Using PAPAYA for eHealth – Use Case Analysis and Requirements

Ala Sarah Alaqra
Karlstad University
Karlstad, Sweden
alaaalaq@kau.se

Eleonora Ciceri
MediaClinics Italia
Trento, Italy
e.ciceri@mediaclinics.it

Simone Fischer-Hübner
Karlstad University
Karlstad, Sweden
simofihu@kau.se

Bridget Kane
Karlstad University
Karlstad, Sweden
bridget.kane@kau.se

Marco Mosconi
MediaClinics Italia
Trento, Italy
m.mosconi@mediaclinics.it

Sauro Vicini
MediaClinics Italia
Trento, Italy
s.vicini@mediaclinics.it

***Abstract*—This paper presents an eHealth use case based on a privacy-preserving machine learning platform to detect arrhythmia developed by the PAPAYA project that can run in an untrusted domain. It discusses legal privacy and user requirements that we elicited for this use case from the GDPR and via stakeholder interviews. These include requirements for secure pseudonymisation schemes, for allowing also pseudonymous users to exercise their data subjects rights, for not making diagnostic decisions fully automatically and for assurance guarantees, conformance with specified standards and informing clinicians and patients about the privacy protection. The requirements are not only relevant for our use case but also for other use cases utilising privacy-preserving data analytics to classify medical data.**

***Index Terms*—Privacy Preserving Data Analytics, Arrhythmia Detection, Requirements, GDPR, User Centred Design**

I. INTRODUCTION

The use of machine learning to conduct medical data analytics in eHealth applications has been emerging recently and promise great benefits for medical diagnostics. To protect the patients' privacy, the medical data to be analysed need to be adequately protected, especially if the data analytics is outsourced, for cost reasons for example, to be performed in the cloud. The PAPAYA (Platform for Privacy Preserving Data Analytics) H2020 EU project [1] aims to address the privacy concerns when data analytics are performed by untrusted third-party data processors, such as cloud providers. Consequently, the PAPAYA project designs and develops a platform with dedicated privacy preserving data analytics modules that will enable data owners to extract valuable information from protected (e.g. encrypted) data, while being cost-effective and providing data accuracy. The PAPAYA platform will be piloted for eHealth uses cases that have been defined for the project. However, PAPAYA's project results can only be successfully deployed if its solutions are legally compliant, perceived as privacy-respecting, secure, trusted and usable. To address these design goals, the project has elicited legal privacy requirements as well as user requirements for the project's use cases.

This work was funded by the H2020 Framework of the European Commission under Grant Agreement No. 786767.

This paper has the objective to present: i) PAPAYA's eHealth use case to detect Arrhythmia, involving a privacy-preserving neural network to perform data classification on encrypted data, and ii) the legal privacy and user requirements that we elicited for this use case, via legal analysis and semi-structured interviews with eHealth stakeholders.

The research contributions of this article are the following:

- We present an eHealth use case that utilises a novel Platform for Privacy Preserving Data Analytics, which is currently piloted by the PAPAYA project.
- We show how legal privacy principles of the EU General Data Protection Regulation (GDPR) translate into privacy requirements, including options to implement these requirements for this or similar eHealth use cases based on privacy-preserving data analytics.
- Based on the analysis of interviews that we performed with stakeholders, following a user centered design approach, we elicit user requirements to increase transparency and trust in PAPAYA.

II. ARRHYTHMIA DETECTION USE CASE

A. Motivation

Cardiac arrhythmia is a condition characterised by irregular heartbeat. In most cases, arrhythmia does not represent a serious problem, as it causes either no symptoms or some bearable distress. Nevertheless, some types of arrhythmia may greatly affect patients' health status, leading even to life-threatening situations. Due to its potential severity and its spread across the population, it is of vital importance to monitor patients at risk and detect possible arrhythmia at its onset, so as to preserve their health conditions.

Arrhythmia detection is performed by analysing patients' ECG signals for a period (usually up to 24 or 48 hours) and identifying irregularities in rate, mechanism or duration. Data are acquired with a portable device, which reads the patient's ECG and stores it for later evaluation. Once the observation period is over, a cardiologist can download the data on her computer and perform the analysis.

