Bio-curity

D4.2 Trust Model Technology Architecture

DELIVERABLE VERSION	V1.0
WORK PACKAGE	WP4 Data sovereignty, trust and security
TYPE	Document
CONTRACTUAL DATE OF DELIVERABLE	30/06/2024
ACTUAL DATE OF DELIVERABLE	01/07/2024
RESPONSIBLE PARTNER	TNO (TNO 2024 P11375)
DISSEMINATION LEVEL	Public
STATUS	Final

Eureka

ECSEL

1. Table of content

1. Table of content 2
1. Introduction 4
1.1 Scope 6
2. Methodology 7
2.1 Criteria selection 7
2.2 Requirements gathering and architecture design 8
3. Requirements 10
3.1 Functional Requirements 10
3.2 Non-Functional Requirement 11
4. Patient-centric and Decentralized Approaches 13
Blockchain 13
Self-Sovereign Identity 13
Persoonlijke Gezondheidsomgeving 14
Discussion 14
5. Architecture 16
5.1 Technology 16
Self-Sovereign Identity 16
Personal Health Environment 16
5.2 Overview 17
Orchestrating Server 17
Personal Health Environment 18
Sensor provider 18
Researcher 18
5.3 Processes 18
6. Criteria 27
6.1 Consent management 27
Current consent models 28
Purpose for processing 29

Bio-curity consent model 30
6.2 Access control 31
Attribute-Based Access Control 32
Authorization of access 33
6.3 Confidentiality 34
Implementing confidentiality in Bio-curity 35
6.4 Right to share data 35
State-of-the-Art 36
Implementing rights to share data in Bio-curity 37
6.5 Encryption 37
6.6 Data Pseudonymization 38
7. Discussion and Conclusion 39
8. References 40

Document history

Version	Date	Author/editor	Description
v1.0	28-06-2024	Dayana Spagnuelo, Kevin Witlox	Deliverable complete
V0.9	14-06-2024	Dayana Spagnuelo, Kevin Witlox, Joanna Morozowska	Deliverable ready for review
V0.6	06-06-2024	Dayana Spagnuelo, Kevin Witlox, Joanna Morozowska, Burcu Bektaş Güneş	Draft of Methodology, Architecture, Criteria
V0.4	22-05-2024	Dayana Spagnuelo, Kevin Witlox	Draft of Introduction, Patient-centric and decentralized approaches
V0.2	19-04-2024	Dayana Spagnuelo, Kevin Witlox	Draft of Requirements, Outline of document

Document review

Date	Reviewer	Description
15-06-2024	Peter Langenkamp	Review internal at TNO
23-06-2024	Serhat Taşkale	Review Kafein

Partners involved

Company	Name and surname	Email address
TNO	Kevin Witlox; Dayana Spagnuelo	kevin.witlox@tno.nl; dayana.spagnuelo@tno.nl
MEDrecord	Joanna Morozowska; Burcu Bektaş Güneş; Jan-Marc Verlinden	joanna@medrecord.io; burcu@medrecord.io; jan-marc@medrecord.io
Lostar	Murat Lostar	murat.lostar@lostar.com.tr
Medron Tech	Eren Mert; Fatih Alparslan	eren.mert@medrontech.com; fatih.alparslan@medrontech.com
Vestel	Ilhan Kaya; Ismail Ebeperi	ilhan.kaya@vestel.com.tr; ismail.ebeperi@vestel.com.tr
Kafein	Serhat Taşkale	serhat.taskale@kafein.com.tr

1. Introduction

The introduction of new medical technologies, such as sensors that measure pathophysiological and behavioral processes and can be offered as a wearable, has accelerated the process of collecting patient data for relevant clinical decisions. These technologies allow for several medical developments, such as the creation of digital biomarkers, which in turn can be used in the better care of patients. However, these advancements necessitate the intensive sharing of sensitive data, raising significant privacy and security concerns. How to support such data sharing in a trustworthy and secure manner is the subject of this work.

When we talk about healthcare data, one important term is the Electronic Health Record (EHR). Every healthcare institution records data about their patients, documented by healthcare professionals. Currently, in most of the countries worldwide, including the Netherlands, there is no single Electronic Health Record of a patient. Every healthcare institution records their own version for each of their patients. As such, a patient will accumulate an EHR at each institution they are treated, and there is no easy method for either the patient or the healthcare institution to see the full overview of the patient's health record, without manually collecting all the pieces of information scattered across the institutions.

The lack of a unified method of sharing data about a patient also hinders innovations in healthcare, such as the use of the aforementioned medical sensors. A patient could significantly benefit from keeping a Personal Health Record (PHR), an extension to the EHR, in which the patient themselves may contribute to the record. In a PHR, a patient could contribute occasional questionnaires on pain-levels or mental health, or provide the data generated by medical wearables. The absence of a unified EHR system, and the lack of patient access to any EHR make this concept difficult to develop.

Currently, the sharing of EHRs between institutions is often still being done via either scanning or by retyping physical dossiers. The Netherlands has recently started to require that healthcare institutions share EHRs between institutions electronically [1], though the exchange is still in the process of being standardized. This standardization of the EHR will help, but the challenge remains that these copies of the EHRs are statically stored at each institution.

Sharing of Electronic Health Records (EHRs) in Turkey faces similar challenges to those seen in other countries. While efforts have taken place to digitize health records, many healthcare institutions still rely on manual processes as in the Netherlands (such as scanning or retyping physical dossiers). Turkey has made strides in creating a more integrated healthcare system with initiatives like the e-Nabız platform[1], which allows patients to access and manage their health records electronically. However, interoperability between different healthcare providers' EHR systems has not been investigated in depth, and the exchange of EHRs between institutions is not yet fully standardized. This lack of standardized, electronic sharing mechanisms poses significant barriers to achieving a comprehensive and unified view of a patient's health record across different healthcare providers.

It is clear that the current data sharing capabilities in healthcare are insufficient. However, it is not clear how to tackle these issues. There currently exists many barriers to sharing data in healthcare [2]. For one, given the sensitive nature of electronic health records, many organizations are weary about setting up data exchanges. Secondly, different healthcare providers employ different EHR management systems, which are not interoperable.

To tackle these issues, we take an innovative decentralized approach for the use-cases in the Bio-curity project. We envision a decentralized healthcare ecosystem in which the patient has more control over their patient information. By placing the patient central in the IT landscape, numerous benefits can be unlocked. The single point of truth allows a patient to always have an up-to-date record, as well as provide the patient with the ability to share this information with any healthcare institute that requires it. Furthermore, by giving the patient access and control over their health record, the patient can

[1] https://enabiz.gov.tr/

contribute valuable information. This decentralized approach will allow us to implement the data sharing envisioned in the Bio-curity project, stimulating privacy, confidentiality, trust and scalability, which shows a possible solution for the wider IT healthcare landscape.

1.1 Scope

To scope this document, we will specifically deal with the question of how a patient can share the insights generated with the data of their medical wearables. Within the use-case we define two concrete goals for which we would like to share this data:

- **Medical Research**: How do we make sensor insights available for researchers?
    - To allow for research and development as well as real-world evidence collection for novel innovations, researchers wish to be able to collect data, and often patients are willing to help. The challenge is that large amounts of data over multiple patients need to be computed on, and consent needs to be carefully managed.
- **Patient Insight**: How do we make sensor insights available for the patient?
    - In order to provide the patient with tailored insights and advice, the patient should be able to gather insights from all its medical wearables in one place. The challenge is that this data is sensitive and subject to strong data protection, and thus needs to be handled carefully.

This deliverable focuses on the preconditions and criteria required for a trusted data exchange. The criteria required for a trusted data exchange found in deliverable D4.1 are further selected and expanded upon. Furthermore, this document provides a high-level architecture to illustrate the relationships between the components of the data exchange. Note however, that it falls outside the scope of this deliverable to provide the technical details on how the interfaces between the components should be implemented.

2. Methodology

We take two approaches to work through the problem presented in Bio-curity. 1) we derive functional requirements from the project proposal, the use case descriptions and discussions with partners, and 2) we derive relevant criteria from the legal literature.

2.1 Criteria selection

This work follows from our previous work in deliverable 4.1, which dives into the regulatory framework and technical standards (the corpora, composed by 14 documents) to gather security and assurance criteria relevant to healthcare systems, and the processing of health data. This is done through review of requests found in the corpora, and extensive thematic analysis. This results in more than 60 relevant criteria, which are grouped into four main categories: 1) data management and quality; 2) rights and interests of data subjects; 3) information security; and 4) transparency and accountability. Additionally, two supporting categories emerged from the analysis, which do not directly contribute or restrict the implementation of the Bio-curity platform, but that are relevant for its operation once implemented: 5) healthcare and ethics; and 6) automated processing.

Our past work elaborates on the completeness and validity of the criteria list with respect to needs partners defined for the Bio-curity platform. Particularly looking at the risks and mitigation strategies gathered at the beginning of the project (Deliverable 1.2), selecting those risks with a score of “high” or “critical”, and checking that the criteria contributes to the respective mitigation strategies.

Testing the validity of criteria with respect to the desires of partners for the Bio-curity platform is the scope of the current deliverable (4.2). With the understanding that for practical reasons, the complete list had to be refined before designing an appropriate architecture, part of the current work was to prioritize the criteria list. To do so, we devised and executed the following steps:

1. We interpret the use case descriptions provided in Deliverable 2.1 in the light of the criteria list. The use cases place emphasis on *data* and its *accuracy* and *quality*, provision of *information* and explanations, data subject *rights*, and *data transfer*. A list of ten preliminary criteria is refined according to their relevance in this step.
2. As the main goal of the Bio-curity platform is to facilitate qualified data exchange, we select among the criteria contributing to the mitigation of risks (see D4.1 for more details) those which directly relate to data exchange. Here, topics of *access* control and *authorization*, *consent*, and the *identifiability* of data appear to be the most relevant. Another six criteria emerge in this step.
3. For completeness, we cross-reference the top five most occurring criteria which contribute to the mitigation of risks (see D4.1) with our preliminary list emerging from the previous two steps. Three criteria did not appear in our preliminary list yet: security measures, *encryption*, and *confidentiality*. “Security measures” is too high-level and many aspects of it are also present in the remainder of our list, therefore we opt to leave it out the list of more specific criteria. Two new criteria are added in this final step, for 18 pre-selected criteria.

The list of 18 pre-selected criteria is composed of, in alphabetical order: Access control, Authorized access, Compliance to information provision (former "Information to be provided"), Confidentiality, Consent management, Data accuracy, Data anonymization, Data format, Data management, Data

pseudonymization, Data quality, Data quality management, Data transfer (cross border), Encryption, Logging for auditing, Right to restrict access, Right to share data, and Storage limitation. See Deliverable 4.1 for a description of those.

To further refine the pre-selected list, we consider how each criterium can be covered by the Bio-curity platform. In this exercise, we look at how the architecture will address each criterion, whether roles and responsibilities need to be devised to support each criterion (this is the goal of Deliverable 4.3, more details are presented there), and the documentation which is relevant.

What emerges from this exercise is the knowledge that some criteria can be fully addressed by supporting documentation, or by each Bio-curity partner individually. This means they do not contribute to the design of the technology architecture, the goal of our work. For instance, addressing criteria on the topic of data quality, and format seem to require agreements between partners, and the definition of standards accepted by the platform. But from the perspective of the technical architecture, it is sufficient to know data exchanged will be in a commonly recognized format. We do not dive into this layer of abstraction.

