Privacy Respecting Data Sharing and Communication in mHealth: A Case Study

Michael Pleger, Ina Schiering

Privacy Respecting Data Sharing and Communication in mHealth: A Case Study

Michael Pleger[0000−0001−5398−3694], Ina Schiering[0000−0002−7864−5437]

Ostfalia University of Applied Sciences, Wolfenbüttel, Germany
{mic.pleger, i.schiering}@ostfalia.de

Abstract. The increasing usage of mobile devices within the healthcare domain and the social sector requires a closer look on privacy and security aspects of those systems. Major data breaches have been made public in recent years, which resulted in privacy issues for unaware users, ranging from loss of control over their data, to identity theft cases, up to discrimination because of political, religious or sexual orientation. To this aim, privacy enhancing technologies enabling privacy-preserving communication and data sharing are investigated in the context of an mHealth solution.

Keywords: mhealth · data sharing · attribute-based encryption · privacy enhancing technology · zero knowledge

1 Introduction

Mobile technology like wearable devices and mobile phones, for example, utilize existing communication technologies to provide a vast amount of unique services, all within a single device. While these devices are used for personal health-monitoring, they can be accounted as digital, mobile health solutions (*mHealth*).

Safety and security concerns of mHealth applications have been broadly addressed [11,14,24,27,15], where, among others, unencrypted data gets transmitted to remote servers without notifying the user of their purpose and without compliance with data regulations applicable for mHealth apps, e.g. GDPR [5]. In several contexts, mHealth applications need to gather health related data from patients or vulnerable people, which are considered as special categories of data according to Article 9 GDPR. To protect such data from unwanted access (both locally on a device, as well as remotely stored on servers) cryptographic methods like encryption can be generally used. While a secure communication channel to a remote server might protect the transmission itself, it leaves data stored on a remote server completely unencrypted. To protect such data against unauthorized access and prevent data breaches, it should be stored encrypted as well. End-to-end encryption can be managed to complete such task, but requires a generally more complex approach concerning key management and distribution.

The aim of this paper is to investigate privacy patterns and privacy enhancing technologies for privacy preserving communication and data sharing in the context of an existing mHealth solution. Within the healthcare domain such

data typically is highly sensible due to their nature in providing insides for professional assessments. Reports of medical- or physical conditions, vital signs or blood-level measures are information that should be protected and limited to eligible personal. As an example, if a doctor receives a radiograph from a hospital stay of one of his patients (e.g. a public figure), it does contain identifying information (e.g. the patients full name, birth date, date of x-ray, etc.). What happens, if such scan would be leaked to the public it could have significant, unclear consequences for the patients. Therefore the communication between the doctor and clinic should be secured, as well as the access to such data within the clinic limited without proper authentication. Furthermore the electronic data storage should be encrypted as well. Within a mHealth-domain, on the other hand, if the patient is using a wearable self-monitoring devices (e.g. a fitness tracker) that collects vital signs (step count, heart rate) over a period of time and shares these information with the devices manufacturer for processing, there is typically not such a strict data protection requirement. It is not uncommon to buy such devices with a privacy policy attached, that requires the publication of personal data towards the manufacturer or application provider, in order to use the device itself, since often times the data processing is done on remote servers due to the limited capabilities these devices face.

The selection process for privacy enhancing technologies (PETs), based on the given requirements in the selected use case, as well as the intended purpose with such PETs, is provided within the methodology section of this publication. The analysis focuses on selected approaches addressing untrustworthy service providers and the risk of data breaches at the service provider level respectively. Hence, which privacy patterns can be utilized to improve privacy in data sharing towards untrustworthy platform providers in the context of an mHealth application.

Two prototypes, based on the analysis, were developed to further evaluate the usability by less technological skilled people. After that the shortcomings for both prototypes are then be provided within the discussion section, along with a brief overview of further technologies, that could be utilized to improve the privacy aspects of data sharing within mHealth.

2 Background

Smartphone based mHealth applications and *Assistive Technologies* (AT) enable users to monitor their own vital-signs and e.g. allow them to provide data to physicians remotely for quality and improvement assessments without the need of personal, eye-to-eye contact [14]. This technology enables users to engage in healthcare related activities on their phones rather than using traditional, stationary personal computer systems, while being flexible in their location [19]. Smartphone based mHealth applications provide various types of services, e.g. tracking behavior for developmental disorders, information and treatments on cognitive disorders, as well as mood, eating and sleep disorders among others [27,15,19], which can be very useful for patients at home.

The downside of this technology has been continually picked up by ongoing discussions regarding user's privacy, since (mHealth) applications tend to collect and share significant amounts of data to provide their services, even if the purpose of such collections is not always disclosed [11,24,27]. Common identifiers like name, age, weight and address might be collected, which according to data protection regulations, can be classified as *Personally Identifiable Information* (PII). Furthermore, depending on the application and what service it provides, the collection of fitness data (including location data), medical results or personal images, treatment plans or -schedules might be necessary.

While users could be enlightened about the usage of such data through privacy policies on an informational level (e.g. what data is collected, where is the data processed, who is processing the data and with whom is the collected data being shared), this would not protect the data itself on a computational level, e.g. when the data is stored on a server or within a database. This needs to be addressed with technical means, like data encryption, -separation and access-structures within the realm of ICT, or information and communications technology.

To set the scope regarding privacy within mHealth in the context of this paper, a typical use case will be provided as follows (see Illustration in Figure 1). Whenever patients require services from within the healthcare domain, they generally are required to provide personal identifiable information (PII) like name and address, age, their social security number, healthcare account number or other highly sensitive data like blood type or medical conditions, in order to receive services by healthcare professionals. For example, if a cognitive impaired patient requires assistance to achieve daily tasks (e.g. to stock up groceries), through a mobile health application instead of a caregiver or personal care assistant, sensible information like time schedules, preferences and what items have been bought could be gathered by that application. This could leave the patient vulnerable to information loss, since the item list could also contain prescriptions, which could disclose the patients medical conditions. On the contrary, using an application instead of a human assistant might increase the anonymity in regards to the patient, which could be beneficial (depending on the conditions).

