit processes and help to establish regulatory compliance but also advise on applying standards for data quality or correspond to requirements for security. Technical aspects connected to data capture are prioritized, but other tasks related to data management are loosely divided between engineers and management positions.

The ad hoc nature of data-related skills and roles becomes visible also while measuring performance. If the feedback loop on the performance (i.e., progress and outcomes) of data management procedures are not set, it is difficult to identify capacity gaps. For most interviewees, the concept of monitoring or measuring value created by data remained distant. In very few cases, KPIs were mentioned in terms of business value such as economic gains. Measuring outcomes is also in the interest of policy-makers to get feedback on the effect that the regulations aim for.

Another aspect of organisational interoperability are the processes and methods of handling data in a service provider´s organisation. These are tools for managing and using data by establishing internal control over data flow. Only after internal control is established, can interoperability be created with external actors and partners along the service chain. Differencies in data access conditions cause challenges when bringing data from different sources together (EIT Health 2021). As the roles of data controllers and data processor, and corresponding legal duties may change with each additional source, it creates a lack of transparency towards the data subjects while also inreasing complexity for internal data management function *(Ibid.)*.
Interviewees point out that user consent models used in different countries could set barriers for data interoperability. Where the digital health solution supports a patient through several aspects

of their journey through the health system, consent may be requested at each data collection point and by several different actors, as each may be required to demonstrate that they have legitimate grounds to handle data from users. Stakeholders highlight cases when some data are controlled by the end user who wants to exercise their right (in accordance with the GDPR) to control personal data, the person may request deleting data from an app but may not achieve to delete the same data from the databases held in the healthcare system (EIT Health 2021).

One remedy is consistent use of meta-data to identify data points on a given user. However, it is not sufficient if only one provider in the service chain is able to do it. Regardless of the original source, there should be a way to follow data changes over its lineage (life cycle) and exert control over what metadata gets added to the data at each stage. It would be helpful for business organization to rely on a common metadata catalogue. However, national metadata catalogues currently exist only in a few countries in Europe, such as Finland and Sweden (Bogaert 2022).

Technical interoperability refers to the "inclusion of interface specifications, interconnection services, data integration services, data presentation and exchange, and secure communication protocols" (European Commission 2013). When applying standards, these should preferably be available in an open format i.e. allowing integration with external information systems. Open technical specifications should be adapted to the specific context in which they will be used. Wide use of internationally recognized standards and open technical specifications allows to achieve interoperability of data, information systems, and services even in different jurisdictions. An example of such open system is the OpenEHR (Open electronic health record). It is a technology for digital health consisting of open specifications, clinical models and software that can be used to create standards and build interoperability solutions (OpenEHR Foundation 2022).

The capabilities connected to enabling technical interoperability are data infrastructure and data lifecycle management, including quality management. The discussion on data sharing infrastructure focused on secure processing environments for health data. The interviews revealed low awareness on central data exchange infrastructure that would be provided by the European Commission for data processing purpose. Diverse platforms are currently in use that come with different readiness for integration with other systems. Interviewed companies expected that technical standardization should come from public procurers who are interested that innovative solutions are compatible with information systems they already have. Service providers should be ready to integrate with any infrastructure for data processing on the market where they operate in.

The strive towards open formats increases along with the company´s ambition to scale to new markets and avoid cumbersome and time-consuming integration.

Any interoperability assets meant for cross-border health data exchange need to be localized in the national contexts in member states. The first elements, the European Patient Summary and ePrescription cross border information services have leveraged on common European technical specifications, the eHealth Digital Service Infrastructure (ECHAction 2021). The Patient Summary defines a minimal dataset with the most generic aspects of a person's healthcare history. However, these emerging unified solutions may not be relevant for service providers who do not need to establish connection to patient summary but provide other digital personalised services. In addition to conducted interviews, other stakeholder discussions have highlighted expectation on the facilitating role and responsibilities of health data access bodies which are foreseen to be founded in each member state to facilitate access to electronic health-related data (TEHDAS 2022a). However, as of December 2022, there were just a few countries in Europe ready to establish national contact points that enable access to public databases and mediate data permits (Bogaert 2022). Hence most service providers need to seek data access from disparate databases and send in requests from each individual registry.