Upon further investigation other criteria were deemed out of scope of the project. For instance, “data anonymization” seems desirable at first in a platform where sensitive data such as health data is shared. However, achieving real anonymization requires in depth and continuous analysis, and from a regulatory perspective, anonymizing health data does not exclude the fact that it remains sensitive, and therefore subject to processing restrictions[2]. Instead, we opt to process health data based on explicit consent, which is one of the exceptions for processing sensitive data[3], and to focus on pseudonymous data, which is better aligned with the objectives of the project. Another example is “data transfer (cross border)”; our project is composed by partners from the Netherlands and Turkey, which are subject to different data protection regimes, and do not currently benefit from an adequacy decision [3]. Due to legal complications, we opt as a consortium not to explore cross border data transfer in this project.

Our final list is composed of six criteria which will have a strong influence on the components and protocols which our platform will need to support. Each criterium is inspiration for non-functional requirements presented in Section 3 and is further investigated for current practices and state of the art, presented in Section 6: consent management (6.1), access control (6.2), confidentiality (6.3), right to share data (6.4), encryption (6.5), and data pseudonymization (6.6).

2.2 Requirements gathering and architecture design

The previous section discusses the process of selecting the criteria relevant in the consideration of the architecture. These criteria lay out the framework of the aspect to consider during the design of the architecture. To shape the design of the architecture itself, a requirement engineering process was performed.

[2] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 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) [2016] OJ L 119/1, Art. 9

[3] GDPR, Art. 9(2)(a)

For the requirement engineering processes, the stakeholders are the consortium partners in their various roles. These roles include healthcare institutions, sensor manufacturers, and software vendors. Due to the nature of the project, the requirement engineering process is not straightforward. First and foremost, the output (the architecture) aims to be innovative and push the boundaries of what is currently possible (technology push). Secondly, though all stakeholders are interested in the end-result, no stakeholder alone will have to consider the entire design. Inherent to the assignment, the output is a decentralized architecture describing not a central interface or service, but a set of interactions between different stakeholders.

The two main sources of the requirements are documentation available in the project, and interviews and discussions with the stakeholders. The descriptions of the use-cases in the project proposal and the deliverable "*D2.1 Description of use-case and stakeholder interviews*" were used to derive requirements required to support the use-cases. Secondly, through interactive sessions with sensor manufacturers and software vendors further requirements were developed. In these sessions, discussions were held based on how data currently flows in existing systems and based on different proposals developed to feed the discussion.

The functional requirements derived from this requirement engineering process are found in section 3.1.

3. Requirements

This section lists the functional and non-functional requirements derived using the methodology described in section 2.

3.1 Functional Requirements

The requirements in this section have been gathered specifically for the Bio-curity project. They are the result of the description of the use-case (D2.1), the high-level requirements (D2.2), criteria elicitation and stakeholder interviews conducted by WP4 with consortium partners, and conversations with WP3.

#	Requirement
F1	Raw sensor data may not be shared with the patient
F2	Raw sensor data is processed into processed sensor data by the sensor provider
F3	Sensor insights need to be shared to the patient in real-time
F4	Sensor insights must be stored under control of the patient
F5	Sensor providers should not need to be able to link a sensor to a patient
F6	Sensor providers must verify that patient consent upon sharing data for research
F7	Patients must be able to share their sensor insights with healthcare professionals
F8	Patients must be able to share their health data with third-party service providers
F9	Researchers must be able to discover the data sources
F10	The platform should be interoperable with other healthcare systems

F1 Raw sensor data may not be shared with the patient
Source: Stakeholder interviews
Raw sensor data emitted from a wearable device may be restricted due to issues of business competition and Intellectual Property Rights.

F2 Raw sensor data is processed into processed sensor data by the sensor provider
Source: Collaboration with WP3
The raw sensor data is of little use and needs processing into insights or new biomarkers to be of value to both a patient and a healthcare professional.

F3 Sensor insights need to be shared to the patient in real-time
Source: D2.1 Description of use-case, High-Level Requirement DPS_001
In the context of the Dutch use-case, the ultimate goal is to provide real-time feedback on behavior and lifestyle to help the patient combat the causes of the MetS syndrome. Though the data stream need not have a high sampling frequency, it should be an uninterrupted data stream.

F4 Sensor insights must be stored under control of the patient
Source: High-Level Requirement NFSSP_001
The sensor insights are health data about the patient, and should thus be stored under control of the patient.

F5 Sensor providers should not need to be able to link a sensor to a patient
Source: High-Level Requirement NFSSP_002

The sensor provider does not need insight into a patient's sensor data, and as such does not need to be able to link a sensor to a patient.

F6 Sensor providers must verify that patient consent upon sharing data for research
Source: Law & Regulation
If a sensor provider receives a request to share sensor data about a patient, the sensor provider must verify that the patient consents to its data being shared.

F7 Patients must be able to share their sensor insights with healthcare professionals
Source: D2.1 Description of use-case
The prevention, detection and alleviation of MetS symptoms is always done together with a healthcare professional. They require access to the sensor data or its insights to help the patient.

F8 Patients must be able to share their health data with third-party service providers
Source: D2.1 Description of use-case
One of the goals of the Bio-curity project is to track the development of the metabolic syndrome across the population and use this information for decision making. To support this, the patient's sensor insights must be able to be shared with such a platform.

F9 Researchers must be able to discover the data sources
Source: D2.1 Description of use-case and stakeholder interviews
In order to support medical research, researchers must be able to find the data sources on offer.

F10 The platform should be interoperable with other healthcare systems
Source: D2.1 Description of use-case and stakeholder interviews
An important general requirement is that the platform should be interoperable with other healthcare systems. The lack of interoperability is the main technical obstacle in integrating the current healthcare systems.

3.2 Non-Functional Requirement

Non-functional requirements are derived from the selection of criteria.

#	Requirement
NF11	Platform must support a consent model which can represent valid consent according to current regulations and standards.
NF12	Platform must support an access control model that adequately protects resources, according to regulatory demands and technical standards.
NF13	Platform must comply with regulatory and standard demands for confidentiality.
NF14	Platform must support the right to share data, according to current regulations and standards.
NF15	Platform must support encryption protocols where necessary, that adequately protect resources, according to regulatory demands and technical standards.
NF16	Platform must support data pseudonymization when necessary, according to regulatory demands and technical standards.

NF11 Platform must support a consent model which can represent valid consent according to current regulations and standards.
Source: Criteria 'consent management' (D4.1)
Platform supports consent management by: establishing and documenting a process by which it can demonstrate whether, when and how consent to process data has been obtained, requesting (where appropriate, in written form and dated) and keeping record of explicit permission to process data for a specific purpose, and allowing easy withdraw of consent at any time. Additionally, supporting information of expected consequences and risks to the health of the patient in case proposed examination, treatment or procedures are not performed.

NF12 Platform must support an access control model that adequately protects resources, according to regulatory demands and technical standards.
Source: Criteria 'access control' (D4.1)
Platform adopts an access control system that allows granular rules to represent different categories of electronic health record required by different health professional, restriction of data access to authorized parties, and overrule of restrictions to protect vital interests of people.

NF13 Platform must comply with regulatory and standard demands for confidentiality.
Source: Criteria 'Confidentiality' (D4.1)
Platform ensures confidentiality (prevent sensitive information from unauthorized access attempts) of data by: ensuring no health information about a patient is provided to or accessed by others without patient's consent; data exchange is secured along the entire path between sender and receiver; data is encrypted prior to exchanging; protecting against unauthorized access or unlawful processing; protecting commercially confidential information, trade secrets, and intellectual property rights.

NF14 Platform must support the right to share data, according to current regulations and standards.
Source: Criteria 'Right to share data' (D4.1)
Platform supports transmitting data to a recipient of choice and sharing data with a third party.

NF15 Platform must support encryption protocols where necessary, that adequately protect resources, according to regulatory demands and technical standards.
Source: Criteria 'Encryption' (D4.1)
Encryption is used to safeguard rights and freedoms of natural persons where anonymization may significantly affect the purpose pursued.

NF16 Platform must support data pseudonymization when necessary, according to regulatory demands and technical standards.
Source: Criteria 'Data pseudonymization' (D4.1)
Pseudonymization (de-identification procedure by which personally identifiable information fields within a data record are replaced by one or more artificial identifiers) is used to safeguard rights and freedoms of natural persons where anonymization may significantly affect the purpose pursued. The information necessary to reverse pseudonymization is only available to authorized parties, and failure to respect pseudonymization is subject to penalties.

4. Patient-centric and Decentralized Approaches

In this section, we consider and survey innovative approaches that could help support a decentralized and patient-centric ecosystem. We will review a variety of approaches and discuss their applicability.

Blockchain

Blockchain is a technology that offers a *decentralized, immutable ledger*. At the core of Blockchain, is the so-called *consensus* mechanism, allowing a large set of peers to reach consensus on the state of the ledger. With this decentralized consensus mechanism, there is no need for a central authority, which is the key innovation of the technology. The ledger offers an auditable immutable history of all logs entered on it.

Since the rise in popularity of blockchain technologies, much academic research has been done to develop proofs of concepts for its use in the medical domain [4]. The decentralized approach of blockchain technologies together with its strong auditability guarantees appears to be an interesting match for the fractured ecosystem that is the IT landscape in healthcare.

Initial proposals dating back to 2016 envisioned storing entire Electronic Healthcare Records on the blockchain, but this has some obvious drawbacks [5]. Most problematically from a technical point of view, blockchain solutions are not nearly scalable enough currently to support this. Furthermore, concerns can be raised concerning privacy, data loss (due to human error), and the flexibility of such solutions.

Instead, many proposals focus on employing the blockchain for data validation, auditing and authorization when dealing with electronic health records [5]. In these works, electronic health records are generally stored off-chain, with pointers and hashes stored on the blockchain to ensure data integrity and auditability [6]. Most focus in literature is on employing blockchains for either decentralizing Electronic Health Records (EHRs) or decentralizing Personal Health Records (PHRs), though other uses have been explored, such as data sharing in research or managing decentralized patient trails [7].

By decentralizing electronic health records, and giving (partial) control to the patient, the health data ecosystem becomes patient-centric [8], as opposed to the current institution-centered and siloed approach. Blockchain could support digital access rules on patient data, where patients decide how and when data is shared. By having a centralized overview of a patient's medical record on a blockchain, the data becomes available and findable for the ecosystem. The medical record of patient represented on the blockchain also provides a single patient identity, that other institutions can leverage for interoperability.

Self-Sovereign Identity

Self-Sovereign Identity is a concept originating from the larger identity community. With the introduction of blockchain technology and its rise in popularity, the identity community saw the opportunity to leverage the technology to achieve true online identity.

The goal of self-sovereign identity is to provide users with a digital identity that they own and control themselves. More concretely, self-sovereign identity should enable digital credentials that function similarly to existing physical credentials, like your driver's license. When a government agency has

issued you a driver's license, you can show this license to a car rental company to prove ownership of said driver's license, without needing the government agency to be involved in this transaction. This so-called trust triangle, between the Holder, Issuer, and Verifier is at the core of self-sovereign identity.

Most Self-Sovereign Identity solutions use blockchain as an underlying technology, and as such, the technologies have overlapping benefits. SSI provides a decentralized and interoperable network for exchanging credentials, which could be employed for patient-centric interoperability in healthcare. The technology could cross the gap between siloed health data spread over healthcare institutions and provide patients with a method to gather and share their health data. The extra benefit of SSI is its *verifiable* credentials, which could help bring trust in the healthcare ecosystems. Healthcare providers are weary to share data, as improper consent management could result in legal consequences.