As a generalization, the issues for privacy can be seen by mHealth solutions and -services that do not have access to larger, established medical solutions and electronic health record systems, to process the patients data in a privacy preserving way. Instead, smaller or dedicated service providers are used to process and exchange information with the healthcare domain, which could expose sensible personal identifiable information to a broader scale.

Such data required by healthcare professionals, which contains sensitive information, is then often being processed through dedicated software applications, in order to provide services to the patient. Since healthcare professionals often times rely on software developed by third parties, they do not have (and should not require) the knowledge of how such personal information are been processed or stored. What they require is the sole outcome of the data processing, therefore

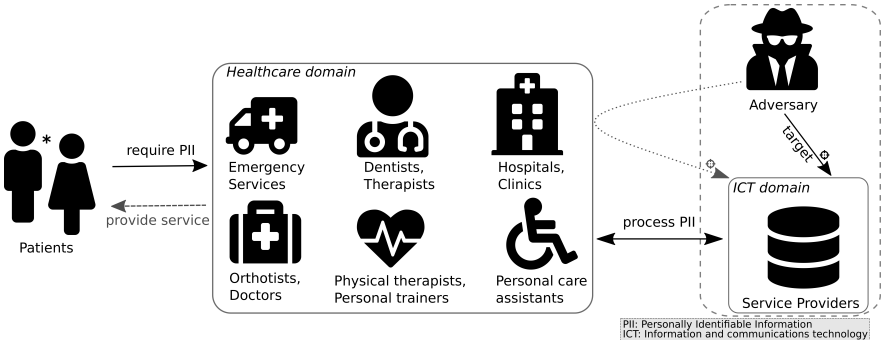

Fig. 1. Generalized data processing flow in mHealth, where healthcare professionals require third party providers for their services. In such case adversaries typically target service providers within the ICT domain directly

their understanding of information processing can be seen limited, since these are typically not tech-savvy users. For example, rehabilitation centers and personal care assistants are specialized to provide a service (e.g. consultations) to their patients and utilize applications to fulfill their tasks.

A distinction has to be made between healthcare (and electronic health) and mHealth. While the former is based on professional healthcare and -systems (e.g. patient management software within clinics, clinical or laboratory software, etc.) the latter one often times refers to mobile applications, that tend to assist or self-monitor instead of treat them. Such applications are, for example, Apple Fitness, Google Fit, FitBit or Strava.

In regards to mHealth applications on mobile devices (and the increasing availability of such), healthcare professionals tend to have difficulties with what kind of application should be recommended to patients. Since a vast amount of these applications have not been created by medical experts, healthcare professionals, like doctors, have to evaluate applications by themselves to validate their functions. If the application's functions vindicate the intended purpose, without proper validation of the safety and security mechanism (or certification thereof), massive privacy and security risks need to be addressed.

While a healthcare professional might be able to validate the purpose of such mHealth applications, this does not hold true for the data protection against service providers, in regards to unauthorized access, while maintaining the user's control over the data. Furthermore it might be non-trivial to use systems that deliver strong privacy implementations without advanced training.

As context for the investigation presented here, an mHealth application called *RehaGoal* is considered. It has been developed to provide treatment capabilities to patients with an impairment of executive dysfunction, which could be received either through an accident or a congenital disease. Such an impairment could result in patients being unable to perform consecutive tasks on their own. For example, if someone is asked to water all plants in a specific room, a healthy

person would water each plant successively through a pattern (e.g. start from left and go to right) until each plant has been watered sufficiently. This behavior would not be the case for a person diagnosed with executive dysfunction, where the motion to stop and reflect the acted behavior might be unavailable for such person. This could result in watering all plants multiple times in a row, until the motion to stop is either remembered or the person is notified by external means (e.g. another person). To change or treat such false behavior, the concept of *Goal-Management Training* has been used to define therapies for affected patients.

Individual therapies can be provided in a supervised environment, with assisted interactions between patients and therapists (e.g. in a clinical setting). It could also be used individually in a private environment, whenever personal interactions with healthcare professionals might not be feasible or possible (e.g. during a pandemic) [20].

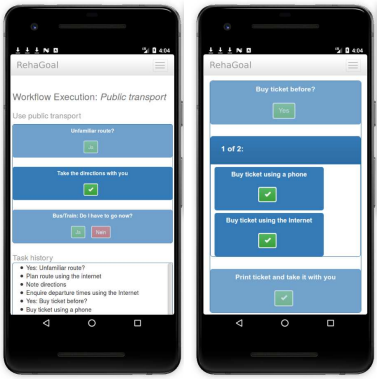

Fig. 2. Representation of the RehaGoal application to support people with cognitive disabilities during daily activities

Developed as a hybrid, cross-platform web-application with a strong focus on mobile devices (e.g. smartphones, see Figure 2), its intended purpose is to assists patients in completing a defined goal, by serving a sequential list of single tasks as a workflow. Users are able to create, execute, schedule or share workflows inside the application. Workflows can be designed with a graphical editor by placing task-blocks on a virtual canvas. Workflow tasks are typically modeled on a textual basis, but images can be added for increased accessibility as well. RehaGoal users can be differentiated into two roles: *patients* and *therapists*. The latter typically designs workflows for each patient individually on a stationary device (e.g. a desktop computer or notebook) by splitting an intended goal into sub-goals or -tasks. Patients on the other hand can be considered as the main users, which run the application on smartphones or other mobile devices like tablets to execute workflows in the context of a given therapy.

Although no vital data of patients are collected, the individual workflows might include names of persons or addresses, required doses of medicals for a given therapy or photos might show people and the personal environment of patients. Also, only the fact that a person uses the app on his/her mobile device allows conclusions, that the user has a certain disease or impairment (thus this can be seen as metadata).