Providing health services by digital applications that collect data from multiple sources creates challenges for maintaining data quality. To align standards for describing the qualities of software, the EHDS proposes a system for voluntary labelling of wellness applications, i.e., digital health applications that are not certified as medical devices. The discussion among stakeholders evolves around the extent of the data to be included in these labels, to notify on the quality and utility of data for re-use purposes. Stakeholders also suggest that in cases where the infrastructure for digital health services includes the use of tools (hardware and software) held by several parties, criteria for certifying data quality must be established along with ensuring data security (TEHDAS 2022a, EIT Health 2021).

The capabilities connected to enabling semantic interoperability are data architecture and data models along with semantic classifications.

Semantic frameworks are necessary to increase the portability and usability of patient information (European Commission 2013). "Reading" data from different sources and applications relies on using standardised formats such as modelling and coding standards (ECHAction 2021). Classifications are used to support data use across the health ecosystem. WHO has developed

reference classifications that can be used to describe the health state of a person at a particular point in time, such as International Classification of Diseases (ICD), International Classification of Functioning, Disability and Health (ICF) and the International Classification of Health Interventions (ICHI) (World Health Organization 2019).

HL7 provides a framework (and related standards) for the exchange, integration, sharing, and retrieval of electronic health information. HL7 standards support clinical practice and the management, delivery, and evaluation of health services, and are recognized as the most commonly used in the world (Health Level Seven International 2022). HL7 FHIR standards reference existing terminologies, classifications and coding standards, such as ICD or SNOMED (OECD 2019).
The interviews revealed that standardization is not prioritised by most technology providers, rather by individual experts who are familiar with clinical (i.e. business) processes and are concerned how accurately these are captured by information systems. Respondents who are involved in linking technology standards with clinical terminology even stated that if data is not in coherence with semantics, the data does not yield any value.

The interviewees supported a wider use of industry standards, to be promoted by standardisation associations (such as HL7 and IEEE) and other collaborative networks on digital health. However, it was upheld that the regional and national policy makers have a specific role in supporting the deployment of the standards in practice. There needs to be an informed and agreed-upon choice of all actors in any given context, which standards frameworks to promote and adhere to. This can only be achieved by persistent coordination by an institution which has a strong mandate - either by rule-making, reimbursement of health and care services, or both. WHO has recommended that a systematic approach to the adoption of eHealth standards for data exchange and interoperability needs to be taken, with a national body in each member state clearly identified to govern this process (World Health Organization 2016).
Moreover, communication and training are necessary to reach the critical adoption of common standards by stakeholders and thus build interoperability of technological platforms and services.

4.1. Policy recommendations to boost innovation capacity

Based on the findings collected in the course of the case study, the author presents some recommendations for improving data management capabilities. While the ambition to create value

by data, generate growth and provide benefits for customers drives the organization internally, the technology providers must also adapt to enablers and challenges in the external context. They are expected to ensure compliance with regulations, deploy standards and meet requirements. The following policy recommendations target the boundary space of combining opportunities coming from external context with business strategy and management decisions.

- Encouraged by EC regulations, data sharing and re-use are expected to significantly grow in the near future. Service providers recognise the benefits of these enablers and express their intention to increase sharing and re-using data. However, business actors look to policymakers in national as well as EU-level institutions to facilitate access to data. There is also need to clarify the roles and responsibilities of the health data access bodies in each member state which are foreseen to facilitate access to electronic health-related data.
- The policies should build on responsible data sharing initiatives driven by industry, promoting open technologies such as the Open Electronic Health Record. Community initiatives would in turn benefit from national and EU level decisions on harmonising how personal health information can be shared. For example, the institutions could agree on appropriate tools and security standards that would be used to process sensitive health data in de-identified format, depending on the purpose of re-use.
- Service providers acknowledge that protecting users´ health data is key to the enhancing trust in new health technologies. At the same time, the regulations grant individuals control who accesses their data and for which uses, to restrict access, and to give consent for transmitting information from one health care provider to another. For service providers, it is difficult to balance these aspects in their internal data policies and business processes. There is need to raise competences as well as make significant investments to enable user´s control in order to ahieve safe and smooth data portability.
- It is recommended to agree on a commonly used health data classifications or alternatively, recommend and disseminate standards developed by industry networks and coordinate their use in public as well as private procurement and provision of digital services.
- Service providers need support in their endeavours to scale up on the market and provide the best possible solutions for users. This can be facilitated by procurers, addressing issues such as recognition and reimbursement of innovative technologies and services across borders, similar to German DIGA system as an example.