Persoonlijke Gezondheidsomgeving

The Persoonlijke Gezondheidsomgeving (PGO), translated the personal health environment, is an initiative from the Dutch government to provide patients with access to their electronic health records. The idea of personal health environments is not new, but the effort is interesting in the Dutch domain as the initiative is supported by government and the legislative process.

As is true in many countries, medical information of patients in the Netherlands is scattered in silos at multiple institutions. If a patient seeks help at a new healthcare provider, its medical history often needs to be copied over by hand [9] or tests are preformed again to get the same data.

On the 1st of July, 2020 [1], the Dutch government enacted a law that gives patients the right to view their Electronic Health Record. This ability must be supported by healthcare providers either by 1) providing a downloaded pdf, 2) providing access through an online portal, 3) or allowing the EHR to be downloaded into a PGO.

The information exchange in the Dutch healthcare system is being standardized by Nictiz, and interoperability between PGOs is organized via the MedMij association. Currently, PGOs are mostly one-way traffic, allowing patients to transfer their scattered electronic health record from multiple sources into their PGO. The ambition is that patients in the future will be able to add their own data, such as sensor data or questionnaires, to provide back to healthcare providers.

Discussion

Despite the promises of blockchain technology and it's decentralized approach, the technology may not be the best fit for the healthcare domain. Specifically, privacy is of concern with blockchain technologies. A blockchain used as a distributed ledger is by definition readable to all participants. Whether actual health information is stored on the blockchain, or only authorization and access logs, both options bring privacy concerns. Furthermore, blockchains are known to have scalability issues, which would need to be carefully considered.

The PGO acts as a patient-controlled store of health information about a single patient and provides a healthcare specific interface to visualize and interact with this data. The PGO integrates through MedMij with the existing healthcare ecosystem to obtain its patient information. However, the PGO currently has no ability to share data itself with other parties. Furthermore, the current data sharing

capabilities offered through MedMij are rigid, and only participants in MedMij may participate in data sharing.

Self-Sovereign Identity provides a decentralized identity and a flexible way of sharing verifiable identity information through the use of verifiable credentials. This patient-centric approach could bridge the gap in interoperability by placing the patient itself in the center of data exchanges about said patient. However, current Self-Sovereign Identity implementations are not well suited to store arbitrary information such as sensor data, nor are they designed to implement interfaces specifically for healthcare.

In short, these approaches do not provide a drop-in solution for data sharing in healthcare, but they have interesting characteristics that support a radically different patient-centric approach.

5. Architecture

In this section we will lay out the architecture to orchestrate data exchanges for sharing data and insights from wearable medical sensors. As discussed in section 1.1, the main goal of the architecture is to facilitate 1) research with insights from medical wearables, and 2) provide the patient with insights into their own health. The purpose of this architecture is to illustrate the relationships between the various partners in the consortium and how the trust required for data exchanges can be implemented through identity management, consent management, and access control.

First, we will discuss the foundational concepts that will be incorporated into the architecture in section 5.1, followed by an overview of the architecture and its components in section 5.2. In section 5.3, the functional processes supported by the architecture are detailed to show the interactions between the components.

5.1 Technology

The architecture described in this document is a high-level architecture and will not dive into details of any specific implementation of technology. However, we will take the concepts of Section **Error! R eference source not found.** and apply them.

Self-Sovereign Identity

The concept of Self-Sovereign Identity gives us a foundation to decentralize identity management and access control in the architecture. As we are handling sensitive data, it is crucial that data is only handled by those who are authorized to handle it. However, as we are organizing data exchanges across domains (different organizations and institutions) that each have their own internal access control, checking whether a user is authorized to perform an action is not trivial. Self-Sovereign Identity provides a decentralized approach to identity management and access control.

To abstract the concept of self-sovereign identity for the architecture, participants in the system will be said to have an *agent* to manage and use their identities, which acts on its user's behalf. We assume for simplicity that a patient uses a single decentralized identifier (DID) to uniquely identify itself in the ecosystem.

Personal Health Environment

Traditionally, a person within an SSI ecosystem has a *wallet* application. Such a wallet manages decentralized identifiers (DIDs), provides storage for verifiable credentials and is the agent that acts on the user's behalf in identity-related transactions, such as receiving verifiable credentials or presenting verifiable presentations. These functionalities, though crucial for SSI, do not cover all that is required for a patient-facing application in healthcare. An application managing health data for a patient also needs to be able to store health information. Though it is possible to store any data as a verifiable credential, patients and applications may benefit more from storing and sending health data in health-specific data formats focused on semantic interoperability. Furthermore, such an application would require a domain specific interface, to present the information to the patient in a meaningful way. In the architecture we will consider an abstract role called the personal health environment (PHE), which implements an SSI wallet application, is able to storage health information, and provides a healthcare specific interface.

5.2 Overview

In Figure 1, an overview of the components of the architecture is depicted. The parties relevant to the architecture are:

1. The Sensor Provider, who manufactures the medical wearable that patients use and processes their data.
2. The Personal Health Environment Provider, that implements and fulfils the personal health environment role for the patient.
3. The Orchestrator, that facilitates setting up data exchanges (process 3 and 4 in the diagram).
4. The Healthcare Institution employing Researchers that wish to use data for medical research (process 3 in the diagram).
5. The Patient, who wears a medical wearable (process 1 in the diagram) and has a personal health environment and wishes to retrieve insights (process 2 in the diagram) and participate in medical research (process 4 in the diagram).

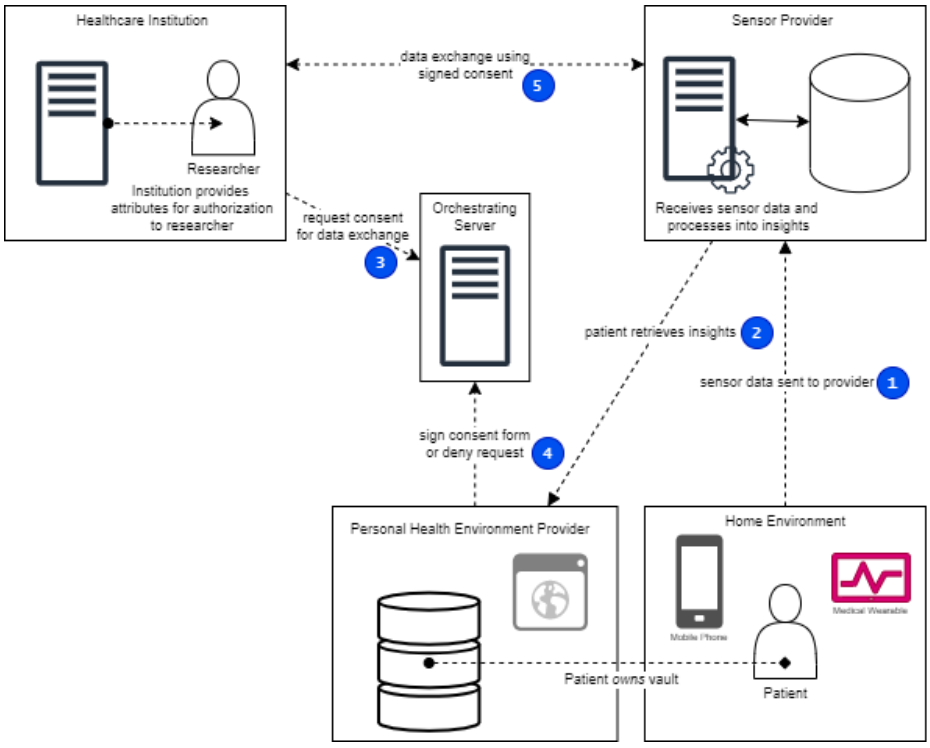

Figure 1 Overview of the Architecture

Orchestrating Server

At the center of the architecture is the orchestrating server, which is the only "tailormade" component of the architecture specifically for the Bio-curity project. The orchestrating server, as the name suggests, helps in orchestrating the data exchanges. Specifically, the server will be the platform through which data providers can advertise their offerings, and for researchers to find datasets and patients willing to participate in research.

The orchestrating server will not store or process any personal or medical information of patients. Creating such a centralized database would bring great privacy, scalability, trust and security risks. Furthermore, this would pose a large burden to the party tasked with running the orchestrator, considering the legal implications of storing and processing of medical data.

Do note that even without storing any personal or medical data, the party deploying the orchestrator still has a sensitive role in organizing the ecosystem and can infer information about patients based on the data exchanges it sees. For example, the orchestrator can infer that a patient participating in medical research into diabetes might have diabetes.

Personal Health Environment

In a patient-centric architecture, the patient will need a place to store their health data and an interface to manage it. In the architecture, we will abstractly describe this role as the *personal health environment, as discussed in section 5.1*.

The *personal health environment* should hold (part of) the Electronic Health Record of the patient for the patient to be able to participate in research, as this data is used to determine whether a patient is eligible to participate. Therefore, the *personal health environment* is assumed to integrate with the healthcare systems of medical institutions and retrieve the patient's data[4].

Sensor provider

The sensor provider is the role associated with the manufacturer of a medical wearable device. The sensor provider also processes the raw data produced by the device into insights, either on the device itself or through their infrastructure. It is assumed that the insights produced from the raw data by the sensor provider may be shared with the patient (and further shared with their healthcare professionals).

The architecture supports an arbitrary number of sensor providers, but to ease with exposition, the architecture discusses a single party.

Researcher

The Researcher must be employed through a medical institute or research institute that participates in the Bio-curity consortium. The researcher needs an appropriate Researcher credential issued by their institute for them to authorize themselves in the system. How such a credential could look is discussed in section 6.2. Rules should be laid out in the governance process to determine when an institute can authorize a researcher.

5.3 Processes

This section dives into the details of the functional processes supported by the architecture. The processes are ordered in chronological order where relevant and further explain the architecture's details and design decision.

[4] In the Netherlands, this functionality is implemented in the PGO through the MedMij network.

Patient: Register wearable with sensor provider

From a patient point of view, the first interaction required is registering themselves and their wearable with the sensor provider. We assume that the patient has readily received the wearable device, for example during a doctor's visit.

The goal of the registration process is to link the patient to the device it has received. For the patient to be able to receive insights from the sensor provider, the patient must be able to prove physical ownership over the device to the sensor provider, to ensure that only authorized users (including the patient themselves) may request the insights.

To this end, the patient confirms the device identifier to the sensor provider and presents their patient identifier. The process has been laid out in a sequence diagram in Figure 2.