An important requirement for neuropsychological trials and usage of the app afterwards in rehabilitation scenarios is the ability to exchange workflows between patients and therapists. While each patient might receive individual treatments and intensities with the application, it could occur that pre-defined workflows are applicable to multiple patients. Furthermore, if the application is rolled out through small to large scale therapy institutions (e.g. clinics), the ability to distribute workflows is required. Sharing can be done either locally through an export on the device or alternatively the data exchange can be realized via a remote server. A visualization of that requirement can be seen in Figure 3, which illustrates the life-cycle for workflows, which where designed by therapists and provided to patients. The inner arrows picture a local file exchange, while the outer ones attached with the cloud symbols represent the server method. In case the rehabilitation is commenced within a larger institution (or e.g. patients with severe impairment), a mutual handler can be utilized to assist patients to load and start their workflows, or schedule multiple workflows for a given therapy and provide feedback to the therapist. While in this context, patients, therapists and the mutual handler are not considered as tech-savvy, the use of an secure and privacy respecting data exchange solution must be very simplistic. A workflow export, that requires a unique password to be defined, which has to be communicated through a secure channel and entered prior to importing, can be seen as to cumbersome and not suitable for deployment. Furthermore, the protection against unauthorized access and control over the data by the user should not be relayed to the end user, but instead be mitigated through information processing means within the application itself.

During the development process of the RehaGoal application both privacy and accessibility were accredited of utmost importance. Therefore *Privacy by Design* principles have been incorporated, which were based on *privacy design strategies* (PDS) [2] in conjunction with the definition of *privacy protection goals* (PPG) [9]. Furthermore the concept of data minimization was adhered to, wherever possible [6]. The RehaGoal application has been utilized during neuropsychological intervention-studies to evaluate whether the usage of *Assistive Technology* can improve the participation of patients with cognitive impairments [20].

A central characteristic of mHealth scenarios is that the organizations are typically therapists or physicians, working in relatively small practices, hospitals or rehabilitation facilities, which are (in general) not able to host software services. Hence external service providers would be typically needed, thus it is important to analyze Privacy by Design approaches concerning untrusted platform service providers.

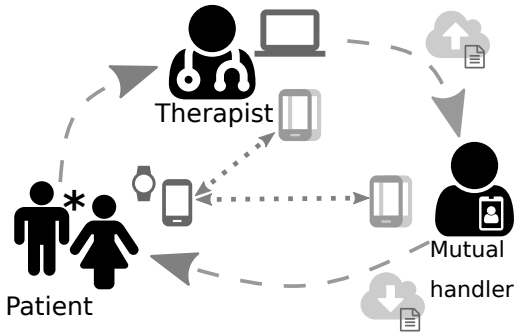

Fig. 3. Workflows created by therapist are shared either directly with patients, or in case of a larger deployment (e.g. within an institution) via a mutual handler that can assist in importing, scheduling or executing workflows for impaired patients

Standard security measurements would be authentication of users in combination with a role and right model, network encryption and the use of pseudonyms instead of usernames, allowing an identification. Though access to the cloud based database is limited by authentication, a data exchange solution should not solely be used to protect data from unintended access. Unlinkability is, in addition, an important goal. In the RehaGoal context, user accounts on the remote server are based on pseudonyms generated by therapist rather than the patient's name and do not contain addresses or other highly sensitive information.

To protect against unauthorized access through server- or database breaches, as well as rogue platform service providers (e.g. system administrator) and furthermore to hide potential metadata, workflow encryption should be used. But this requires further evaluation on how such an encryption scheme could be both, designed and implemented, because the selection depends on regulations and the perspective. A straightforward solution would be the encryption of any data by the users themselves, including the key management and -exchange. Although this solution would provide secure means regarding the remote server, it lacks the ability to be scaled up efficiently. Larger businesses or institutions (e.g. healthcare clinics or hospitals) could require the usage of a remote server-oriented workflow exchange with an elaborated role-based access model, where therapists and doctors of different departments can treat patients in unison.

Since the RehaGoal application (and the therapies thereof) can be situated within the mHealth domain, the following distinctions will be made to reflect stakeholders, with the data protection terminology defined in the GDPR. In this context data subjects will be referenced as users (patients and therapists), which provide data to the RehaGoal application by e.g. generating and executing workflows. A remote server will be provided by an platform service provider, which can be seen as a processor of the provided data, since these servers are being used for workflow file exchanges. The development and research group of the RehaGoal application can be seen as both, data controller and -processor as

well, since usage-metrics might be collected in case patients have joined selected neuropsychological studies.

3 Related Work

To address and promote privacy friendly solutions, Colesky et al. [3] documented a collection of privacy pattern for multiple categories, e.g control of data, minimization or enforcement of policies. Theses have been classified based on previous work of Colesky et al., where privacy design strategies have been extended with privacy tactics intended to address privacy issues, by *Privacy by Design* during the development process. Further research has been commenced to characterize such privacy patterns [16] and on how these patterns could be applied to the healthcare domain [1]. Since authentication and encryption are necessitated for secure data sharing in mHealth, the usage of such patterns [8] is evaluated.

While proper encryption standards to protect data communications on the internet have been developed and used for a long time [13], they generally require an elaborated key distribution and -management scheme. Those communications rely on a trust-model that is based on certificates and public key cryptography. To mitigate the trust-model, newer technologies like Blockchain, advanced by crypto-currencies, have been introduced and used to authenticate and store data online [4,12]. A shift from Single-Factor-Authentication (SFA) to Two-Factor-Authentication (2FA) or even Multi-Factor-Authentication (MFA) has increased the security aspects of such processes [23].

On the other hand, processing encrypted data within the cloud can be achieved by advanced cryptographic schemes, like homomorphic encryption, which permits a limited evaluation of encrypted data without the need for a decryption key [7]. Attribute-based encryption (ABE) extends asynchronous encryption, based on public/private keys and incorporates an access structure within the encrypted data itself, mitigating the need for elaborated key distribution schemes [8]. This concept has been extended with approaches for distributed [21] and decentralized [18] implementations.