Unsurprisingly, the manual analysis of a very long ECG signal is infeasible for a cardiologist alone: due to the massive amount of collected information, its manual scrutiny would probably be affected by human error. This is why patients are often instructed to keep a diary of all performed activities and experienced symptoms: in this way, the cardiologist can isolate the areas of the signal that likely contain irregularities, and concentrate on them. Nevertheless, this approach, though minimising the cardiologist's effort and stress, does not give us the certainty of having analysed exhaustively all arrhythmia in the collected ECG signal. To be sure of this, the only solution is to trust machines to perform most of the automatable work, i.e. to go through the whole ECG sequence and classify all beats (via, e.g. a machine learning algorithm) in various arrhythmia classes. The rest of the analysis is left to the cardiologist: she would validate the input provided by the machine and interpret it to formulate a diagnosis.

Unfortunately, even this automatic approach comes with its own limitations. Not all laboratories may have the needed computational power, leaving the staff with a huge amount of unprocessed information on the one hand, and on the other hand the need to fulfil the requests for analysis promptly. Also, laboratory staff might not have the right competences to write an arrhythmia detection algorithm. Thus, it becomes necessary to outsource the analysis to an external environment with more resources and solutions: this would solve the technical issue and guarantee the appropriate performance. In any case, fulfilling this strategy can be done only with the application of appropriate technical and organisational measures to comply with the existing data protection regulations, as it would imply moving patients' personal data from a trusted environment (e.g. the laboratory premises) to an untrusted one (e.g. the public cloud). This need is aggravated by the necessity of outsourcing health-related data, which is classified as a *special category* (GDPR, Article 9), making its processing particularly critical in terms of data protection.

B. Implementation

To demonstrate the effectiveness of the proposed strategy, we designed a service that, being empowered with the computational power made available by external providers, allows patients to easily perform a cardiac check-up in short time and with little organisational effort. As the service is distributed through pharmacies, patients can use it to obtain a report from a cardiologist without going to a hospital. The service is integrated with the PAPAYA platform to perform privacy-preserving machine learning for arrhythmia detection in untrusted environments: in this way, large amounts of (protected) ECG data can be processed in a timely manner and with high performance. To preserve privacy, the PAPAYA platform uses a neural network model for data classification that can be executed over encrypted data by using advanced cryptographic schemes, such as homomorphic encryption or secure multi-party computation [2], [3]. This guarantees that at no point of time the PAPAYA platform or the cloud server hosting it can see any medical data in plain format (i.e. unencrypted).

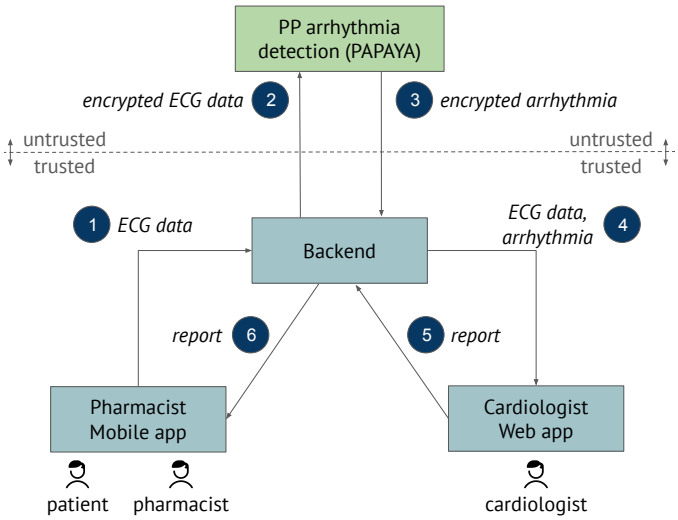

Fig. 1. Service architecture and integration with PAPAYA

The data flow between architectural components is depicted in Figure 1. At first, the pharmacist registers the patient to the service via the mobile application (installed on a tablet), and provides her with an ECG monitoring device that will record ECG data for the following 24 hours. When the monitoring period is over, the patient returns the device and the pharmacist sends the acquired ECG data to the service backend (step 1), which is then encrypted and outsourced for analysis (step 2). At this point, the privacy preserving arrhythmia detection tool running on the PAPAYA platform in the untrusted domain analyses the encrypted ECG data and identifies the ECG areas with irregular heartbeat, deriving encrypted arrhythmia analysis results. Then, it sends the results back to the service backend (step 3). The arrhythmia analysis results are decrypted and sent (together with the cleartext ECG data) to the cardiologist Web app (step 4), with no patient identifiers, so as to allow the cardiologist to visualise them and create a report accordingly. Once the cardiologist finalises her work, the produced report is returned to the backend (step 5) and finally forwarded to the pharmacist mobile application. The pharmacist is notified about the availability of the report, and consequently shares it with the patient.