- Support instruments should be accessible for all actors in the digital health and care domain, however targeted according to the specifics of their roles in the ecosystem. Since there is a lack of general data literacy and moreover, specific skills related to handling and analyzing data, there is need to train staff in service providing organizations and share best practices to learn from capacity-building experience across EU countries.
- The support can take the form of additional funding made available for market actors, by instruments that link research and development with business goals to boost innovation; by strengthening institutions such as accelerators and innovation hubs that can advise how to align business operations with new requirements and opportunities. Next Generation funding schemes could be used to enhance digital and data-related capabilities of all actors in the health and care systems.

Conclusion

For analyzing deployment of innovation, capabilities of market actors should be aligned with enablers created by context such as policies and regulations. This case study explored what configuration of data management capabilities in service providing organizations is needed in view of the new provisions for regulating data economy in Europe. Readiness to adapt to new opportunities will affect the capacity for adopting, upscaling and transferring technological solutions in the provision of digital health and care services.

The empirical analysis was conducted based on case-study method. Innovative, data-intensive services are varied in scope and character, and the group of organizations that put these services on the market is diverse. Interviews with representatives of service provider organizations allowed to collect insights and reflect on their current capabilities. It became evident that data governance in these organizations is subject to significant challenges stemming from regulatory context, specifically around ensuring legal, organizational and technical interoperability.

The European Commission has proposed data-related regulations with an aim to "enable scaling up innovative solutions and elaborate the pan-European interoperability of technological innovation" (European Commission 2020c). Rather than setting strict norms to comply with, the focus is on creating enablers to gain better access, share and move data between stakeholders, sectors and member states.

As the study demonstrates, stakeholders share the policymakers view that regulatory enablers will provide businesses with opportunities for innovation and business growth. However, it became evident from research that business organizations´capabilities are not yet aligned with regulatory enablers. More effort and investment is needed to start benefitting from the opportunities.

As organizational data management functions and practices vary, so do the capabilities. The capabilities framework and indicators developed for the study covers a comprehensive scope of data governance aspects as described in field-specific literature.

The collected insights on the current state of capabilities are by no means representative for the whole group of data-intensive service providers and generalized conclusions cannot be drawn.

However, collected evidence shows that limited data capabilities in the organization present challenges for sharing data with other stakeholders, either for the primary purpose of care or re-use purposes. Data exchange depends on the ability to have authority and control over data, which

includes safeguarding its integrity and quality, labelling and describing data, and securely granting access rights for data subjects and re-users. These data management functions pave the way for re-use, making data FAIR: findable, accessible, interoperable and reusable (TEHDAS 2021a).

The more effort and investment is put into improving data management capabilities, the better are the outcomes in terms of value - such as data accuracy and timeliness, secure access and user control over their data. For the business outcomes, the value created by fit-for-purpose data can be invested back in the organization, for improving the services and scaling these up on a larger market. The study showed that theory for innovation capabilities can be used to elaborate an assessment framework for data management capabilities and contribute to management practice. The framework of capabilities and indicators proposed by the author can support technology providers in digital health and care domain for further analysis on their data management capabilities and undertake appropriate managerial actions.

Further research could focus on the capabilities for handling personal data for the purpose of delivering personalized care. The potential for scaling up innovation could be supported by guidance and tools such as a code of conduct for data governance to be used by different stakeholders in the health care ecosystem.

Market actors who are capable of seizing the opportunities in the regulatory context have better capacity for innovation. In turn, data-driven innovation that accelerates data mobility within the EU will produce better health outcomes for everyone. The impact created by contextual enablers can only be measured after regulations are implemented. Therefore, further studies are needed to ascertain whether the regulatory package will have a positive effect on the data economy in Europe.