To protect sensitive information in a privacy preserving manner, intervention studies typically rely on methodologies that combine pseudonymization and encryption schemes to reduce e.g. linkability between data sets [10,22]. Polymorphic encryption and Pseudonymisation (PEP) by Verheul et. al [26] is a novel approach suitable for such applications, based on a similar idea of ABE where the decryption relies on policies rather then fixed keys.

4 Methodology

Patients and therapist of RehaGoal, or users of any mHealth solution in general, use an application to fulfill an intended purpose (e.g. to execute a treatment plan or monitor their vital signs). The user's knowledge, in regards to information security and data privacy, can not be expected in this domain and should, in a best case scenario, only contribute to the validation that an application or

policy preserves the users privacy. Since complex applications, that require user interactions to set privacy settings, tend to be difficult to get used to, they should provide a user centered privacy setting by default. Otherwise this behavior could burden users with overwhelming settings of which they do not know the ramifications for, require advanced training or would raise the user acceptance threshold, using the application in general.

To identify, which privacy patterns or privacy enhancing technologies could be utilized to improve privacy in data sharing, specifically against untrustworthy platform providers, an overview of suitable solutions had to be created. Figure 4 illustrates four steps, that have been established. The initial step (*Identify*) was to gather an overview of data sharing solutions, based on common proposals, current research and best practices. This selection was not limited to the mHealth domain and listed, among others, file exchange through messages (e-mail), peer-to-peer solutions (P2P), network based file shares (e.g. Network File Share [NFS], Server Message Block [SMB]) or role-based access structure concepts on private and public clouds.

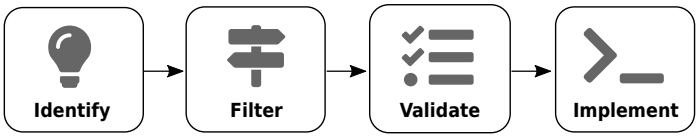

Fig. 4. Established steps, that have been performed to generate an overview of sharing solutions (1.), narrowed down to a subset based on privacy design strategies (2.), adjusted to the requirements by the RehaGoal use case (e.g. low-powered mobile devices) (3.) and realization of prototypes to evaluate the functionality and conduct usability studies (4).

The second step (*Filter*) was conducted to narrow down the subset of solutions based on their classification within the privacy design strategies, with emphasis to minimize trust towards platform providers. For example, in this step solutions that utilize a peer-to-peer network have been excluded, since this scheme only provides a mean to transport data (with package validation), but does not provide further means to protect the actual data itself. Furthermore, this solution does not provide a reasonable mean to delete already shared data, reducing the intervenability in case users do want to remove their data, significantly. Therefor only solutions that utilize a (de-)centralized client/server architecture where processed further. Within the RehaGoal use case, one requirement was the protection of sensible health data, that should adhered to on an end-to-end basis. This means that only patients and e.g. their therapists should be able to decrypt the shared data respectively. The utilization of service providers for remote data sharing should only require the storage of data by those. Data processing should not be outsourced from the mHealth application itself. Following

that, shared data should not provide any insight for service providers into how users workflows look like, or what they intend to solve.

The third step (*Validate*) imposed further requirements to the filtered subset. All proposed solutions are required to be used on low-powered, mobile devices. Furthermore, solutions should be, if possible, utilize established web-standards and assessed cryptographic libraries, since the usage is intended for a production environment. The mHealth solution should be usable by users with medical conditions or impairments. This means that the need for a key management (in case data is encrypted) should be easy to use without the need for users to choose complex (e.g. secure) passphrases for each data exchange. The need for advanced user training should also be avoided. Instead the application and their processes should be as self-explanatory as possible. And lastly workflows need to be shared with multiple recipients or re-used with others. Since cryptographic computations on mobile devices are rather costly (e.g. battery usage, wait time), they should not be encrypted for each individual separately. Instead an encryption scheme able to encrypt once (for either all or an intended group instead) should be considered. This process generated an overview of suitable privacy enhancing technologies and data sharing schemes, that are listed in Table 3.

After that was the realization (*Implement*) of selected concepts within RehaGoal, in order to evaluate the functionality for both users and service providers, the computational costs and resources requirements. This step provided artifacts that can be used to conduct usability studies in the future.

While this provided the general steps that have been passed through, the following text will provide more context on the filtering and validation process.

Privacy by Design is generally focused on design principles and tends to be difficult to be overcome from a software development standpoint. In order to select suitable privacy patterns, which provide the ability to preserve privacy within a data sharing scheme, especially towards platform service providers, a general filtering process has been commenced. The selection process had two phases:

The first phase identified applicable privacy patterns for data sharing on a conceptual level. Specifically patterns should be intended to improve data confidentiality for RehaGoal users, along with reduced linkability features for the platform service provider. Based on the privacy design strategy tags, that have been used within the catalog provided by Colesky et al. [3], a pre-selection of suitable privacy enhancing technologies had been commenced. The selection has then been cross-referenced with the classification matrix identified by Lenhard et. al [16], to focus on data oriented categories, namely *MINIMIZE, AGGREGATE, SEPARATE* and *HIDE*. Table 1 lists the included privacy design strategies, based on Colesky et. al [2] with the reasoning and an example as how they can be applied in the stated use case.