The described data flow involves several stakeholders, each one playing a role in the GDPR perspective. Specifically: i) the patient is the data subject; ii) the service provider (i.e. the one that distributes distributing the service via pharmacies) and the pharmacist are joint data controllers; iii) all the remaining actors (i.e. the cardiologist and the external cloud provider running the PAPAYA platform) act as data processor.

III. REQUIREMENTS

The elicitation of legal privacy requirements is based on a legal analyses of GDPR, complementing reports and commendatory by the Art. 29 Data Protection Working Party, and

the Fundamental Rights Agency of the European Commission. For eliciting user requirements for the use case, we followed a user-centred approach by involving stakeholders in eHealth via semi-structured interviews, to consider their perspectives and understandings in regard to our use case. Ethical approval for conducting the interviews was provided by an ethics advisor at Karlstad University. Information about initial requirement elicitation can be found in [4]. In the following subsections, we provide an overview of important legal privacy and user requirements that were elicited for the PAPAYA platform and our use case. A selection of key use case specific requirements are coded with PRQ (Privacy Requirements) or URQ (User Requirements) and a number, followed by description, acceptance criteria, while other more general privacy requirements are only briefly summarised due to space limitation.

A. Legal requirements

This section briefly presents and discusses the essential privacy requirements pursuant to the GDPR for the use case, with a focus on general data processing principles, and principles for protecting transparency and control for the data subjects. Moreover, we discuss how those requirements can be implemented by organisational and technical means to achieve data protection by design pursuant to Art. 25 GDPR.

The material scope of the GDPR is restricted to the processing of personal data, where personal data are defined in Art. 4 (2) as any information relating to an identified or - directly or indirectly - identifiable natural person ('data subject'). The GDPR further defines 'pseudonymisation' in Art. 4 (5) as "the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person". Data encryption fulfils the definition of pseudonymisation. This also means that processing of encrypted ECG signal data by the PAPAYA platform must be regarded as personal data processing, for which the European Data Protection legislation needs to be applied. Not only the ECG signal data collected by the pharmacy and sent to the PAPAYA platform in encrypted form for data analysis fall under the scope of the GDPR, but also any derived metadata, particularly the analysis report about arrhythmia classification derived by PAPAYA's machine learning.

ECG signals are uniquely identifying biometric "fingerprints" [5], and thus ECG signals cannot be anonymised by simply redacting the direct identifiers of the patient. Moreover, as medical data, and thus special categories of data pursuant to Art. 9 GDPR, ECG data require special protection.

1) General privacy principles: The general privacy principles specified in Art 5 GDPR define the basic privacy requirements that PAPAYA needs to fulfil, which include the data processing requirements in Art 5 (I) for: (a) lawfulness, fairness and transparency; (b) purpose specification and limitation; (c) data minimisation; (d) data accuracy; (e) data storage limitation; (f) data integrity and confidentiality; as well as the requirement for accountability specified in Art. 5 (II).

It should be noted that PAPAYA's privacy-preserving data analytics are already enforcing data minimisation via encryption and thus pseudonymisation, and that the deployed homomorphic cryptographic schemes will not impact data quality negatively [2]. Nonetheless, for allowing the service provider at the backend to link the raw ECG signal data of a patient with the corresponding arrhythmia analysis results, another layer of pseudonymisation is needed by a scheme that is replacing any identifying patient (anagraphic) data by pseudonyms that will be attached to the ECG signal data and the derived arrhythmia analysis results. This pseudonymisation must be implemented either by the pharmacy or by the service provider and must then allow the pharmacy or service provider to securely link the pseudonymous diagnosis report and any other medical data back to the patient for forwarding the final report that the pharmacy receives in step 6 to the corresponding patient. If personally identifying data is already replaced by the pharmacy with pseudonyms, a higher degree of data minimisation can be reached, as in this case real patient identities can also be hidden from the service provider and are only known by the pharmacy.