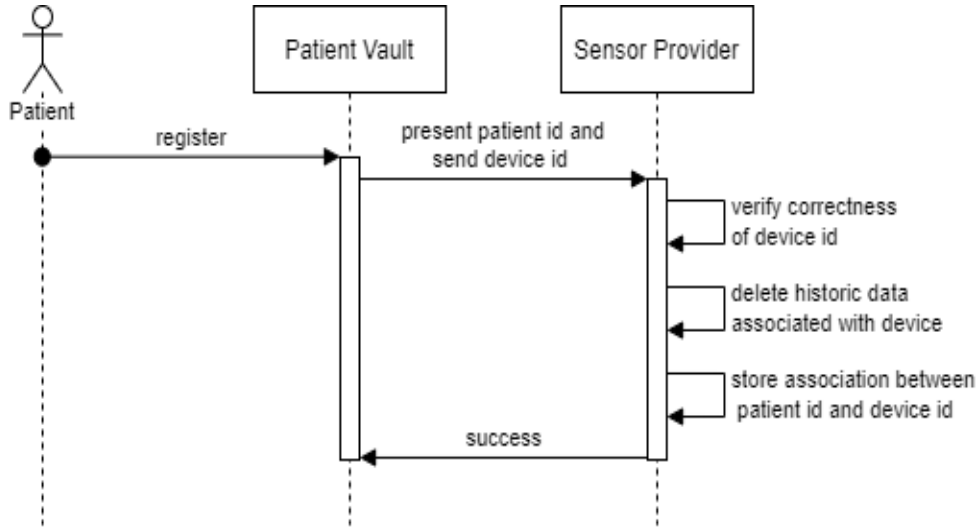

Figure 2 Sequence diagram of the registration process

Note that the device identifier is likely static and not secret, and thus anyone knowing the device identifier could register the device. To secure this process and ensure only those in physical possession of the device may register it, a passcode could be stuck onto the device that must be provided by the patient. This allows the sensor provider to reuse the same device, as it can assign a new passcode to the device upon return.

Patient: access medical wearable insights

Having established the registration process above, we can support the use-case of a patient retrieving insights derived by the sensor provider from the sensor data, as depicted in Figure 3.

The figure gives a rather simplified view of the process, focusing solely on access control. The process assumes that the *personal health environment* offers a user interface to initiate the importing of insights. Furthermore, per the functional requirements (F3), the insights from the medical wearable sensors are provided to the patient in real-time. Though the access control in the figure above does not limit this requirement, transferring data constantly and in real-time will require more complex software interfaces between the *personal health environment* and sensor provider components, not detailed in this deliverable.

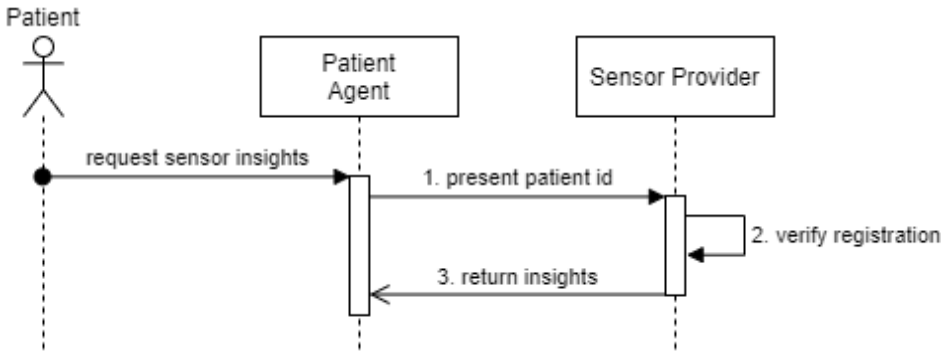

Figure 3 Sequence diagram of process for requesting insights

Researcher: request approval for research

The second main goal (see 1.1 Scope) is to make patient data available for researchers. To support this, we need to verify several preconditions. Though a patient may consent to their data being used for medical research, we need to verify that this consent is not abused. In this sequence, we verify that the researcher is authorized within the Bio-curity consortium to perform research, and verify that the research proposed is lawful and ethical through a governance process involving an Ethical Board.

Authorization

We do not allow any random person claiming to be a researcher to access patient information. Only personnel designated as researchers by an institution onboarded to the Bio-curity consortium may perform research requests. We depend on Self-Sovereign Identity to simplify the implementation of access control. The researcher is issued a Researcher credential by their own institute stating their authorization related attributes (such as, “is employed at institute X” and “is researcher”). These credentials can easily be verified by the orchestrator by checking that the signature on these credentials is valid and from a known and onboarded institute (step 2 in diagram). What the contents of such a Researcher credential could contain and when the credential is sufficient for authorization for what types of research, is discussed in section 6.2.

Research Definition

The definition of the research has both a functional and a governance purpose. First, the research definition should specify the type of data the researcher wants to access, and from which patient group. A research may for example be interested in the heart rate of patients diagnosed with obesity. The types of data the researcher needs and the query on which to select patients need to be specified in a standardized fashion such that the data can be retrieved once the research is approved.

Ethical *Board*

The second purpose of specifying both data required, and the patient group is for the governance process involving the ethical board. An Ethical Board needs to be assigned by the Bio-curity consortium to judge whether a particular research request is lawful and ethical. Whether this process should be manual of (partially) automated must be decided by the consortium and is left out-of-scope for this deliverable. However, certain guardrails should likely be put into place, such as:

- Ensuring that the patient set resulting from the query has a minimum size (to prevent a researcher collecting data on a single or handful of patients).
- Limit the amount of research requests one may submit to prevent abuse.

Once the Ethical Board has approved the research definition, a ResearchApproval credential is issued by the orchestrator to the researcher (step 5).

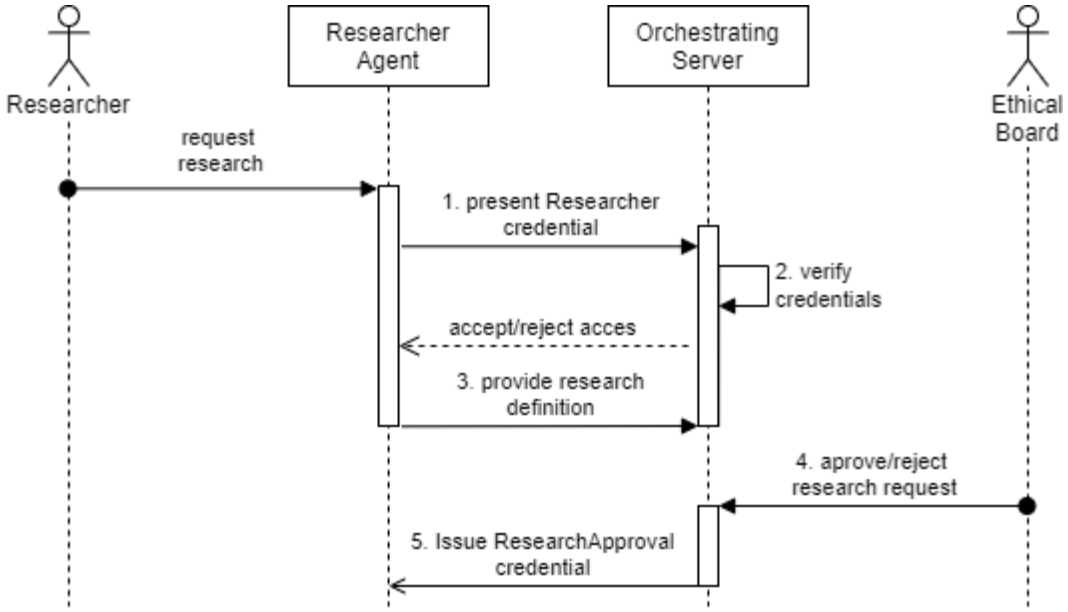

Figure 4 Sequence diagram of the research approval process

Researcher: request consent from candidates for research

The researcher has now defined their research, which has been approved by the Ethical Board. The researcher can now initiate the process of gathering consent from patients that would be eligible to participate in the research. The researcher has already specified the patient group the researcher is interested in, e.g. patients diagnosed with diabetes. The sequence diagram below shows how the orchestrator finds research candidates and asks for consent.

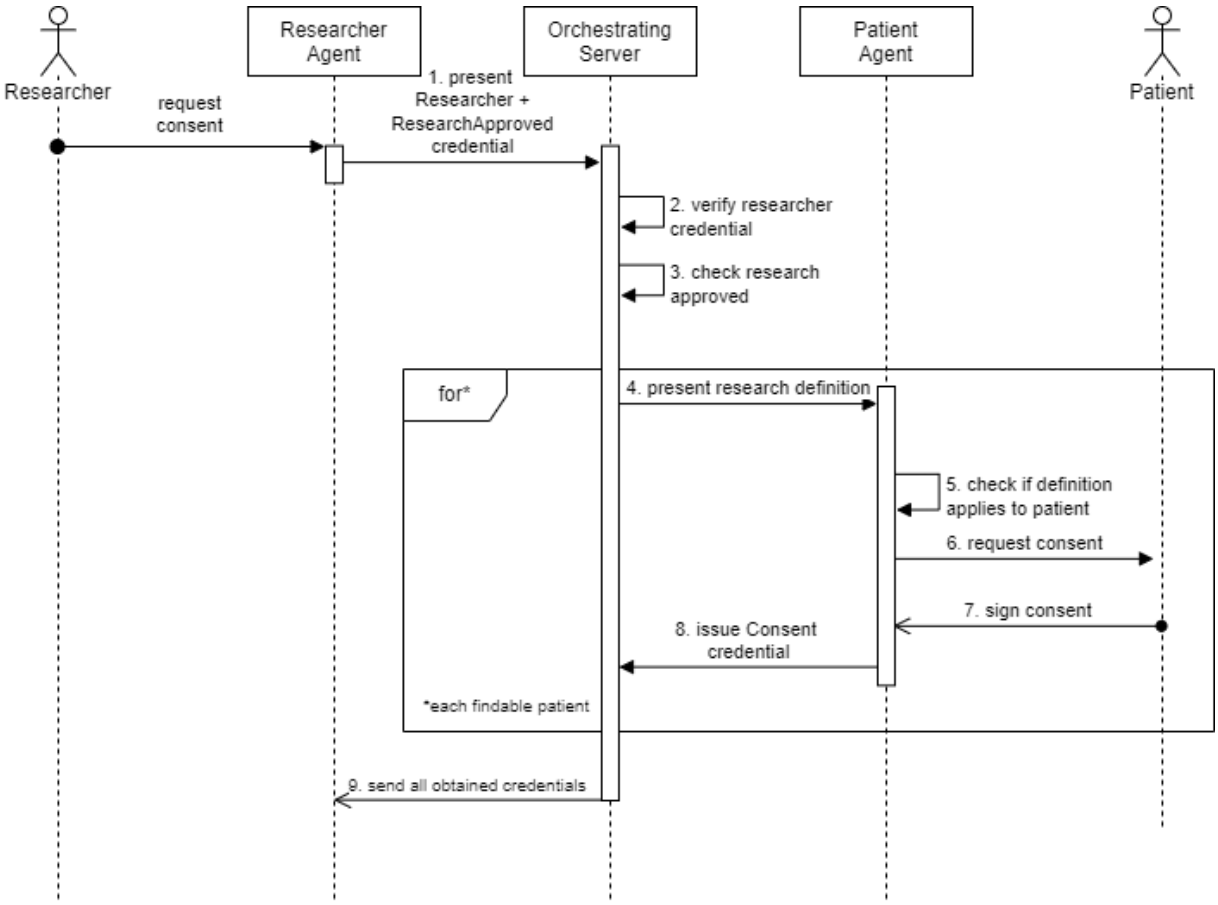

Figure 5 Sequence diagram of request consent process

In step 1, the researcher authorizes itself to the orchestrator by providing its identity, the Researcher credential and its ResearchApproved credential. The orchestrator verifies the identity of the researcher (step 2) and that the research definition has been approved by the ethical board, by verifying the ResearchApproved credential (step 3). If all is in order, the orchestrator starts requesting patients to participate in the research. For each patient registered to the orchestrator, the orchestrator requests consent for the research. The personal health environment automatically determines whether the patient falls in the target group of the research (step 5). If so, the personal health environment can either automatically or by instruction of the patient issue a Consent credential to the researcher (indirectly via the orchestrating server) (step 6-8). If the patient does not fall in the target group, the personal health environment should ignore the invitation. The Consent credential must state to which research the consent applies. Once all eligible patients have responded, or after a certain time, the

researcher collects all Consent credentials (step 9). These credentials provide the researcher with the authorization to request the data of interest at the sensor provider.