LIST OF REFERENCES

Alhassan, I., Sammon, D., Daly, M. (2016). Data governance activities: an analysis of the literature. *Journal of Decision Systems,* 25:sup1, 64-75.

Abraham, R., Schneider, J., vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management* 49, 424-438.

Akhtar, P., Khan, Z., Tarba, S., Jayawickrama, U. (2018). The internet of things, dynamic data and information processing capabilities, and operational agility. *Technological Forecasting and Social Change* 136, 307–316.

Barker, J. M. (2016). *Data Governance: The missing approach to improving data quality*. A Dissertation Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Management in Organizational Leadership with a Specialization in Information Systems and Technology. University of Phoenix.

Bharadwaj, A. R. S. (2000). "A Resource-Based Perspective on Information Technology Capability And Firm Performance: An Empirical Investigation," *MIS Quarterly,* vol. 24, no. 1, 169–196.

Bogaert, P. (2022, 5 December). *European Health Data Space* [Conference presentation]. EIT Matchmaking event 2022, Tallinn, Estonia.

Bräutigam, T., Aholainen, M., Cunningham, F., Geus, M., Toivanen, M., Kukorelli, F. (2022). *EU regulation builds a fairer data economy. The opportunities of the Big Five proposals for businesses, individuals and the public sector.* SITRA Working paper.

Cepeda, J., Arias-Perez, J.(2019). Information technology capabilities and organizational agility: the mediating effects of open innovation capabilities. *Multinational Business Review* 27(2), 198–216.

Chan, C. M., Teoh, S.Y., Yeow, A., Pan, G.(2019). Agility in responding to disruptive digital innovation: case study of an SME. *Information Systems Journal* 29 (2), 436–455.

Colombo, F., et al. (2011). *Help Wanted? Providing and Paying for Long-Term Care*. Paris: OECD Publishing. Retrieved from https://doi.org/10.1787/9789264097759-en, 11 May, 2022

Competence Center Corporate Data Quality (CC CDQ). *Data excellence model.* Retrieved from https://www.cdq.com/events-insights/research/data-excellence-model, 8 January, 2022.

DAMA International (2009). *The DAMA guide to the data management body of knowledge (DAMA-DMOK)*. New Jersey: Technics Publications, LLC.

Davenport, T. H., Harris, J. G., De Long, D.W., Jacobson, A. L. (2001). Data to Knowledge to Results: building and analytic capability. *California Management Review Reprint Series,* Vol 43, No 2.

Digital Europe. (2021). *Digital Europe recommendations on the European Health Data Space.* Retrieved from https://www.digitaleurope.org/resources/digitaleuropes-recommendations-for-the-european-health-data-space/, 22 February, 2022.

Digital Europe. (2020). *Improving lives and managing diseases through a data driven EU healthcare system*. Retrieved from https://www.digitaleurope.org/resources/improving-lives-and-managing-diseases-through-a-data-driven-eu-healthcare-system/, 19 February, 2022.

ECHAction. (2021). *Interoperability Guide.* Retrieved from http://ehaction.eu/interoperability-guide/about-the-guide/, 2 April, 2022.

EIT Health. (2021). *Learning from health data use cases. Real-world challenges and enablers to the creation of the European Health Data Space*. Retrieved from https://eithealth.eu/wp-content/uploads/2021/11/EHDS_report.pdf, 2 January, 2022.

Enterprise Data Management Council. (2014). *Data management capabilities assessment model* (DCAM). Retrieved from https://edmcouncil.org/page/aboutdcamreview#, 3 January, 2022.

European Commission. (2022a). *European Declaration on Digital Rights and Principles for the Digital Decade.* Retrieved from https://digital-strategy.ec.europa.eu/en/library/declaration-european-digital-rights-and-principles, 4 April, 2022.

European Commission. (2022b). *Proposal for a regulation: the European Health Data Space.* Retrieved from https://ec.europa.eu/health/publications/proposal-regulation-european-health-data-space_en, 3 May, 2022.

European Commission. (2021). *Public consultation on Data Act. Summary report on the public* consultation. Retrieved from https://digital-strategy.ec.europa.eu/en/public-consultation-data-act-summary-report, 4 January, 2022.