[**PRQ1:** Secure Pseudonymisation]
Description: One of the joint data controllers, preferably the pharmacy, MUST implement a secure pseudonymisation scheme for the patient's medical data.
Acceptance criteria: Best practices for secure pseudonymisations (e.g. based on secure hash functions) MUST be followed (as provided by ENISA's report on guidelines for secure pseudonymisation [6]).

2) Lawfulness, consent and ex ante transparency: Personal data must be processed lawfully, which means that at least one of the following legal grounds of Art. 6 applies. For our use case, consent by the data subject serves as the legal basis, which pursuant to Art 4 (11) GDPR needs to be freely given, informed, specific, and confirmed by an affirmative action. Moreover, the consent must, as a legal basis for processing of medical data, i.e. special categories of data, be explicit, which requires a very clear and specific statement of consent, e.g. by a written confirmation.

To implement an informed consent, privacy policy information should be provided to the data subject in a "concise and transparent" (Art. 12 GDPR) manner for avoiding information fatigue. The Art. 29 Data Protection Working Party has for this reason suggested to use multi-layered privacy notices [7], which enable data subjects to easily navigate to the particular section of the privacy notice that they are interested to read. In addition to policy information required by Art. 13 GDPR (incl. identity of the data controller and data processing purposes), the first (top) layer of such a layered policy notice should contain any other information on the processing which has the most impact on the data subjects and that enables them to understand the consequences, while further layers could provide additional information about the privacy-preserving technical measures, as we will discuss further below (see

URQ5).

3) Data subject rights: Legal requirements can be derived for data subject rights for access pursuant to Art. 15 (ex post transparency) and for intervenability pursuant to Art. 7, 16-24, allowing data subjects to "intervene" with the data processing by requesting to correct, block, delete or to export their data or to object to the data processing. These rights apply for data that are processed either by one of the (joint) controllers directly or by the PAPAYA platform taking the role of a data processor, and it is the obligation of the controllers to enforce the rights for the data subjects. Most data subject rights apply not only for the data that the data subjects have disclosed to the data controller, but also for data that have been derived from that data, e.g. via machine learning. An exception is the right to data portability that only applies for data that the data subject provided directly to the data controller, and not for data that have been derived/inferred from those data by any data processing on the PAPAYA platform. Nonetheless, the patient could request to obtain an electronic copy of all her derived data, including the arrhythmia analysis results from the service provider derived by the PAPAYA platform, by exercising her right to access pursuant to Art. 15 (III).

According to Art. 11 (II) GDPR, even if the data controller cannot identify the data subject, the data subject can still exercise her data subject rights if she provides additional information enabling her identification. This can be relevant for the case where the pharmacy pseudonymises the ECG signal data before they are sent to the service provider in step 1. In this case, the service provider as the data controller may not be able to directly identify the patients to whom the medical information processed by the service provider belongs. Nonetheless, the patient could also in this case exercise her data subject rights with the service provider by proving that she is the holder of a specific pseudonym, which she could for instance also prove with zero knowledge (i.e. without revealing any additional identifying information) to the service provider, e.g. by using attribute based credentials [8] issued by the pharmacy.

[**PRQ2:** Answering data subject rights request for pseudonymous users]
Description: Pseudonymous patients MUST be able to securely exercise their data subject rights with the service provider without revealing any identifying data.
Acceptance criteria: Secure pseudonymous authentication schemes are implemented.

For our use case, in particular the right of not being subject to a decision based solely on automated processing pursuant to Art. 22 (I) GDPR, including profiling, can be of relevance, because the PAPAYA platform automatically performs data classification. Exceptions from this right include the authorisation by explicit consent. This right, however, only applies for decisions based solely on automated processing, i.e. without any human intervention. It does, therefore, not apply for decisions that are not based solely on PAPAYA's automated data analysis, such as in our use case, as long as the final diagnostic decisions will be done by the medical doctor based on PAPAYA's analysis results and other information. According to [9], Article 22(1) establishes not only a data subject right, but also a general prohibition for decision-making based solely on automated processing which significantly affects the data subject, which applies whether or not the data subjects take an action to actively invoke this right regarding the processing of their personal data. Therefore, we require:

[**PRQ3:** No fully automated diagnosis decisions]
Description: The diagnosis decisions SHOULD be done by a medical doctors based on both the raw ECG signal data and the automatically generated arrhythmia analysis result from the PAPAYA platform.
Acceptance criteria: The doctor MUST review the patient's ECG data and the automatically generated arrhythmia analysis result in combination for deriving a diagnosis report, unless the patient has explicitly consented to an automatically generated diagnosis report.