Privacy

The major obstacle to overcome is that, by our requirements, we do not wish to store the medical information of all patients interested in potentially joining a study. This impacts the architecture, as this means the orchestrator alone cannot determine which patients would be eligible for the research, as the orchestrator does not have access to a database with medical information on all patients. Having such a database is not desirable either, as it poses a high security and privacy risk to the patients stored within, given that such a database of information would be a highly valuable target to adversaries.

Instead, we choose to let the personal health environment check (step 5) whether the patient is eligible to participate in the research. The personal health environment has access to the Electronic Health Record of its patient, and thus can determine whether the research definition applies to the patient, provided that the research definition is properly standardized.

Without further assurances, this would allow a malicious participant to erroneously claim to fall within scope of the research definition. This could be prevented by leveraging the verifiable claims concept of Self-Sovereign Identity to the Electronic Health Record; A medical professional could sign the Electronic Health Record to provide assurance that the EHR is legitimate.

Consent

If the personal health environment decides that the research definition applies to the patient, the patient may choose to consent or not. How exactly this process looks like depends on the consent model, which is discussed in section 6.1. In short, the patient decides to consent on an institute and/or purpose level, and thus is not required to consent to each individual research request. Once a preference has been set by the patient to the personal health environment, the agent can match the research definition against the user-specified preference and make an automated decision.

Findability

Part of the question of consent is whether a patient is findable by the orchestrator in the first place. This is easily solved by introducing a registration step for the patient at the Orchestrator. Not only does a registration process provide a good opportunity to educate the patient about the project and let the person indicate whether and for which research he or she wishes to be “findable”, it also is required technically for the Orchestrator to be able to locate a findable patient.

Revoking consent

At some point, a patient may decide that a previously made decision regarding consent should be changed. The patient can update their consent preferences in the personal health environment to change when consent is signed for new research requests. Previously signed consent credentials can be revoked, such that verification of the consent credentials will fail if they are used. The implementation of the revocation mechanism depends on the implementation of self-sovereign identity.

Researcher: use consent to retrieve data from sensor provider

Once a researcher has received the Consent credentials from the participating patients, the researcher is ready to perform their research. When it comes to the actual data sharing process, there are two options:

1. The researcher retrieves the data of interest and performs analysis locally.
2. The analysis is performed using Privacy Enhancing Technologies and only the analysis result is shared with the researcher.

The first option is the simplest and is depicted in the sequence diagram of Figure 6. However, this will mean that the researcher retrieves all data from the provider, which may not be desirable depending on the use-case and/or the sensitivity of the data. The second option would be to employ Privacy Enhancing Technologies, such as confidential computing, or federated learning and multi-party computation (in case of multiple data providers), to only share the results of an analysis with the researcher.

The use of Privacy Enhancing Technologies to obtain strong privacy guarantees is promising, but out-of-scope for this deliverable. In this section we focus on the sequence diagram for option 1.

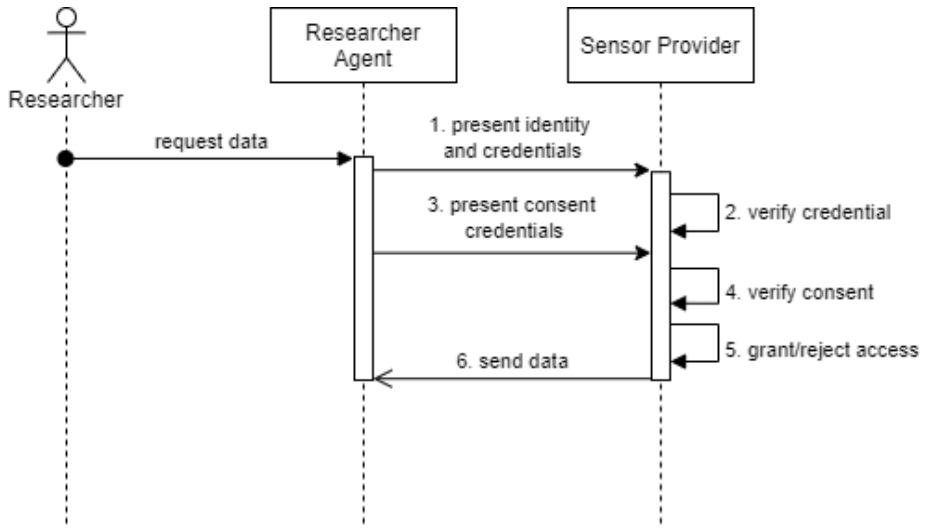

Figure 6 Sequence diagram of process for retrieving data for research

Authentication and Authorization

As always, the researcher needs to be properly authenticated and authorized before receiving any data. As the diagram shows, the orchestrator is not required directly in this interaction, as self-sovereign identity allows the sensor provider to independently verify the correctness of the credentials provided by the researcher.

The researcher needs to prove various statements to the sensor provider (step 1, step 2):

1. The Researcher has a valid Researcher credential, signed by an onboarded institution.
2. The ResearchApproved credential is valid and matches the research definition.
3. The research definition was submitted by the researcher.

If any of these statements do not hold, the sensor provider should reject the transaction. Next, the researcher provides the consent credentials they obtained via the orchestrator from the patients who consented to participate in the research (step 3). Here, again various statements need to be proven (step 4):

1. The consent credentials are valid.
2. The consent credentials provide consent for the purpose covered by the research definition
3. The consent credentials provide consent for the institutions and researchers involved in the research definition.

After this step, all the necessary preconditions have been met, and it is up to the sensor provider to grant or reject access to the data. Whether the sensor provider should be allowed to reject access for approved research, and on what grounds, must be set out in the governance of the consortium.

Patient: provide Home Monitoring access to medical professional

Figure 7 depicts the processes of a patient providing access to the insights of their medical wearable to a medical professional. The process is a slightly simpler version of the process "Researcher: use consent to retrieve data from sensor provider". In this process, the patient is assumed to have direct contact with the medical professional, which simplifies the access control and makes the consent more direct.

The medical professional requests insights (processed results from sensor data) from the patient. To provide this access, the personal health environment issues a Consent credential (step 2).

Using the Consent credential, the agent of the medical professional can authorize at the sensor provider and retrieve the data (step 3-5).

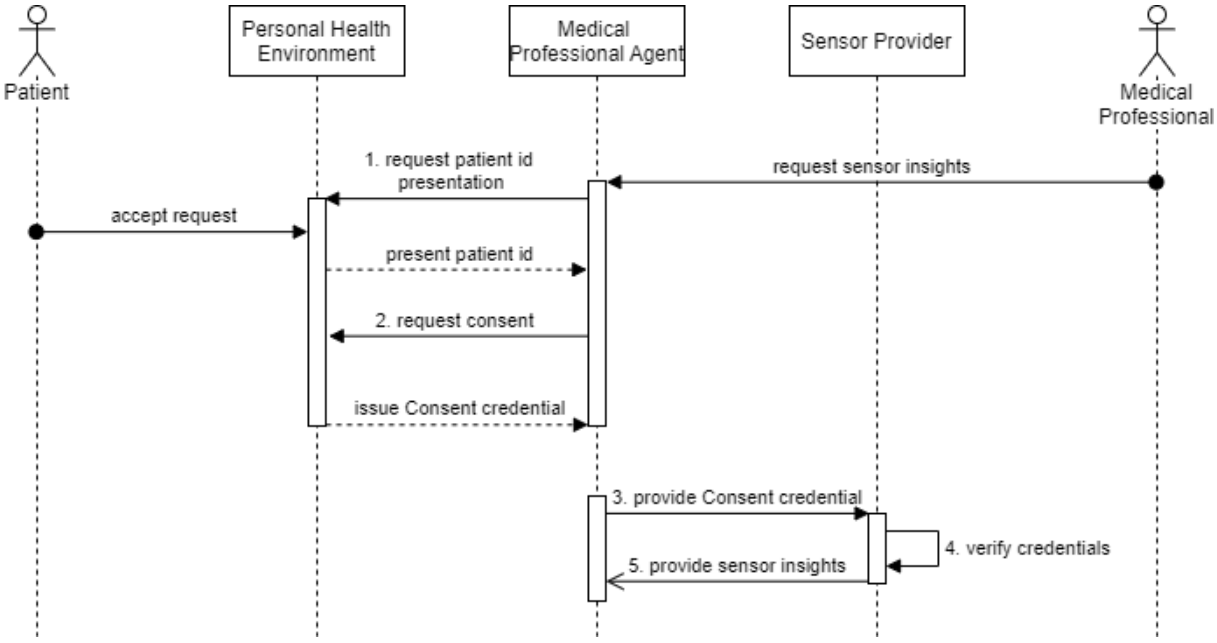

Figure 7 Sequence diagram of a patient sharing sensor data with a medical professional

Real-world Example

To picture how this process might look in a real-world setting, imagine a patient visiting a doctor during a regular check-up. The patient has a wearable device sending sensor data to the sensor provider. The doctor asks the patient for access to this data and presents the patient with a QR-code to initiate a connection. The patient scans the code and agrees to issue consent to the doctor. The doctor can then use this consent to request the data from the sensor provider.

6. Criteria

6.1 Consent management

Consent is a central component to several regulations and standards. In general, there are three scenarios where consent appears relevant:

1. Primary use of data (medical treatment). This is consent collected during a consultation with health practitioners for a specific treatment or administration of medicaments. Consent of this type can take the format of a conversation.
2. Clinical trial. This is normally collected during interviews or other types of interactions with trial participants. This type of consent request may be presented in many formats, including in hard copy.
3. Data processing. This is consent which allows the processing of data in its most comprehensive format, that is, collection, visualization, filtering, analyzing, transforming, sharing, and so on. This type of consent is normally digital, in written format.

All three scenarios present situations where consent is equally important, and mandatory by regulations, but only data processing (#3) is relevant to the design of a platform for qualified data exchange. This is the type of consent we explore in this section.

Despite being a central component to several regulations, consent becomes particularly relevant for the Bio-curity platform considering GDPR's demands. That is because under the GDPR, health data is considered sensitive, and the processing of such data is prohibited in general. A few exceptions apply, such as processing for the establishment, exercise or defense of legal claims in courts[5]; when there is substantial public interest in the area of public health[6]; or the data subject has given explicit consent for such processing[7]. Consent appears to be the only applicable exception for the legal processing of health data in Bio-curity.

What emerges from legal requirements is that consent for data processing must adhere to the following to be considered valid:

1. Explicit consent must be given for one or more specific purposes;
2. If the data subject has a fiduciary (guardian, legal representative) the rights accrued to the data subject are exercised by the fiduciary, including consent;
3. Controller must be able to demonstrate consent was given;
4. Controller should establish and document a process by which it can demonstrate whether and how consent has been obtained;
5. Data subject has the right to withdraw at any time;
6. Before giving consent, the data subject is informed of the right to withdraw at any time; and
7. Withdrawing consent is as easy as giving it.