European Commission. (2020a). *A European Strategy for Data COM/2020/66.* Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066, 10 January, 2022.

European Commission. (2020b). *Communication: Shaping Europe's digital future.* Retrieved from https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/shaping-europe-digital-future_en, 4 January, 2022.

European Commission. (2020c). *European Health Data Space Initiative. Combined evaluation roadmap/Inception impact assessment*. Retrieved from https://www.vci.de/ergaenzende-downloads/2020-12-23-inception-impact-assessment-european-health-data-space.pdf, 4 January, 2022.

European Commission. (2020d). *Data Act. Inception impact assessment.* Retrieved from https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13045-Data-Act-&-amended-rules-on-the-legal-protection-of-databases_en, 2 February, 2022.

European Commission. (2020e) *Proposal for a regulation of the European Parliament and of the Council on European data governance (Data Governance Act) COM/2020/767.* Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0767, 2 March 2, 2022.

European Commission. (2020f). *White paper on Artificial Intelligence - A European approach to excellence and trust.* Retrieved from https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en, 3 February, 2022.

European Commission. (2018a). *Communication on enabling digital transformation of health and care in the Digital Single Market; empowering citizens and building a healthier society*. Retrieved from

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A233%3AFIN, 11 January, 2022.

European Commission. Directorate-General for Communications Networks, Content and Technology, Scaria, E., Berghmans, A., Pont, M., et al (2018b). *Study on data sharing between companies in Europe: final report.* Retrieved from *https://data.europa.eu/doi/10.2759/354943* , 3 April, 2022.

European Commission. (2017a). *Communication on the mid-term review of the implementation of the Digital Single Market Strategy* (COM (2017)228). Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52017DC0228, 21 November, 2022.

European Commission. (2017b). Regulation (EU) 2017/745 on medical devices (MDR). Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0745, 3 March, 2022.

European Commission. (2016). Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Retrieved from https://eur-lex.europa.eu/eli/reg/2016/679/oj, 3 March, 2022.

European Commission. (2013). *eHealth European Interoperability Framework Study Report.* Retrieved from https://www.digitalhealthnews.eu/download/publications/3474-ehealth-interoperability-framework-study, 23 April, 2022.

European Commission. (2011a). *Synthesis report on the public consultation on the European Innovation Partnership on Active and Healthy Ageing*. Retrieved from https://www.kowi.de/Portaldata/2/Resources/fp/EIP-AHA-consult-report.pdf 21 May, 2020.

European Commission. (2011b). Directive 2011/24/EU of the European Parliament and of the Council on the application of patients' rights in cross-border healthcare. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32011L0024, March 3, 2022.

European Observatory of Health Systems and Policies. (2021). *State of health in the EU: Companion report 2021*. Retrieved from https://eurohealthobservatory.who.int/publications/m/state-of-health-in-the-eu-companion-report-2021, 1 May, 2022.

Fu, X., Wojak, A., Neagu, D., Ridley, M., & Travis, K. (2011). Data governance in predictive toxicology: A review. *Journal of Cheminformatics,* 3(1), 1–16. Retrieved from http://dx.doi.org/10.1186/1758-2946-3-24, 20 April, 2022.

Health Level Seven International. *Introduction to HL7 Standards*. Retrieved from https://www.hl7.org/implement/standards/index.cfm?ref=nav, 25 April 2022.

International Organization for Standardization.(2016). *Medical devices — Quality management systems — Requirements for regulatory purposes* (ISO standard No.13485:2016). Retrieved from https://www.iso.org/standard/59752.html, 25 April, 2022.

International Organization for Standardization. (2019). *Medical devices — Application of risk management to medical devices* (ISO standard No. 14971:2019). Retrieved from https://www.iso.org/standard/72704.html, 25 April, 2022.

Kalmuk, G., Acar, A. Z.,(2015). The effect of quality orientation and learning orientation on firm performance. *Research Journal of Business Management* 2 (4), 455–487.

Karo, E., Kattel, R. (2016). Innovation and the State: Towards an Evolutionary Theory of Policy Capacity. *Working Papers in Technology Governance and Economic Dynamics* no. 72.