Process oriented categories have largely been excluded, since privacy towards service providers should be addressed, rather than the user's means to *ENFORCE, DEMONSTRATE, CONTROL* or *INFORM*. Lastly the results had been compared to Aljohani et. al's findings [1], who identified privacy patterns

Table 1. Overview of the included privacy design strategies with their intentions and implications for the given use case

Included PDS	**Reasoning**
HIDE	Since this improves both unlinkability and confidentiality, the strategy seemed most promising. For example by encrypting the workflow data before sharing it with the platform service provider, a general obfuscation could be accomplished.
CONTROL	This strategy allows users to have the ability to choose who can process or access their personal data. By encrypting workflow data through the *HIDE* strategy and the management of related encryption keys by the RehaGoal users themselves, this could provide useful means to improve the intended purpose.
SEPARATE	Tries to prevent correlation between data as much as possible by either distribution or isolation. Since the remote server is intended to be realized as a single instance the isolation tactic could be utilized to process (store and provide) workflow data without access or correlation to further information (e.g. content of other workflows of each user).

in a similar mHealth use case on a larger, nation-wide scale. Since the RehaGoal application is focused on individual deployments for e.g. rehabilitation centers with limited resources, up to larger institutions with role-based access structure, the up-scaling aspects of the suitable PETs had to be addressed as well.

Table 2. Excluded privacy design strategies

Excluded PDS	**Reasoning**
DEMONSTRATE INFORM CONTROL ENFORCE	Privacy design strategies that focus on Transparency and Intervenability have largely been excluded from analysis, since these focus on privacy regarding users rather then platform service providers.
MINIMIZE	Requires the minimal data collection and procession to fulfill an intended request with a specific purpose limitation. Since only workflows should be exchanged between users, this strategy could not been used in this instance.
ABSTRACT	The strategy's intention is to limit details by grouping or summarizing data together. Similar to MINIMIZE, this is not applicable in the given use case.

Within the second phase of the selection process, the privacy patterns have been evaluated with existing privacy enhancing technologies (PET). In contrast to privacy design strategies, which are renowned as guidelines in the complete development life cycle, PETs can be seen as solving "only one specific privacy problem in already implemented software" [1]. To enhance applications that gather raw outcome measurements, containing sensitive or personal information, an anonymization function could be used to remove or replace such information prior to distribution. An aggregation function could also be seen as a PET, since e.g. data can be grouped together to provide insight while still managing to mask individuals personal information.

5 Analysis

The following section will provide an overview of which patterns have been implemented within an mHealth application, to increase the privacy aspects in regards to an untrusted platform service provider. Both are designed to enhance the functionality of a remote server sharing concept, where therapists and patients are able to share workflow data securely. Note that the mentioned remote server is configured in a privacy preserving way. To achieve this all logging functionalities have been disabled or limited. For example if a client requests a resource of this server, typically the IP address gets captured in the log file of the web server. In our case this type of logging has been removed. Therefore the server can only store the connection details while the connection is established. After the transmission has ended, no information about what was exchanged can be restored. In regards to data storage, the only identifier that is actually stored within the database along with the encrypted data is the pseudonym of the user, that belongs the data. But since the pseudonyms are generated out of scope of the server (and the service provider respectively), no identifying information (quasi-identifier, or combinations of non-identifiers) can be gained. Table 3 lists three selected concepts together with the privacy design strategy, which their usage represents. As mentioned previously, many novel and experimental cryptographic schemes like *Homomorphic encryption* (HE) and *Polymorphic Encryption and Pseudonymisation* (PEP) exist, but their implementations were not be considered for this use case deployment yet, because only data storage and not the processing thereof is required here. Further details on what these technologies do, or what they imply to improve privacy, will be discussed later on.

Based on the overview provided by Colesky et. al [3], the concepts *Private Link* and *Encryption with user-managed keys* have been chosen. Both have been classified within the selected strategies in the first phase. Furthermore *Attribute-based encryption* has been selected as a holistic approach to secure data sharing within RehaGoal. In contrast to the first two PET's, this concept provides further privacy improvements, not only for platform service providers, but also to RehaGoal users by providing enhanced intervenability due to the role-access structure embedded in ABE schemes. To differentiate the concepts and their

Table 3. Overview of the selected privacy PET's classified within the privacy design strategy terminology by Colesky et. al (*excluded due to the experimental state)

PET	Privacy design strategy
Private Link	HIDE
Encryption with user-managed keys	CONTROL
Attribute-based encryption	SEPARATE, HIDE, CONTROL
Homomorphic encryption*	ABSTRACT, HIDE
Polymorphic Encryption and Pseudonymisation*	SEPARATE, HIDE, CONTROL

intended purpose, they will be sectioned into individual or general deployments for *Private Link* and *Encryption with user-managed keys*. On the other hand *Attribute-based encryption* will be listed within deployments with existing role-based structures.

5.1 For individual or general deployments

A privacy enhancing technology, which has been adopted broadly and can typically be found within applications like cloud storage services and social media platforms, is called *Private Link*. If, for example, a user intends to upload a video to their social media profile, without releasing it to the general public, they could mark it as private or unlisted, such that the platform or service will generate a link to that resource and provide that link to the user. Accessing such will only be possible with knowledge to that particular link. Further examples can be partially found in cloud-based file exchange services (e.g. Nextcloud, Dropbox or Google Drive), where users can selectively share files by generating a unique link.

The privacy aspects of this PET are associated with the *HIDE* strategy, whereas the purpose is to limit the access to only those who received such a link. This PET enables users to selectively disclose a shared resource with a number of recipients, as long as such private link is retained secret.

With a valid link created, the data owner can share the secret with every person he or she intends to. Since this privacy PET is based on a shared secret it has to be transmitted between other users. Figure 5 shows the general life cycle of this PET, where the controller returns a private link, which can then be send to individuals through any communication channel (email, direct message, etc.) to access the resource.

In order to prevent unintended recipients from accessing the resource, a link should not be guessable and therefore be unique or randomly generated. This mechanism can be extended by adding limitations to the private link implementation, e.g. a time constraint for the link validity or that the link only provides a specific version of the resource.

By combining the aforementioned PET with *Encryption with user-managed keys*, access-control can be gained for RehaGoal users to improve intervenability.