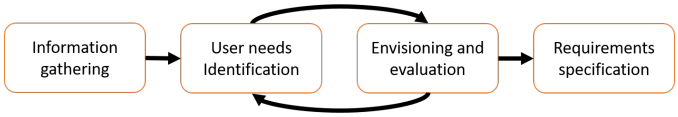

Fig. 2. Requirements analysis processes [13]

B. User requirements

In the process of software development, eliciting user requirements is important and plays a key role in addressing usability functions [10]. In the PAPAYA project, we have followed a user centred design approach in order to design and develop our tools with usability in mind. User centred design (UCD) is multidisciplinary and collaboration is important for the development of usable systems [11], [12]. UCD is growing in practice and is shown to improve usability and product usefulness. In order to understand the needs and requirements of users, a thorough analysis must take place. In our work, we follow the four stages by Maguire and Bevan of requirements analysis as shown in Figure 2 [13]. Details are described in the following subsections corresponding these stages.

1) Information gathering: The first step in our requirement's elicitation and information gathering is the identification of stakeholders involved in the use case. We selected stakeholders who may have an influence on applications and systems used for health data analysis, who can take the role of the cardiologist in the use case, and/or can provide feedback about the patients' expectations. The stakeholders were medical professionals, eHealth researchers as well as technical experts in eHealth.

In total we interviewed 14 stakeholders, who were in Australia (2), Italy (2), Ireland (4), Sweden (4) and United Kingdom (2). Their expertise varied between medical (5), research (3), technical (1), combination of medical and research (3), and combination of research and technical (2). They all had 3+ years of experience in their fields. When selecting stakeholders for our study, we had their knowledge, experience, responsibilities and abilities in mind. According to

the use case, we sampled stakeholders who are familiar with ECG tests, and are knowledgeable of analysis that is done on ECG data. With their consent, interviews were audio-recorded and notes were taken.

2) User needs identification: We collected data about user needs through our empirical approach of semi-structured interviews. We questioned our stakeholders' opinions, needs, and concerns using the ECG use case. Having the use case in focus provides a realistic example for the stakeholders to voice their needs and concerns as well as clarify uncertainty. In addition, user requirements relating to the functionality of the system are provided according to the specific use case.

3) Envisioning and evaluation: This stage was following the data collection of the interviews, and the specification of user needs. We conducted workshops with Human Computer Interaction (HCI), privacy and security, and eHeatlh experts for analysing the gathered data. In these workshops, the evaluation, discussions, and brainstorming resulted in guidelines and requirements of the use case.

4) User Requirements specification: User requirements were derived and elicited based on the analysis and the evaluation of our interview results. In the following paragraphs we introduce our elicited user requirements with regard to key results from our user study. With the exception of **URQ5**, requirements are defined as optional and non-functional requirements to be implemented in production. We define point **URQ5**, on informing patients, as a mandatory requirement.

a) Sensitivity of ECG signals and the need for protection: ECG signals are considered by our interviewees as sensitive data and need to be protected during analysis, especially if linked to individuals with pseudonyms. While some might argue that since the patient identifiers are usually separate from the ECG signal data, and that the ECG signal alone is not identifiable; most of the interviewees regard the ECG to be sensitive requiring strong protection, and even likened the ECG to biometric data such as a fingerprint. Therefore:

[**URQ1:** Communicating outsourced data encryption]
Description: It SHOULD be communicated to stakeholders that outsourced data to the PAPAYA platform are sufficiently protected.
Acceptance criteria: Information about data protection SHOULD be available by introductory tutorials, consent forms, or contact expert at the organisation.

b) Trust in PAPAYA*'s analysis on encrypted data:* Many interviewees, especially those with some basic technical knowledge, are skeptical on how the data analysis on encrypted data is plausible, they highlight the need for testing and certifications in order to trust PAPAYA. Users and professionals should be assured that PAPAYA can conduct a data analysis on encrypted data as claimed. The solutions should be successfully tested, validated and certified by recognised authorities for providing assurance certifications. Moreover, a couple of participants mentioned that research publications proving the soundness of PAPAYA's analytical methods could be made available to interested stakeholders.