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM,* 53, 148–152. Retrieved from http://dx.doi.org/10.1145/1629175.1629210, April 3, 2022.

Kim, H. Y., & Cho, J.-S. (2017). Data governance framework for Big data implementation with a case of Korea. *IEEE 6th International Congress on Big Data*, 384–391.

Kopanakis, I., Vassakis, K., Mastorakis, G. (2016). Big Data in Data-driven innovation: the impact in enterprises' performance. *Proceedings of 11th Annual MIBES International Conference.* 257–263.

KPMG International. (2018). *Data governance: driving value in health care.* Retrieved from https://home.kpmg/xx/en/home/insights/2018/06/data-governance-driving-value-in-health.html, 10 January, 2022.

Kütt, A., Kala, K., Marcus, F., Hinsberg, H. (2022). *Data Governance Guidebook.* Retrieved from https://innovation4ageing.tehnopol.ee/tools-and-outputs/project-outputs/, August 12, 2022.

Lawson, B., Samson, D. (2001). Developing innovation capability in organisations: a dynamic capabilities approach. *International Journal of Innovation Management* 5 (3), 377–400.

Li, H., Wu, Y., Cao, D., Wang, Y.(2021). Organizational mindfulness towards digital transformation as a prerequisite of information processing capability to achieve market agility. *Journal of Business Research* 122, 700–712.

Lu, Y., Ramamurthy, K. (2011). Understanding the link between information technology capability and organizational agility: An empirical examination. *MIS Quarterly,* vol. 35, no. 4, 931–954.

Lundvall, B. (2007). National innovation systems – analytical concept and development tool. *Industry and Innovation,* 14(1), 95-119.

Lyon, D. W., Ferrier, W. J. (2002). Enhancing performance with product-market innovation: the influence of the top management team. *Journal of Management Issues* 14 (4), 452–469.

Martins, H. (2021). *EU health data centre and a common data strategy for public health.* European Parliamentary Research Service, Scientific Foresight Unit (STOA). Retrieved from https://www.europarl.europa.eu/RegData/etudes/STUD/2021/690009/EPRS_STU(2021)690009_EN.pdf, 8 March, 2022.

Oderkirk, J. (2021). *Survey results: National health data infrastructure and governance.* OECD Health Working paper no 127. Paris: OECD Publishing. Retrieved from https://doi.org/10.1787/55d24b5d-en, 10 May, 2022

OECD. (2019). *Health in the 21st Century: Putting Data to Work for Stronger Health Systems.* Paris: OECD Publishing. Retrieved from https://doi.org/10.1787/2074319x, 9 December, 2022

OECD. (2015). *Health Data Governance: Privacy, Monitoring and Research - Policy Brief.* Paris: OECD Publishing. Retrieved from https://doi.org/10.1787/9789264244566-en, 10 May, 2022

OpenEHR Foundation. (2022). Open industry specifications, models and software for e-health. Retrieved from https://www.openehr.org/about/what_is_openehr, 30 April, 2022.

Otto, B. (2011c). Organizing data governance: Findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems,* 29(3), 45–66.

Ramamurthy, K., Lu, Y. (2011). Understanding the link between information technology capability and organizational agility: An empirical examination," *MIS Quarterly,* vol. 35, no. 4, 931–954.

Schlieter, H., Marsch L. A., Whitehouse D., Otto L., Londral A.R., Teepe G. W., Benedict M., Ollier J., Ulmer T., Gasser N., Ultsch S., Wollschlaeger B., Kowatsch T. (2022). Scale-up of Digital Innovations in Health Care: Expert Commentary on Enablers and Barriers. *Journal of Medical Internet Research*, 24(3):e24582.

Shuradze, G., Wagner, H-T. (2016). Towards a Conceptualization of Data Analytics Capabilities. *49th Hawaii International Conference on System Sciences ((HICSS).* Retrieved from https://www.researchgate.net/publication/301281556_Towards_a_Conceptualization_of_Data_Analytics_Capabilities, 14 March, 2022.

Skouloudi, C., Malatras, A., Naydenov, R., Dede, G. (2020). *Guidelines for Securing the Internet of Things.* Retrieved from https://www.enisa.europa.eu/publications/guidelines-for-securing-the-internet-of-things, 24 April, 2022.