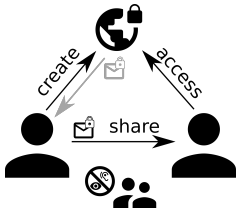

Fig. 5. *Private Link* pattern; access to shared resources is only available through a unique, private address, rather than a public registry or website-index

The basic concept demands that users are responsible to generate and manage their own, strong encryption keys used to encrypt data, prior to storing or transferring it to the data controller. The process consists of three stages, illustrated in Figure 6, beginning with the user generating secret encryption keys, which will, in turn, be used to encrypt any data intended to be stored or transmitted. Once shared with the data controller, these can not access the encrypted information, since the keys are only accessible by the user.

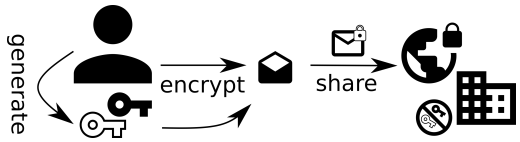

Fig. 6. To increase users authority over who can access data *Encryption with user-managed keys* can be used in conjunction with the *Private Link* pattern to encrypt information prior of sharing it

The combination of the enhancing technologies *Private Link* and *Encryption with user-managed keys* provided enhanced privacy aspects for user data shared with platform service providers in RehaGoal. Figure 7 shows the process of how workflow sharing has been implemented with this solution. The `RehaGoal-GUI` as the applications user interface provides the feature to share workflows remotely. The data of each selected workflow will be processed by the `ExchangeService`, which will then call the `EncryptionService` to generate a random, unique encryption key. This key will be return together with the encrypted workflow data to the ExchangeService. After posting the encrypted workflow data to the `RemoteServer` and receiving a valid URL, the `ExchangeService` will finally return the ExchangeURL with the decryption key appended by a fragment identifier. This ensures that once the URL is clicked, the content behind the fragment identifier will not be transmitted to the `RemoteServer`, but rather be processed within the `RehaGoal-GUI` (or any web-browser) itself.

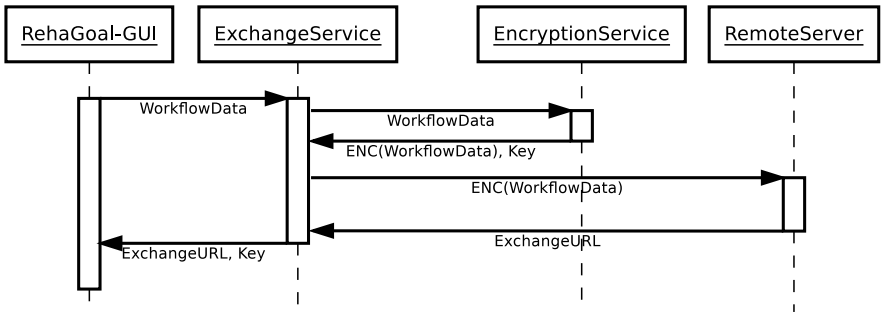

Fig. 7. Sequence diagram of the implemented prototype where data is stored remotely fully encrypted without access to the decryption key for the remote server

5.2 Deployments with existing role-based structures

In the realm of data sharing with multiple users, a functional encryption mechanism should be considered, where data can be encrypted once and decryption is enabled for all users with a secured access structure. Such an encryption scheme, introduced by Sahai et al. [25] in 2005 is called *Attribute-based encryption* (ABE). It was extended into key-policy and ciphertext-policy schemes by Goyal et al. [8] in 2006. While multiple kind of attribute-based encryption schemes exist, they are typically distinguished between key-policy (KP-ABE) and ciphertext-policy (CP-ABE) based schemes. Within the former system, policies are related to the private key of the user, while the attributes are linked to the ciphertext. Therefore retrieval of the plaintext is only possible for users, which fulfill the policy with their private keys. In contrast to KP-ABE, in a ciphertext-policy attribute-based encryption scheme, the user's private key contains the attributes instead of the ciphertext. Vice versa the ciphertext contains the policy, hence user attributes have to fulfill the policy of the ciphertext.

The basic concept consists out of an attribute authority, which provides private keys for each user and public parameters relevant for the encryption. Users can encrypted data using public parameters and specify an access policy within the ciphertext. Such policy consists out of attributes which define a Boolean formula. As an example for RehaGoal, the following attributes have been defined by the attribute authority (among others): `Role Organization, Studies`. If a user wants to encrypt a message for two different subjects: once for all members of the organization A and secondly that only therapists from the organization B should be allowed to decrypt the message. A policy could therefore be represented as follows: `"Organization: A" OR ("Organization: B" AND "Role: Therapist")`. Users not matching those attributes will not have permission to decrypt and access the data. Figure 8 shows the process for both, valid and invalid, user attributes.

This ABE scheme improves data privacy for RehaGoal users in multiple ways. The concept lets users choose which attribute (e.g. user role) is required to decrypt data. Due to the encryption, it also hides any sensitive workflow

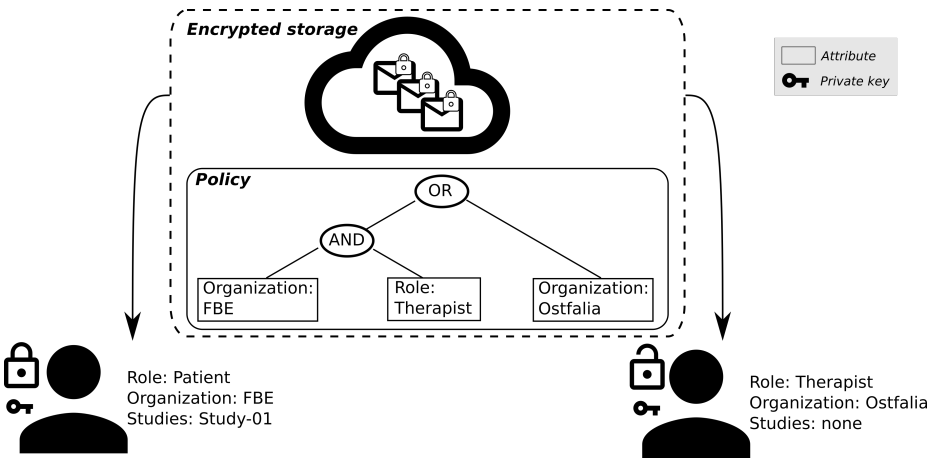

Fig. 8. Basic principle of a ciphertext-policy ABE scheme in RehaGoal, where user attributes are correlated with the policy of an encrypted object. If the policy is fulfilled the plaintext can be decrypted by the user's private key

information towards the platform service provider, similar to *Encryption with user-managed keys*. Furthermore does it provide means to separate data between multiple platform service providers, since Attribute-based encryption can also be deployed on a decentralized basis. Also this solution provides abilities for advanced key management, which is required by Attribute-based encryption schemes. For example, in case authorized therapists do change their occupation, and it is required to revoke the deployed key, mechanisms introduced by Lewko et. al [17] could be used.