[**URQ2:** Assurance guarantees]
Description: Assurance guarantees confirming that analysis on encrypted data on the PAPAYA platform works as stated SHOULD be provided to doctors and other stakeholders using or working with the PAPAYA platform
Acceptance criteria: The privacy preserving data analytics modules of the PAPAYA platform SHOULD be tested, validated, and certified. Assurance certification proofs by recognised authority, and /or reports on validated research study SHOULD be made available to stakeholders.
[**URQ3:** Conformance of PAPAYA to specified standards]
Description: The PAPAYA platform SHOULD demonstrate conformance to standards
Acceptance criteria: Conformance testing of the platform's privacy preserving data analytics modules MUST be conducted and test reports SHOULD be satisfactory and available to stakeholders.

c) Clinicians perspective on information about privacy protection: In our interviews, we queried clinicians and medical staff about the level of information they would like to know about privacy protection. Doctors do not necessarily want to be experts in encryption, or understand the process fully. However, most stated that it is important that they have a general level of understanding that satisfies them. They also needed to know that the privacy-preserving data analytics process on encrypted data is safe, of clinical value, and is reliable. Clinical staff want to know enough about the privacy-preserving process so that they are able to answer questions from patients, even though questions by patients may be rare.

[**URQ4:** Informing clinicians about privacy protection]
Description: Clinicians SHOULD receive basic information on PAPAYA in regard to the technical privacy protection, data quality guarantees, and about experts who could be contacted for details.
Acceptance criteria: User-friendly information SHOULD exist in a form of tutorials, leaflets. Contact information of an expert SHOULD also be available.

d) Informing patients: When communicating and informing patients about their data protection and technical details, most interviewees thought that it is necessary to have information available to patients. Doubts on whether patients will ask or understand the information were raised by clinicians. Patients need to be informed that their data are processed in compliance with GDPR and that their data that are outsourced to the PAPAYA platform are fully protected by technology. Thus, patients should be given at least access to some information about the technical privacy protection measures employed so that they can understand the consequences of outsourcing the data analysis to the PAPAYA platform. This information could be made accessible via layered policy notices mentioned in section III-A2 that allow interested users to easily navigate to that information from a top policy layer.

[**URQ5:** Informing patients about privacy protection]
Description: Information about technical privacy protection of the PAPAYA platform SHOULD be available to patients.
Acceptance criteria: Usable layered policy notices (including

those that appear as part of consent forms) SHOULD provide this information and/or information leaflets SHOULD be available on demand.

e) Transparency of privacy impacts and utility trade-off: From our results, several participants praised the notion of having a privacy impact assessment (PIA) for the use of the PAPAYA platform conducted by the service provider. They noted that the service provider must also show reports containing details of the PIA method, process, and the qualifications of the evaluator. In addition, the evaluator must be an independent expert not associated with the service provider. Independent expert PIA can be expected to help build trust in the claims made by the service provider. Privacy and utility benefits, and trade-offs, should be clearly communicated to users and stakeholders as well.

[**URQ6**: Communicating privacy impacts, benefits and utility trade-offs]
Description: The Privacy Impact Assessment (PIA) results SHOULD show privacy benefits and utility trade-offs of the PAPAYA platform. Information about the evaluation, evaluators, costs, and other factors SHOULD be be made available
Acceptance criteria: A PIA SHOULD be conducted by an independent qualified expert. Detailed information about the PIA results and reports SHOULD be made available by the user interface to different stakeholders (including clinicians and patients), e.g. via layered privacy notices, or by other means.

IV. Related work

Remote cardiac monitoring in hospitals or in telemedicine contexts has been discussed in the literature for some years now. While most of the works offer plain processing of ECG data based on the most common parameters (e.g. RR intervals) [14], [15], some others apply modern techniques based on machine learning to conduct more structured analyses [16]. Unfortunately, as often happens in the state of the art related to the health sector, although health-related data is considered as one of the most sensitive categories of data, scientists confront the analysis challenges first, and then deal with the privacy concerns. This is why there is a surge of studies on ECG analysis techniques, and a lack of application of privacy-enhancing technologies to them. Some of the works in the literature actually already apply encryption prior to analysis when data is outsourced to untrusted domains [17], but the pool of available literature considering privacy thins when machine learning falls in the picture. In [18], the authors apply full homomorphic encryption on ECG data for analysis in a public cloud, however they are not describing a use case or any legal privacy or user requirements. User perceptions and requirements for other types of novel privacy-enhancing eHealth use cases have for instance been discussed by us in [19], which similar to this study also show that users with more technical expertise also require information about assurance guarantees to trust the claimed privacy enhancing properties.