Sultana, S., Akter, S., Kyriazis, E. (2022). How data-driven innovation capability is shaping the future of market agility and competitive performance? *Technological Forecasting & Social Change*, 174(8):121260. School of Business, University of Wollongong.

Teece, D., Peteraf, M., Leih, S. (2016). Dynamic capabilities and organizational agility: risk, uncertainty, and strategy in the innovation economy. *California Management Review* 58 (4), 13–35.

TEHDAS. (2022a). *European health projects meet to discuss European health data space proposal.* Retrieved from https://tehdas.eu/news/european-health-projects-meet-to-discuss-european-health-data-space-proposal/, 12 August, 2022.

TEHDAS. (2022b). *Report on secondary use of health data through European case studies*. Retrieved from https://tehdas.eu/app/uploads/2022/08/tehdas-report-on-secondary-use-of-health-data-through-european-case-studies-.pdf, 3 March, 2022.

TEHDAS. (2021a). *Why health is a special case for data governance*. Retrieved from https://tehdas.eu/app/uploads/2021/06/tehdas-why-health-is-a-special-case-for-data-governance-2021-06-23.pdf, 3 March, 2022.

TEHDAS. (2021b). *Catalogue of EHDS services for secondary use of health data*. Retrieved from https://tehdas.eu/app/uploads/2021/12/tehdas-catalogue-of-ehds-services-for-secondary-use-of-health-data-2021-12-09.pdf, 4 March, 2022.

TEHDAS. (2021c). *Identifying those data quality features that could be legally bound and providing advice to the European Commission.* Retrieved from https://tehdas.eu/app/uploads/2021/11/tehdas-identifying-those-data-quality-features-that-could-be-legally-bound-2021-11-11.pdf, 4 March, 2022.

Van Veen, E.–B. (2018). Observational health research in Europe: understanding the General Data Protection Regulation and underlying debate. *European Journal of Cancer Volume* 104, 70-80. Retrieved from https://www.sciencedirect.com/science/article/pii/S0959804918314023, 4 April, 2022.

World Health Organization. (2019). International Classification of Health Interventions (ICHI). Retrieved from https://score.tools.who.int/tools/optimize-health-service-data/tool/international-classification-of-health-interventions-ichi-131/, 22 April, 2022.

World Health Organization. (2016). *From innovation to implementation: state of eHealth in the WHO European region.* World Health Organization. Regional Office for Europe. Retrieved from https://apps.who.int/iris/handle/10665/326317, 3 December, 2021.

World Health Organization. (2009). *Practical guidance for scaling up health service Innovations.* Retrieved from https://apps.who.int/iris/handle/10665/44180, 3 December 2021.

Yin, R. K. (2003). *Case study research: design and methods.* (5th ed.) Thousand Oaks: Sage Publications.

Zolnowski, A., Christiansen, T., Gudat, J. (2016) Business model transformation patterns of data-driven innovations. *Twenty-Fourth European Conference on Information Systems (ECIS), İstanbul, Turkey*. Retrieved from https://aisel.aisnet.org/ecis2016_rp/146, 5 April, 2022.

APPENDICES

Appendix 1. Interview structure

I Introduction

In your business activities, how important is the use of data from external sources (such as government, research institutions, other businesses)?

- o Extremely important
- o Important
- o Moderately important
- o Slightly important
- o Not at all important

II Questions according to categories of data management capabilities

CATEGORY	QUESTIONS
Data Governance	1. What procedures have you established to ensure compliance with EU legislation - please refer to the specific regulation that is of relevance for your business? (e.g. in areas of privacy and data protection policy, data exchange policy, IoT or AI-enabled processes etc) 2. Do you coordinate compliance throughout the service chain, with vendors and partners?
People, roles, processes	3. Have tasks been assigned for data management functions, incl functions related to sharing (B2B and B2G)? 4. Is your team equipped with sufficient skills to perform these tasks?
Performance and monitoring	5. Which metrics (KPIs) do you use to understand the value created by data management, including data sharing and re-use? 6. Are these KPIs distributed across the organization? Through the service chain?