6 Discussion

Appropriate data sharing schemes should be chosen, depending on the scale of which data sharing is required. Both proposed solutions provide decent means to improve privacy in mHealth applications. Attribute-based encryption provides a very powerful and feature-rich solution, based on functional encryption by managing a fine-grained access policy. This mitigates both, security and privacy concerns, significantly. Major drawbacks of Attribute-based encryption are the lack of available web-based standard libraries, as well as functional attribute-revocation systems within a ciphertext policy based scheme. Hence, the revocation of a private key (e.g. a patients who left a therapy) will be difficult for distributed and decentralized schemes and the mitigation of that fact should be further investigated.

On the other hand, *Private Link* in conjunction with *Encryption with user-managed keys* improved the privacy with significant less complexity. It provides privacy by being virtually transparent to the RehaGoal user itself. But the principle drawback of the *Private Link* pattern is obvious: once the address (Ex-

changeURL with fragment identifier and content) is publicly available (e.g. listed on a website, registry, index or within a social media posting) the confidentiality of the resource, in regards to who has access to it, is basically lost. The responsibility to manage the encryption keys offsite through the user however, proposes other challenges for a deployment. Since the platform service provider can not provide any key management for those keys, in case the user looses access to them (and has not yet shared the keys), decryption will be prohibited and the content be lost.

In regards to the performance of those solutions, it highly depends on the capabilities of those devices. Without a user intervention study, it is unknown how much the computational overhead will impact users on a daily basis (e.g. battery drain, wait time, etc.). But since these prototypes are intended to evaluate the proposal of a secure data sharing scheme (given the use case), this has so far not been a critical performance indicator. Especially for larger deployments this should be addressed in a follow up evaluation, possibly paired with an user acceptance study.

In case any of these solutions will be used, where the platform service provider does not have access to the encrypted information itself, metadata like file names, timestamps or system information can still be gathered and processed, since these are typically not encrypted, which might posses minor privacy risks.

Another privacy preserving concept, *Homomorphic encryption* where encrypted data can be processed through platform service providers, without a prior decryption, has not been considered. While *Homomorphic encryption* does not fit within the selected use case, since only data storage is required here, it does provide strong privacy implications. Especially for mHealth providers this technology could be used to analyze sensitive information of patients on a larger scale, e.g. to collect trends within groups of patients without ever loosing control over their data. Reference implementations and libraries are already available. *Homomorphic encryption* has also not been considered within RehaGoal, because it is a web-based application which uses web standards and libraries, whereas no reviewed or audited cryptographic libraries for HE have been released at the time of this writing.

In contrast to *Homomorphic encryption*, where reference implementations exists, *Polymorphic Encryption and Pseudonymisation* has also not been selected in this use case. Similar to ABE, the concept of PEP encrypts data in a polymorphic way, which means that the decryption keys could be formed, depending on relevant attributes, which could be managed by the users. This allows for a larger privacy protection gain, but requires a rather complex system, for which currently only experimental libraries exist.

Further subject of evaluation should be a user acceptance study with different UI designs, in order to see if users actually desire to know, or set the encryption key themselves (e.g. with a password-derived function), or if the encryption scheme is even perceived when utilizing small scale deployments.

7 Conclusion

Several solutions, based on privacy patterns for data sharing, have been examined in an mHealth application context. Those solutions have been introduced and conceptualized within this paper, to provide a basis for the implementation of prototypes. Since these prototypes are part of a complex mHealth system, it is not possible to provide the whole source code under an open license, but it will be considered to publish single modules with a focus on secure data sharing.

The implemented approaches, utilizing privacy enhancing technologies, already improved the privacy of the system, by prohibiting access to (meta-) data during data exchange and -storage for service providers and hence allowing more control. Especially usability aspects and specific requirements concerning role-based access control, in the context of mHealth applications, need to be further investigated. Future prototype evaluations should also investigate resource consumption, as well as required computation time, the amount of storage needed for individual transactions and how additional measurements, like padding of the data, would provide protection against re-identification attacks within the server.

References

1. Aljohani, M., Hawkey, K., Blustein, J.: Proposed privacy patterns for privacy preserving healthcare systems in accord with nova scotia's personal health information act. In: Human Aspects of Information Security, Privacy and Trust. vol. 9750, pp. 91–102. Springer (2016). https://doi.org/10.1007/978-3-319-39381-0_9
2. Colesky, M., Hoepman, J.H., Hillen, C.: A critical analysis of privacy design strategies. In: 2016 IEEE Security and Privacy Workshops (SPW). pp. 33–40 (2016). https://doi.org/10.1109/SPW.2016.23
3. Colesky, M., Hoepmann, J.H., Boesch, C., Kargl, F., Henning Kopp, Patrick Mosby, Daniel Le Mètayer, Olha Drozd, José M. del Álamo, Yod Samuel Martin, Julio C. Caiza, Mohit Gupta, Nick Doty: Patterns, `https://privacypatterns.org/patterns/`
4. Dang, N.T., Nguyen, V.S., Le, H.D., Maleszka, M., Tran, M.H.: Sharing secured data on peer-to-peer applications using attribute-based encryption. In: Computational Collective Intelligence. pp. 619–630. Lecture Notes in Computer Science, Springer International Publishing (2020). https://doi.org/10.1007/978-3-030-63007-2
5. European Parliament: General data protection regulation (GDPR), `https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679`
6. Gabel, A., Ertas, F., Pleger, M., Schiering, I., Müller, S.: Privacy-preserving metrics for an mHealth app in the context of neuropsychological studies. In: HEALTHINF. vol. 5, pp. 166–177 (2020). https://doi.org/10.5220/0008982801660177
7. Gentry, C.: Fully homomorphic encryption using ideal lattices. In: Proceedings of the 41st annual ACM symposium on Symposium on theory of computing - STOC '09. p. 169. ACM Press (2009). https://doi.org/10.1145/1536414
8. Goyal, V., Pandey, O., Sahai, A., Waters, B.: Attribute-based encryption for fine-grained access control of encrypted data. In: Proceedings of the 13th ACM conference on Computer and communications security. pp. 89–98. CCS '06, Association for Computing Machinery (2006). https://doi.org/10.1145/1180405.1180418