V. Conclusions

The presented eHealth use case is currently piloted and will be validated to meet the elicited requirements by the PAPAYA project. We consider the legal privacy and user requirements as important for achieving privacy by design and establishing trust in PAPAYA. Moreover, they can also guide the development, already in earlier stages, of similar eHealth applications which utilise privacy preserving machine learning for medical diagnostics in the future.

References

[1] PAPAYA, "Platform for privacy preserving data analytics," https://www.papaya-project.eu/, eU H2020 project(Accessed on 11/30/2019).
[2] E. Ciceri, M. Mosconi, M. Önen, and O. Ermis, "Papaya: A platform for privacy preserving data analytics," 2019.
[3] Papaya Consortium A, "D3.1 Preliminary Design of Privacy Preserving Data Analytics," 2019.
[4] Papaya Consortium B, "D2.2 Requirements specification," 2019.
[5] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold, "Ecg to identify individuals," *Pattern recognition*, vol. 38, no. 1, pp. 133–142, 2005.
[6] European Information Security Agency ENISA, "Pseudonymisation techniques and best practices," https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices, 2019.
[7] Art. 29 Data Protection Working Party, "Guidelines on transparency under Regulation 2016/679," 2016.
[8] K. Rannenberg, J. Camenisch, and A. Sabouri, "Attribute-based credentials for trust," *Identity in the Information Society, Springer*, 2015.
[9] C. Giakoumopoulos, G. BUTTARELLI, and M. O'FLAHERTY, "Handbook on European data protection law - European Union Agency for Fundamental Rights and Council of Europe, Luxembourg," 2018.
[10] G. J. Browne and M. B. Rogich, "An empirical investigation of user requirements elicitation: Comparing the effectiveness of prompting techniques," *Journal of Management Information Systems*, vol. 17, no. 4, pp. 223–249, 2001.
[11] J. Gulliksen, B. Göransson, I. Boivie, S. Blomkvist, J. Persson, and Å. Cajander, "Key principles for user-centred systems design," *Behaviour and Information Technology*, vol. 22, no. 6, pp. 397–409, 2003.
[12] J.-Y. Mao, K. Vredenburg, P. W. Smith, and T. Carey, "The state of user-centered design practice," *Communications of the ACM*, vol. 48, no. 3, pp. 105–109, 2005.
[13] M. Maguire and N. Bevan, "User requirements analysis," in *IFIP World Computer Congress, TC 13*. Springer, 2002, pp. 133–148.
[14] A. Pantelopoulos and N. G. Bourbakis, "A survey on wearable sensor-based systems for health monitoring and prognosis," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 1, pp. 1–12, 2009.
[15] A. L. Bleda, R. Maestre, J. Corral, and R. Ruiz, "A quality and ergonomic heart monitoring device with user-friendly app for telemedicine," in *Multidisciplinary Digital Publishing Institute Proceedings*, vol. 31, no. 1, 2019, p. 67.
[16] C. Roopa and B. Harish, "A survey on various machine learning approaches for ecg analysis," *International Journal of Computer Applications*, vol. 163, no. 9, pp. 25–33, 2017.
[17] A. Page, O. Kocabas, T. Soyata, M. Aktas, and J.-P. Couderc, "Cloud-based privacy-preserving remote ecg monitoring and surveillance," *Annals of Noninvasive Electrocardiology*, vol. 20, no. 4, pp. 328–337, 2015.
[18] O. Kocabas and T. Soyata, "Utilizing homomorphic encryption to implement secure and private medical cloud computing," in *2015 IEEE 8th International Conference on Cloud Computing*. IEEE, 2015, pp. 540–547.
[19] A. S. Alaqra, S. Fischer-Hübner, and E. Framner, "Enhancing privacy controls for patients via a selective authentic electronic health record exchange service: qualitative study of perspectives by medical professionals and patients," *Journal of medical Internet research*, vol. 20, no. 12, p. e10954, 2018.