[5] GDPR, Art. 9(2)(f)
[6] GDPR, Art. 9(2)(i)
[7] GDPR, Art. 9(2)(a)

Current consent models

In the Netherlands, MedMij is one of the most relevant solutions in health data exchange, providing standards for the secure and reliable exchange of such data between care users and care providers[8]. MedMij offers frameworks which can be implemented by applications such as personal health environment (PHE) that can get certified with the "MedMij label" when meeting certain criteria. Relevant for the topic of consent management is the scheme of agreements provided by MedMij's framework, under which the declaration of consent is described [10]. Consent is requested of the data subject for sharing data for an indefinite amount of time, with the condition that consent is invalidated if the data subject becomes inactive for six consecutive months or longer. Particular to PHEs, consent allows a care provider to add data to the data subject's PHE and does not include further sharing without further permission. Consent can be withdrawn at any moment.

Also relevant to the topic is Mitz, a platform for registering and managing decentralized consent, serving as connector between multiple patients and health care providers[9]. There are several components of Mitz, we focus on the consent catalogue, where all consent options are described [11]. Consent is based on layered disclosure, with 16 situations where consent can be requested to the data subject. For instance, consent can be given so that a specific health data controller (such as a hospital) to make personal data of the data subject available to pharmacies. In this situation, all institutions under the category of "pharmacy" would have access to the data in question. Consent can be given with respect to one specific institution, or a category of those, and with the choice of type of data that can be shared, making the consent model somewhat granular. Consent can be given for 72 hours, or indefinitely, until modified by the data subject. Consent can be withdrawn or modified at any moment.

Finally, PROVES's proof of concept for a health data vault also explores consent models. PROVES is a service that organizes proofs of concepts and pilots to test healthcare innovations, with the goal of testing, in a controlled setting, the functionalities and technical aspects of new solutions[10]. Their proposal is for individual data vaults which are linked to each data subject, to safely store their personal and health data, collecting from multiple data sources, so that data becomes user-centered and more easily controlled by the individual [12]. In here, a consent model is explored to connect the data vault with a PHE to share a specific data category, with choice of consent and time limitations for each category. For example, "medications" can be authorized to be shared for an indefinite amount of time with a general practitioner, while "activities and diagnosis" can be time-bound, or not authorized at all. Allowing for consent choices to be expressed in a fine-grained manner. Consent can be modified and withdrawn at any moment.

A summary of observations about the discussed consent models can be found in Table 1. Note, while this list is not comprehensive, we deem the found solutions sufficient for the sake of discussing potential consent models suitable for the Bio-curity platform, as they cover three levels of granularity. It is also important to note that no current consent model covers consent for non-health care providers, such as sensor providers, or research institutes. We do not see this as a limitation of the current models, as they have the specific purpose of contributing to primary use of data, which is the

[8] About MedMij - Home
[9] Mitz, the online consent facility (mitz-toestemming.nl)
[10] PROVES | VZVZ

provision of health care to individuals. Our platform aims to expand to secondary use of data (such as further research), and therefore one the basic design of each consent model is considered. We add to the table our considerations with respect to envisioned limitation within the Bio-curity platform.

Table 1 - Summary of observations on current consent models

	Observations	Limitations for Bio-curity	Notes
MedMij	• Low granularity of consent, requires less user interaction	• Arguably not sufficient to convey the complexity of data processing in Bio-curity	Consent is for an indefinite amount of time, unless user is inactive for six consecutive months
Mitz	• Medium granularity of consent (category/ individual data provider) • Easy onboarding of new partners under an existing category	• Requires structured process to onboard partners in new categories • No granularity for data	Consent can be for an indefinite amount of time, or 72h
PROVES	• High granularity of consent (one by one) • Medium granularity of data • Period limitation	• Only theoretical, little documentation • Granularity of consent requires more user interaction	

Purpose for processing

When discussing consent, defining specific purposes for processing data appears with emphasis in the regulatory requests (see legal requirement 1). All three consent models discussed previously consider primary use of data, with the purpose of sharing data for providing health care to patients. This is arguably not sufficient for the Bio-curity platform, as our goal is to expand to secondary use of data by, among others, allowing scientific research, providing patient insights, and training machine learning models. In order to further specify purposes, members of the project brainstormed all possible purposes, which are combined into the following categories:

1. Clinical use – data is used for primary care
2. Research – data is used for clinical research & development, among others
3. Public Health Governance and Policy – data is used to aid creation of governmental policies related to public health, disease and outbreak surveillance, or epidemiological studies
4. Commercial and Development – data is used to generate real world evidence to support the development of new products and services
5. Findability – (meta-)data is used in a catalogue to improve findability

Bio-curity consent model

In order to decide which consent model is the most suitable for our platform, a consultation with project members and relevant stakeholders was arranged. The three consent models, the observations, and potential limitations were presented on this occasion, and the most interesting aspects of the models were selected to be incorporated into the Bio-curity consent model. Those are: medium granularity of consent from Mitz (option to choose from category or individual data provider); and medium granularity of data types and period limitation from PROVES.

During this consultation another relevant source emerged, on a cooperative for management of health data (in original Dutch: *Gezond Akkoord*) [13]. This source is used for validation of our proposed consent model. Two findings emerge: the two consent models (Gezond Akkoord, and Bio-curity) seem compatible; Gezond Akkoord works with the concept of an ethical approval, which happens before a consent request is sent to data subjects. This gives stronger assurances of data minimization and fair processing of data. Similar to the ethical review that is required for human research, such ethical approval of data requests would also require the creation and maintenance of an internal ethical board. Nevertheless, due to the high sensitivity of data involved in the Bio-curity platform, the consortium opted to incorporate such concept into the consent model.

Another suggestion emerging from this consultation is that in order to provide data subjects with a valid consent choice, they should be well informed about the consequences of their participation in a new data processing activity, including when relevant the level of engagement required from them, and what data subjects get in return for their participation.

The consolidated consent model is depicted in modeling notation in Figure 8. In a more accessible notation, the Bio-curity consent model can be interpreted as following:

> Consent is provided by a data subject (represented by their decentralized identifier (DID) an identifier of the PHE belonging to a data subject – following privacy design strategies of data minimization and separation [14]) to authorize a (individual or category of) Data Holder to share with a (individual or category of) Data Requester a set of data. Each data is described by their category and can be shared for a limited period (time_limitation). Data Holder/Requester can be individually identified (in which case they have an identifier), or by their category. Each consent is tied to a specific purpose, has a status for their ethical approval, and a status for consent itself. There is supporting information to the consent with respect to the date at which it was requested (date_requested), and the date at which it was consented (date_consented, when applicable).

It is important to note that data categories and data holder/requester categories contain a list of values which should have persistent identifiers in the entire platform. Therefore, they should likely be managed centrally by a governance board, which decides on the extensive list of categories relevant for the platform, and the creation of new categories when necessary. It falls outside the scope of this Deliverable to describe how these categories should be created and governed; this topic will be discussed in Deliverables 4.4 and 4.5.

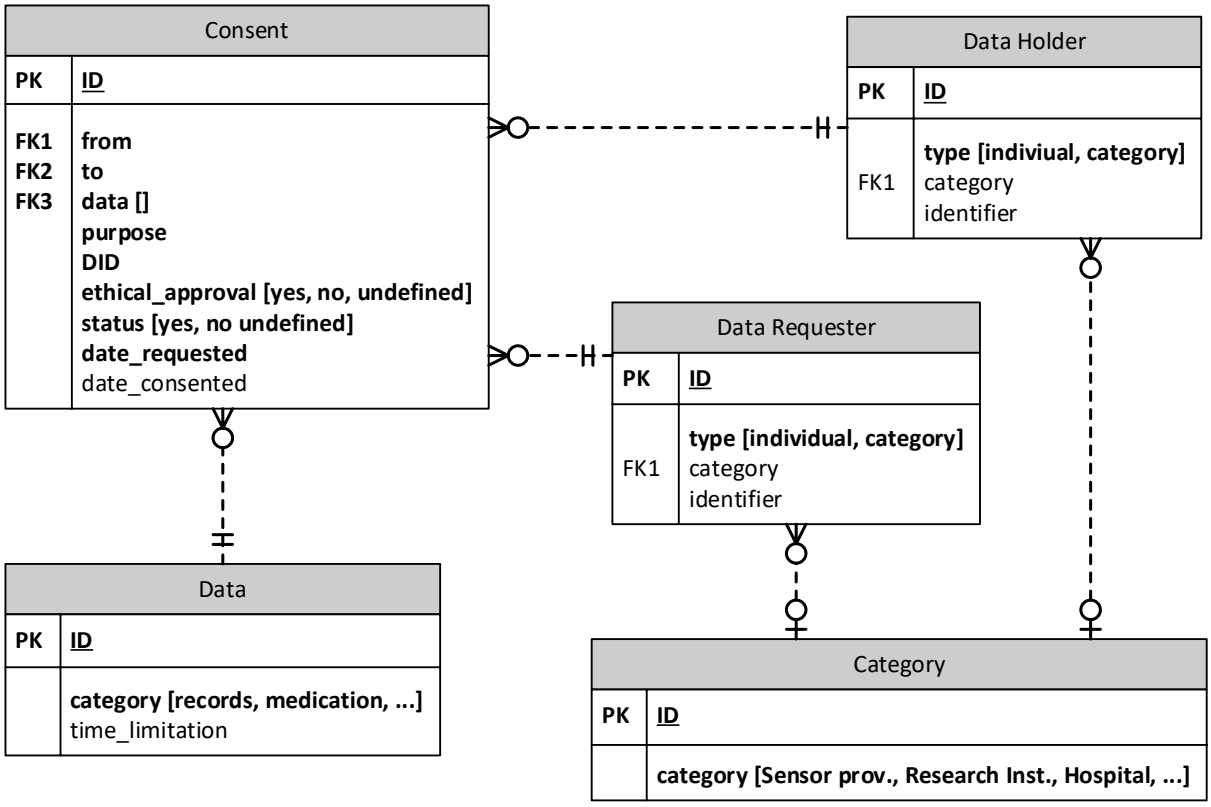

Figure 8 - Bio-curity consent model

6.2 Access control

Access Control is about managing which users have access to which resources, a job performed by system administrators. Manually assigning permissions to specific individuals is costly and error prone, especially for larger organizations where the overview is quickly lost.

Role-based access control (RBAC) is a methodology in access control, first formalized in 1992 and later standardized by NIST in 2004 [15]. It is a simple methodology, where permissions of a user are tied to its role, and these roles are usually tied to the users function within the organization. This reduces the job of specifying permissions per user to specifying the permissions of certain roles, and then assigning those roles to users. RBAC quickly found adoption and was already projected to manage most user's permissions in IT-intensive industries in 2010 [16].

Attribute-based Access Control is a newer [17], alternative methodology for access control in which permission to resources is evaluated dynamically based on the objects and users' attributes, as well as the requested operation and environment conditions. This methodology allows more precise control, as multiple factors can be considered to make a decision, including dynamic factors such as time and location.

Comparison RBAC / ABAC:

- **Flexibility (ABAC)**: The attribute-based approach of ABAC is inherently more flexible. To change or update permissions to a resource, one can simply update the rules rather than

having to modify role permissions. Furthermore, attribute-based rules can be more specific than role-based rules.