9. Hansen, M., Jensen, M., Rost, M.: Protection goals for privacy engineering. In: 2015 IEEE Security and Privacy Workshops. pp. 159–166 (2015). https://doi.org/10.1109/SPW.2015.13
10. Hillen, C.: The pseudonym broker privacy pattern in medical data collection. In: IEEE Trustcom/BigDataSE/ISPA. vol. 1, pp. 999–1005 (2015). https://doi.org/10.1109/Trustcom.2015.475
11. Huckvale, K., Prieto, J.T., Tilney, M., Benghozi, P.J., Car, J.: Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Medicine **13**(1), 214 (2015). https://doi.org/10.1186/s12916-015-0444-y
12. Kan, J., Kim, K.S.: MTFS: Merkle-tree-based file system. In: 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC). pp. 43–47 (2019). https://doi.org/10.1109/BLOC.2019.8751389
13. Khanezaei, N., Hanapi, Z.M.: A framework based on RSA and AES encryption algorithms for cloud computing services. In: IEEE Conference on Systems, Process and Control (ICSPC 2014). pp. 58–62 (2014). https://doi.org/10.1109/SPC.2014.7086230
14. Kotz, D.: A threat taxonomy for mHealth privacy. In: Third International Conference on Communication Systems and Networks (COMSNETS 2011). pp. 1–6 (2011). https://doi.org/10.1109/COMSNETS.2011.5716518
15. Larson, R.S.: A path to better-quality mHealth apps. In: JMIR mHealth and uHealth. vol. 6 (2018). https://doi.org/10.2196/10414
16. Lenhard, J., Fritsch, L., Herold, S.: A literature study on privacy patterns research. In: 43rd Euromicro Conference on Software Engineering and Advanced Applications (SEAA). pp. 194–201 (2017). https://doi.org/10.1109/SEAA.2017.28
17. Lewko, A., Sahai, A., Waters, B.: Revocation systems with very small private keys. In: 2010 IEEE Symposium on Security and Privacy. pp. 273–285. IEEE (2010). https://doi.org/10.1109/SP.2010.23
18. Lewko, A., Waters, B.: Decentralizing attribute-based encryption. In: Paterson, K.G. (ed.) Advances in Cryptology – EUROCRYPT 2011, vol. 6632, pp. 568–588. Springer Berlin Heidelberg (2011). https://doi.org/10.1007/978-3-642-20465-4_31
19. Luxton, D.D., McCann, R.A., Bush, N.E., Mishkind, M.C., Reger, G.M.: mHealth for mental health: Integrating smartphone technology in behavioral healthcare. In: Professional Psychology: Research and Practice. vol. 42, pp. 505–512 (2011). https://doi.org/10.1037/a0024485
20. Müller, S.V., Ertas, F., Aust, J., Gabel, A., Schiering, I.: Kann eine mobile anwendung helfen abzuwaschen? Zeitschrift für Neuropsychologie **30**(2), 123–131 (2019). https://doi.org/10.1024/1016-264X/a000256, publisher: Hogrefe AG
21. Müller, S., Katzenbeisser, S., Eckert, C.: Distributed attribute-based encryption. In: Lee, P.J., Cheon, J.H. (eds.) Information Security and Cryptology – ICISC 2008. pp. 20–36. Lecture Notes in Computer Science, Springer (2009). https://doi.org/10.1007/978-3-642-00730-9_2
22. Neubauer, T., Heurix, J.: A methodology for the pseudonymization of medical data. In: International Journal of Medical Informatics. vol. 80, pp. 190–204 (2011). https://doi.org/10.1016/j.ijmedinf.2010.10.016
23. Ometov, A., Bezzateev, S., Mäkitalo, N., Andreev, S., Mikkonen, T., Koucheryavy, Y.: Multi-factor authentication: A survey **2** (2018). https://doi.org/10.3390/cryptography2010001
24. Papageorgiou, A., Strigkos, M., Politou, E., Alepis, E., Solanas, A., Patsakis, C.: Security and privacy analysis of mobile health applications: The

alarming state of practice. In: IEEE Access. vol. 6, pp. 9390–9403 (2018). https://doi.org/10.1109/ACCESS.2018.2799522
25. Sahai, A., Waters, B.: Fuzzy identity-based encryption. In: Advances in Cryptology – EUROCRYPT 2005. pp. 457–473. Lecture Notes in Computer Science, Springer (2005). https://doi.org/10.1007/1142663927
26. Verheul, E.R., Jacobs, B., Meijer, C., Hildebrandt, M., de Ruiter, J.: Polymorphic encryption and pseudonymisation for personalised healthcare. IACR Cryptol. ePrint Arch. **2016**, 411 (2016)
27. Vrhovec, S.L.R.: Challenges of mobile device use in healthcare. In: 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). pp. 1393–1396 (2016). https://doi.org/10.1109/MIPRO.2016.7522357