Processes and methods	7. Have you defined (=described by schemes, mapped the data flow) your internal processes to capture data from public sources? How is it different than data collection other sources (research, B2B, users)? 8. Are you able to apply safeguards (e.g. pseudonymization and anonymization, differential privacy, generalization, or suppression and randomization) to share personal data with other actors? 9. Do you enable data portability for users (data subjects)? Has there been need for it?
Data architecture	10. What enablers (technical means, organizational process and agreements, semantical standards etc) do you use for interoperable exchange with external stakeholders (incl national and cross-border)? 11. What standards (such as data models or specifications) do you use for data sharing?
Data lifecycle management	12. Does your data hosting policy support secure access (how is it ensured)? 13. Is there a retention policy to erase or review the data stored? 14. Has data quality maintenance process been defined (along with roles and tasks assigned)?
Data infrastructure management	15. Do you keep inventory of applications and software components that are involved in data processing? 16. Do you apply relevant standards (national and EU) (e.g. ISO or other certifications?) on the software for data processing?
General collaboration	17.Which networks do you use discussing data exchange (incl interoperability standards) with peers?

Appendix 2. List of interviews with codes

Code ID	Interviewed person(s)	Description of service/company	Date of interview
Int-ID1	A L, (Spain)	The organization provides technology to collect data on personal dietary requirements such as allergies and link these to nutrition and food delivery in health care provision.	April 11, 2022
Int-ID2	G C, (Slovenia)	Organization is a cluster for digital innovation in healthcare, offering mentoring and accelerator programs for starting health technology companies.	April 13, 2022
Int-ID3	F L, (Sweden)	Organization provides a data sharing platform to enable data portability and data altruism.	April 19, 2022
Int – ID4	P E, (Finland)	The expert has lead several research projects for the active and healthy aging domain. Among other results, he has contributed to defining minimal dataset for frailty prevention guidelines and developed taxonomies and glossaries related to diagnosis. He is currently involved in business process modelling for health and social care services.	April 20, 2022
Int-ID5	M S, (Spain)	The organization is a start-up that provides algorithms-based decision support	April 21, 2022

		technical solutions in health care.	
Int-ID6	A N and F L,	The organization provides an inventory of open service platforms in the Active and Healthy Ageing domain, covering both open platforms and partly-open/proprietary platforms developed by industry, and addresses the interactions between these platforms.	April 25, 2022
Int-ID7	I P, (Spain)	The organization is a technology hub focused in 4 technology areas (Internet/Web of Things, Intelligent Data Analysis, W3C Standardization and Open Data, Vision Technologies) and 3 societal challenges (Digital Growth, Smart Territories and ICT for Active Ageing and Wellbeing).	April 22, 2022
Int-ID8	C S, (Hungary)	The organization is a digital transformation consultancy, helping clients to prepare MDR Technical Documentation, auditing, performing risk management and medical device software LifeCycle Management.	Submitted responses in writing
Int-ID9	N L, (Belgium)	The organization focuses on data-driven research and development. The data sharing platform gives citizens ownership of their own data and	April 28, 2022

		enable to manages their data in safe environment.	
Int-ID10	Á R, (Spain)	The organization provides Enterprise Resource Planning (ERP) solutions to health care organizations.	April 29, 2022
Int-ID11	S B, (Netherlands)	The organization develops technical solutions with special encryption of personal records. The platform facilitates the sharing of information between individuals and organisations.	May 2, 2022
Int-ID12	A D, (Slovenia)	The organization provides an openEHR data platform designed to store, manage, query, retrieve and exchange structured electronic health record (EHR) data.	May 3, 2022

Appendix 4. Non-exclusive licence

A non-exclusive licence for reproduction and publication of a graduation thesis[11]

I Hille Hinsberg (*author's name*)

1. Grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis **Data governance in digital health care: enablers and capabilities**

supervised by Veiko Lember (*supervisor's name*)

1.1 to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;

1.2 to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.

2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.

3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

December 30, 2022 (date)

[1] *The non-exclusive licence is not valid during the validity of access restriction indicated in the student's application for restriction on access to the graduation thesis that has been signed by the school's dean, except in case of the university's right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.*