- **Simplicity (ABAC)**: RBAC requires thoughtful engineering of the roles to support the security and business requirements of the deployment [18]. The effort required to perform role engineering in large organizations may become too costly [19]. ABAC, however, has no such initial design phase.
- **Administration (RBAC)**: Roles, through their names, provide clear descriptions of sets of permissions, easing administration. The dynamic nature of ABAC makes it harder to keep an overview of the access control in a deployment. Furthermore, in some cases the potentially large number of attributes in ABAC must be understood and managed by expert personnel [18]. This is likely the case in the medical domain, where a more central overview of attributes might be desirable.
- **Auditing (RBAC)**: Because of the dynamic nature of ABAC, in some cases it may become infeasible to audit which users have access to a given permission or which permissions have been granted to a particular user [18].
- **Scalability (ABAC)**: In larger deployments, the set of roles is prone to explode [20], as organizations must handle more complex access control rules.

In [18] [20], the authors advocate for combining RBAC and ABAC, by taking RBAC as a base, describing the maximum permissions of a user, and additionally employing ABAC to flexibly and dynamically restrict a user's permissions. This approach has the benefit that the maximum permissions of a user are auditable, while allowing for more dynamic access control through ABAC.

In the Netherlands, the major providers of software products in the healthcare sector [21] (specifically, healthcare information systems, and electronic patient dossier systems), are Chipsoft, a Dutch company with a market share of roughly 70%, followed by Epic, a US-based company, and Nexus, a German company. Due to their proprietary nature, no public information was found on their access control models.

Attribute-Based Access Control

Some initial attributes have been identified as relevant to the Bio-curity platform, according to the needs emerging from the use cases and the consent model. We describe them below and give some brief explanations of how each of them will influence access decisions. Please note however, that the creation of a catalogue of credentials (including attributes and where possible, their extensive list of possible values, otherwise a description of their expected format) is in scope of Deliverable 4.4, and the maintenance of such credentials is in scope of Deliverable 4.5. These will be further elaborated in future tasks. The list presented here is preliminary and might be adapted or increased in the future.

Attributes are defined for subjects, objects, and other relevant contextual conditions. The subject is the person interacting with the Bio-curity platform. Object refers to the resources to which subjects want access, in the case of Bio-curity these are data. Contextual attributes cover supporting information needed, for instance all information related to consent.

Subject attributes:

- Role: Patient, Researcher, Health Professional, Auditor, Ethical board member …

- Type of organization: Research Institute, Hospital, Sensor Provider, …
- ID organization: for consent types which are from one individual institution to another

Object attributes:

- Type of data: record, medication, activities and diagnosis …

Contextual attributes:

- Consent status: yes, no, undefined
- Purpose of consent: Clinical use, Research, …
- Expiration date of consent: to be checked against date of access
- Ownership: decentralized identifier of data subject who the data is about

Authorization of access

Subjects will have access to data according to the permissions their own organization has, which is extracted from the consent model. For example, an individual employed by research institute A has access to all data for which either research institute A specifically or any organization of the type of *research institute* is authorized to process.

Because of the variety of partners within Bio-curity, one problem the project faces with respect to authorization of access is the diverse and inconsistent definition of roles between organizations. It is not an ambition of the project to unify those definitions. Instead, to tackle this problem, we propose a basic authorization profile which follows the limitations imposed by data subjects (represented through valid consents). It remains up to each partner to further limit permissions according to their own practices.

As access in ABAC is defined according to matching of attributes, it remains a responsibility of each partner organization to: a) request the creation of accounts only for the personnel who needs access to the Bio-curity platform; b) manage accounts in order to ensure accounts and attributes are up to date; and c) define relevant attributes for each role. This way, for instance, a hospital B can define that three people should have access to the Bio-curity platform, two of them being doctors who should have access to records type of data of their patients (see #1), and a researcher who can access data for research purposes (see #2), resulting in access to all data matching the following:

1. ((Subject attributes: Type of organization == Hospital) OR (Subject attribute: ID organization == hospital B)) AND (Object attribute: Type of data == Records) AND (Contextual attribute: Ownership ∈ DIDs ([IDs of patients])) AND (Contextual attribute: Purpose of consent == Clinical use) AND (Contextual attribute: Consent status == yes) AND (Contextual attribute: Expiration date of consent >= now)
2. ((Subject attributes: Type of organization == Hospital) OR (Subject attribute: ID organization == hospital B)) AND (Contextual attribute: Purpose of consent == Research) AND (Contextual attribute: Consent status == yes) AND (Contextual attribute: Expiration date of consent >= now)

Every attribute value underscored is provided by the partner organization, for instance the list of patients for a specific doctor. While the Bio-curity platform is only responsible for fetching all data matching the abovementioned rules and extracting the necessary information from valid consents.

Note: the subject attribute "role" is not used in the authorization rules. That is because we anticipate each organization will have a variety of roles with different permissions. Instead, we chose to represent roles to allow for compatibility, and to allow more meaningful logs of accesses, leaving up to each organization to express the permissions of each role in other terms (for instance, listing all purposes they should be involved).

Additionally, it is important to note that the current proposal for the European Health Data Spaces imposes different rules to requests of health data which are at an individual level[11] and those which are at aggregated level[12]. For aggregated level data, requests are simpler and do not require direct access to datasets, only to a given statistical property answer. Our proposed consent and access control models do not explicitly cover both modalities, however those can be implemented by differentiating between processing purposes, for instance "research" can be represented by two more refined purposes: "individual level research" and "aggregated level research". In such cases, ethical approval for the aggregated modality may become simplified, or even deemed unnecessary. The right granularity level, and the policies around the maintenance of such attributes, as in the case of other attributes too, is in scope of Deliverables 4.4 and 4.5.

6.3 Confidentiality

Confidentiality is a fundamental principle in the healthcare domain, ensuring that patient data is protected from unauthorized access and disclosure. It is particularly critical when dealing with sensitive health data, such as that collected by digital biomarkers. In the context of the Bio-curity project, maintaining confidentiality involves a combination of technological measures, policies, and practices that ensure only authorized individuals can access patient information.

Ensuring that the personal health information (PHI) of the patients' is only accessible to those with necessary permissions is how we enforce Confidentiality in the Bio-curity project. This is crucial for several reasons:

- Patient trust: Patients must be able to trust that their sensitive data will be protected and only shared with their consent.
- Legal compliance: General Data Protection Regulation (GDPR) mandates stringent confidentiality measures for handling PHI[13]
- Data integrity: Ensuring that only authorized personnel can access and make modifications to patients' data helps maintain the integrity of health records.

Measures to ensure confidentiality:

[11] Proposal for EHDS, Art. 45 and 46
[12] Proposal for EHDS, Art. 47
[13] GDPR, Art. 5(1)(f)

1. Encryption: All data, whether in transit or at rest, must be encrypted using strong encryption algorithms. For data in transit, Transport Layer Security (TLS) should be employed, while data at rest should utilize Advanced Encryption Standard (AES) with at least 256-bit keys.
2. Access control: Implement robust access control mechanisms to ensure that only authorized individuals can access PHI. This includes Attribute-Based Access Control (ABAC) (see more details in section 6.2).
3. Data masking: Use data masking techniques to obscure data when full data access is not necessary.
4. Control and transparency: Implement means for patients to control how their data is accessed and provides information about it, allowing them to monitor the use of their data. This includes Consent management (see more details in section 6.1).

Implementing confidentiality in Bio-curity

1. Data encryption:
    a. In transit: All data transmitted between the sensors, Bio-curity platform, and users (patients, healthcare providers, researchers) must be encrypted using TLS 1.2 or higher. This ensures that any data intercepted during transmission cannot be read by unauthorized parties [22].
    b. At rest: Data stored on the Bio-curity platform, including databases and backup storage, must be encrypted using AES-256. This ensures that even if physical storage media are compromised, the data remains protected [23].
2. Access control mechanisms:
    a. Attribute-Based Access Control (ABAC): Allows for more granular access control based on user attributes (e.g., department, location) and resource attributes (e.g., sensitivity level of data) [17].
    b. Multi-Factor Authentication (MFA): Requires MFA for all users accessing sensitive data, adding an extra layer of security by combining something the user knows (password) with something the user has (e.g., security token) [24].
3. Data masking:
    a. Data Masking: Implements data masking for processing where full data access is not required. For example, when sharing data for preliminary research analysis, in some cases masked data may suffice.
4. Patient control and transparency:
    a. Data access logs: Provides patients with access to logs that show who has accessed their data and when. This transparency helps build trust and allows patients to monitor their data's confidentiality.
    b. Consent management: Allows patients to manage their consent for data sharing through a user-friendly interface. Patients can easily view, grant, and revoke consent for specific data uses and access by specific individuals or entities.

6.4 Right to share data

The right to share data refers to the patient's ability to control who can access and share their health information. This principle is integral to data sovereignty and ensures that patients have autonomy over their personal health data. Implementing this right within the Bio-curity platform involves a series

of technical and administrative measures designed to empower patients, facilitate secure data sharing, and ensure compliance with legal standards.

Mechanisms to ensure the right to share data:

1. Consent management:
   Explicit consent: Implements a robust consent management system where patients can easily grant and revoke permissions for data sharing. Consent must be informed, explicit, and specific to the purpose of data sharing. This is critical for ensuring that data is shared legally and ethically, as detailed in section 6.1 on *Consent Management*.
2. Data portability:
   Standardized Formats: Ensures that patients can easily transfer their data to other healthcare providers or platforms in standardized formats such as HL7/FHIR. This enhances interoperability and patient control over their health data [25, p. Art. 20].
3. Transparency and communication:
   Clear Information: Clearly communicates to patients how their data will be used, who will have access to it, and the benefits and risks of sharing their data. Transparency helps build trust and ensures informed consent, which is also a recurring theme in sections 6.1 and 6.3.
4. Granular control:
   Specific Permissions: Allows patients to specify which types of data they want to share and with whom. This includes providing options for sharing data at different levels of detail (e.g., aggregated vs. individual data). This level of control is essential for maintaining patient trust and ensuring data security [26].

State-of-the-Art

The current state of the art in the right to share data revolves around robust consent management, data portability, user-friendly interfaces, transparency, and advanced security measures. The General Data Protection Regulation (GDPR) in Europe is a leading example, providing a framework that ensures individuals have control over their personal data. The GDPR requires a legal basis for data processing, one of which is explicit consent, and for the Bio-curity project, explicit consent is the chosen basis. Additionally, the GDPR mandates the right to data portability and transparency in data use.

Technologies like blockchain and self-sovereign identity are being explored to enhance data sharing capabilities while ensuring security and patient control. Blockchain offers decentralized and immutable records that can support transparent and secure data sharing[14]. Self-sovereign identity allows patients to manage their digital identities and credentials, ensuring that they have control over who accesses their data [27].

Interoperability standards such as HL7/FHIR are crucial for enabling seamless data exchange between different healthcare systems [28]. These standards ensure that data can be shared in a structured and consistent manner, enhancing the ability to integrate various health IT systems.

In practice, platforms like MedMij in the Netherlands and initiatives such as the Personal Health Record (PHR) systems in various countries are leading the way in implementing these principles. These

[14] https://www.healthit.gov/topic/scientific-initiatives/blockchain

platforms provide patients with access to their health data and the ability to share it with healthcare providers and researchers securely and efficiently.

Overall, the right to share data is being increasingly recognized and implemented through advanced technologies and regulatory frameworks, ensuring that patients have control over their health information while maintaining high standards of security and interoperability.

Implementing rights to share data in Bio-curity

1. Comprehensive consent management system:
    a. Centralized consent portal: to implement a centralized consent portal within Bio-curity platform where patients can manage all their consents for data sharing. This portal should allow patients to grant, review, and revoke consents easily. This builds on the foundation laid out in section 6.1 on Consent Management.
    b. Granular consent options: to provide patients with granular options for consent, allowing them to specify which data types can be shared, with whom, and for what purposes. This ensures that patients can exercise fine-grained control over their data.
2. Data portability and interoperability:
    a. Standardized data formats: to ensure that data can be easily exported in standard formats such as HL7/FHIR. This facilitates seamless data sharing and integration with other healthcare systems, enhancing interoperability as discussed in section 6.2 on Access Control.
    b. APIs for data transfer: to develop APIs that allow secure and efficient data transfer between the Bio-curity platform and other healthcare systems. Ensure that these APIs adhere to industry standards for security and interoperability.
3. Security measures for data sharing:
    a. Encryption and secure channels: to ensure that all data transfers are encrypted and conducted over secure channels. This protects data integrity and confidentiality during sharing, building on the encryption measures discussed in section 6.5 on Encryption.
    b. Access control for shared data: to implement access control measures to ensure that only authorized individuals can access shared data. This includes validating the identity and permissions of recipients before allowing data access.

6.5 Encryption

Encryption is a technology to help protect user data against unauthorized access. Laws and regulations, targeting both general information security and healthcare domain specific regulations, mandate the use of best-practice protection mechanisms to protect user and patient data.

Within the Bio-curity platform we deal with both data at-rest and data in-use or in-transit. The data is stored at its source at rest, such as in the PHE or in the database of a sensor provider. As the data at-rest is managed by a single institution, it is the institution's responsibility to properly manage confidentiality (discussed further in section 6.3). In general, encryption and other best-practice information security protection mechanisms need to be in place and be part of the information security management system of the organization [29].

Within the context of data exchanges, we must also ensure that data is safely transferred between components and that only those authorized receive the data. For simple data transfers, for example a transfer of sensor data from a sensor provider to a researcher, traditional methods of encryption for data in-transit suffice. Such methods are well-known and standardized and should fit with any technology chosen to implement the architecture with.

More interesting is to consider the encryption of data in-use. Within the scope of the data exchanges considered, a researcher is interested in data because of the value of the analysis produced with the data, and not the data itself. As such, there is no need to give a researcher access to raw data, if aggregated or analyzed data suffices. Not only would not sharing raw data reduce the security and privacy risks associated with handling this data, but it could also increase a patient's trust and willingness to share data, as they can rest assured that their raw data is not accessible to any person.

The set of technologies focusing on this type of data sharing are called Privacy Enhancing Technologies (PETs). Examples of Privacy Enhancing Technologies include homomorphic encryption, multi-party computation (MPC), and federated learning (FL). It could be interesting to explore how these technologies can be incorporated in and supported by the architecture, such that the data exchanges attain even higher privacy standards and reduced information security risks.

6.6 Data Pseudonymization

Data pseudonymization is a practice widely adopted to protect personal data where direct identification to people is not necessary, but maintaining the ability to link different data records on the same person is necessary. In general, data analysis and research can be conducted on pseudonymized data.

In the Bio-curity platform, the adoption of a personal health environment already supplies an automatic solution for data pseudonymization, as (meta-) data is not linked to personal direct identifiers, such as names, or identification documentation number, but to a decentralized identifier, which serves as the "alias" to a person. With decentralized identifiers, data about the same subject can be linked through these identifiers, and additional identity information can be presented by the personal health environment where necessary. However, diving into the attributes and types of metadata which should be stored in such a PHE, falls outside the scope of this Deliverable. This is a task which is more closely related to Deliverable 4.4 on the formulation of a "credentials catalogue", and therefore the exploration of this criterium will be realized there.

To this end, we anticipate at least the following activities will become relevant:

- To identify which types of health data to pseudonymize, an exploration of all relevant health data types and their categorization should be conducted. Expanding on the small selection of health data described in this deliverable (as presented in 6.1 and 6.2). Each category should be studied for their sensitivity and criticality.
- For completeness, a study of state-of-the-art pseudonymization techniques and the best practices applied by the industry should be conducted.
- As well as an assessment of the impact of pseudonymization on the quality and utility of health data, which will feed discussions on strategies to mitigate any adverse effects on data quality. This should be done, as much as possible, in cooperation with Bio-curity consortium partners.

7. Discussion and Conclusion

In this work we explore ways to support handling data from medical wearables and making it available for novel insights. We build upon the security and assurance criteria elicited in Deliverable 4.1, and the Bio-curity use cases described in WP2, to propose a trust model technological architecture that can accomplish the requirements for qualified data exchange in Bio-curity.

The exercise of further exploring the criteria also resulted in input for work conducted in Task 4.3 “Trust model governance framework”. In there, processes and sub-processes related to the Bio-curity use cases are explored. Each process is ensured to have a role accountable for it and others responsible, informed or consulted, where relevant. Collectively, these processes are designed to address relevant technical standards and legal requirements, which are expressed by the criteria.

The work in this architecture resulted in a decentralized and patient-centric architecture for supporting data exchanges, focusing on the aspect of identity management, access control and consent management. The architecture leverages the concepts of Self-Sovereign Identity to achieve this decentralization.

The decentralized approach taken has many advantages. By making the identity of the patient central to the architecture, we provide the ecosystem with patient-centric interoperability. In a future where the patient itself also stores the Electronic / Personal Health Record; the patient becomes the single point for interoperability in the healthcare ecosystem. Instead of having to connect different healthcare systems managed by countless institutions, systems in healthcare will only need to be interoperable with the personal health environment to retrieve an update of the EHR/PHR.

There are, however, many considerations and open questions. To name a few, the scalability of the underlying technology will play an important role in the feasibility of such a decentralized approach. Furthermore, for true interoperability, the role of the PHE needs to be standardized across different providers. The PHE also poses usability questions. Giving too much responsibility to the patient in the data exchange and the management of the EHR/PHR could render the system unusable for the average citizen.

This architecture does not dive into any technological specific details. Furthermore, the architecture mainly considers the management of trust and not that of data. As such, topics such as data standards, semantic interoperability, and interoperable software interfaces are not considered in any detail. Adopting such standards will be crucial when working out the technical details of a decentralized architecture.

Throughout this deliverable we expose a few open questions regarding the management of credentials and the governance plan for the envisioned Bio-curity platform. These are systematized and will be used as input for future work in WP4 tasks “credentials catalogue” and “sustainability plan”.

8. References

[1] "Elektronische inzage medisch dossier," Patiëntenfederatie Nederland, 06 2020. [Online]. Available: https://www.patientenfederatie.nl/downloads/brochures/457-202006-factsheet-elektronische-inzage-medisch-dossier/file.

[2] P. Zhang, W. Jules, S. C. Douglas, G. Lenz and T. S. Rosenbloom, "FHIRChain: Applying Blockchain to Securely and Scalably Share Clinical Data," *Computational and Structural Biotechnology Journal,* vol. 16, pp. 267-278, 01 01 2018.

[3] European Comission, "Report on the first review of the functioning of the adequacy decisions adopted pursuant to Article 25(6) of Directive 95/46/EC," 2024.

[4] A. Hasselgren, K. Kralevska, G. Danilo, P. A. Sindre and A. Faxvaag, "Blockchain in healthcare and health sciences—A scoping review," *International Journal of Medical Informatics,* vol. 134, 01 02 2020.

[5] S. Angraal , H. M. Krumholz and W. L. Schulz, "Blockchain Technology: Applications in Health Care," *Circulation: Cardiovascular Quality and Outcomes,* vol. 10, no. 9, 09 2017.

[6] G. G. Dagher, J. Mohler, M. Milojkovic and P. B. Marella, "Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology," *Sustainable Cities and Society,* vol. 39, pp. 283-297, 01 05 2018.

[7] A. Siquerira, A. F. Da Conceição and V. Rocha, "Blockchains and Self-Sovereign Identities Applied to Healthcare Solutions: A Systematic Review," 25 04 2021.

[8] W. J. Gordon and C. Christian, "Blockchain Technology for Healthcare: Facilitating the Transition to Patient-Driven Interoperability," *Computational and Structural Biotechnology Journal,* vol. 16, pp. 224-230, 2018.

[9] Ministerie van Volksgezondheid, Welzijn en Sport, [Online]. Available: https://www.gegevensuitwisselingindezorg.nl/gegevensuitwisseling. [Accessed 01 02 2024].

[10] MedMij, "Toestemmingsverklaring, Core (Declaration of consent, Core)," [Online]. Available: https://afsprakenstelsel.medmij.nl/asverplicht/mmverplicht/toestemmingsverklaring-core. [Accessed June 2024].

[11] VZVZ, "Appendix | Consent situations," October 2023. [Online]. Available: https://vzvz.atlassian.net/wiki/spaces/MA11/pages/127536044/Bijlage+Toestemmingssituaties. [Accessed June 2024].

[12] R. v. d. Hoek, R. v. Holland, V. Teunissen and F. Horst, "Eindrapportage PROVES Datakluis (Final report PROVES data vault)," MedMij, 2023.

[13] G. Remmers, A. Boorsma, H. Duinkerken and M. v. Lieshout, "Position Paper: Gezond Akkoord," 2021.

[14] J.-H. Hoepman, Privacy design strategies (the little blue book), Nijmegen: Radboud University, 2018.

[15] I. T. L. Computer Security Division, "Role Based Access Control | CSRC," 2016. [Online]. Available: https://csrc.nist.gov/Projects/Role-Based-Access-Control. [Accessed May 2024].

[16] A. O'Connor and R. Loomis, "Economic Analysis of Role-Based Access Control: Final Report," 2010.

[17] V. Hu and et al., "Guide to Attribute Based Access Control (ABAC) Definition and Considerations," National Institute of Standards and Technology, NIST Special Publication (SP) 800-162, 2019.

[18] E. Coyne and T. R. Weil, "ABAC and RBAC: Scalable, Flexible, and Auditable Access Management," *IT Professional,* vol. 15, no. 3, pp. 14-16, 2013.

[19] A. Babko, "RBAC vs ABAC: Comparing Advantages & Disadvantages Before Choosing," 2023. [Online]. Available: https://www.ekransystem.com/en/blog/rbac-vs-abac. [Accessed May 2024].

[20] D. R. Kuhn, E. J. Coyne and T. R. Weil, "Adding Attributes to Role-Based Access Control," *Computer,* vol. 43, no. 6, pp. 79-81, 2010.

[21] KPMG, "Een marktverkenning naar informatiesystemen en digitale gegevensuitwisseling in en met de ziekenhuissector," Autoriteit Consument & Markt (ACM), 2023.

[22] K. McKay and D. Cooper, "Guidelines for the Selection, Configuration, and Use of Transport Layer Security (TLS) Implementations," NIST, 2019.

[23] E. Barker, "Recommendation for Key Management: Part 1 – General," NIST, 2020.

[24] P. A. Grassi and e. al., "Digital Identity Guidelines - Authentication and Lifecycle Management," NIST, 2023.

[25] European Parliament and Council of the European Union, *Regulation (EU) 2016/679 of the European Parliament and of the Council,* 2016.

[26] National Institute of Standards and Technology, "Privacy Framework," 2020.

[27] J. McWaters, *A Blueprint for Digital Identity,* 2016.

[28] HL7, "FHIR Overview," HL7, 26 03 2023. [Online]. Available: https://www.hl7.org/fhir/overview.html. [Accessed 28 06 2024].

[29] Nederlandse Norm, "Medische informatica – Informatiebeveiliging in de zorg Deel 1: Managementsysteem," Nederlandse Norm, 2020.