City Research Online

City St George's, University of London

Citation: Gorp, P. V., Comuzzi, M., Fialho, A. S. & Kaymak, U. (2012). Addressing health information privacy with a novel cloud-based PHR system architecture. Paper presented at the 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 14-10-2012 - 17-10-2012, Seoul, Korea. doi: 10.1109/ICSMC.2012.6378006

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/4083/

Link to published version: https://doi.org/10.1109/ICSMC.2012.6378006

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Addressing Health Information Privacy with a novel Cloud-Based PHR System Architecture

Pieter Van Gorp*, Marco Comuzzi*, André Fialho† and Uzay Kaymak*
* Eindhoven University of Technology, Eindhoven, The Netherlands, {p.m.e.v.gorp, m.comuzzi, u.kaymak}@tue.nl
†Technical University of Lisbon, Lisbon, Portugal, afialho@mit.edu

***Abstract*—Patient Health Records (PHRs) shift the ownership of health data from health providers to patients. Such a shift poses important challenges from the data privacy point of view. Patients would like to be able to selectively reveal information to other stakeholders and, at the same time, be assured that their health information will not be used improperly once shared. Current PHR systems partially fail to satisfy these requirements. In this paper, we show that both requirements can be satisfied fully when adopting a novel cloud-based PHR system architecture. We expain the role of remote virtual machines in this architecture and use interaction models to reason about privacy implications. Finally, we evaluate MyPHRMachines, a prototypical implementation of the architecture: we demonstrate that the system enables the execution of third party genome analysis services on patient-owned genome data while ensuring that (1) such services cannot maliciously store this data and (2) patients can show the analysis results to experts without sharing along their full genome.**

***Index Terms*—Patient Health Record, Privacy, Security, Cloud Computing.**

I. Introduction

Patient Health Records (PHRs) are *a set of computer tools that allow people to access and coordinate their lifelong health information and make appropriate parts of it available to those who need it* [9]. PHR systems differ from traditional Electronic Health Record (EHR) systems, being centered around the patient rather than delivered by individual caregivers.

PHR systems shift the ownership of health information from caregivers to patients, giving support to health information collection, sharing, change, and self-management [8], [18]. PHR systems are likely to decrease the cost of patient information management and, at the same time, increase quality of care, allowing patients to reproduce their complete and accurate medical history in a computerized repository when requested or needed [1], [13]. PHR systems can also empower patients to seek care from alternative caregivers and to engage into emerging social platforms bringing together patients with similar conditions [5]. Finally, PHR systems can empower intelligent assistance services since PHRs are more integrated and longitudinal than conventional patient records.

One of the main barriers to the adoption of PHR systems is posed by privacy and security issues. Patients and, more generally, citizens, show great concern about the privacy and security of their health records [10]. This concern becomes particularly important when patients provide their personal data to public or private institutions providing data analysis or diagnosis services. In this scenario, patients have very few or literally no actual guarantees on how providers of such services will use their personal data. Although personal health data should be strictly used only to satisfy the patients' requests, there are multiple example cases in which such information has been used for different scopes and used for commercial purpose [2].

More generally, PHR systems privacy and security are characterized by a paradox. On the one hand, in fact, digitization of information should decrease the risk that information can be lost or improperly modified. On the other hand, however, the security provided by paper-based records, which are available only at a small number of physical locations, greatly reduces the risk of multi-site unauthorized access entailed by health records available on the Internet [16]. Furthermore, digitized patient health information can be easily duplicated and improperly used by the same institutions with which the patients share information.

This paper focuses on privacy problems related to using PHR systems to support information sharing among patients and caregivers. In particular, we first classify the conceptual and technical issues about PHR privacy of those systems that are currently available on the market. Then, we introduce an innovative PHR system whose design alleviates many of these issues. We focus specifically on allowing patients to selectively disclose information to caregivers, preventing them to misuse the PHR data that the patients share. The PHR system we consider is MyPHRMachines, developed by the authors, which is presented more extensively in [19].

The paper is organized as follows. Section II discusses related work on PHR systems. The analysis of today's security-related vulnerabilities and our proposed solution are discussed in Section III and IV, respectively. The solution is evaluated in the concrete context of secure genomic data analyses. Conclusions and future work are eventually presented in Section VI.

II. Related Work

PHR systems can be classified into *free standing* and *provider-tethered*. The latter are spin-offs of EHR systems in use by specific institutions, while the former are developed independently by a software vendor. Free standing PHR systems promote true patient ownership of health data and pose the challenge of integration with the information systems of several heterogenous health institutions and caregivers, e.g. hospitals, GPs, private clinics, or insurers. Provider-tethered PHRs, on the one hand, promote a *stickier* relationship between the provider and the patient. On the other hand, however, they

do not address the *continuity of care* requirement envisioned for PHRs [8]: an interoperability problem still arises when the patient seeks care from a caregiver outside of the network of the provider by which the PHR system is tethered.

According to [4] in free-standing PHR systems all data have to be (re-)entered manually by patients. When adhering to this strict interpretation, a third category of PHR systems arises - that of so-called *integrated* PHR systems. Such systems integrate EHR data from multiple providers automatically in a comprehensive record. We observe that this can be realized by keeping the data inside the backend provider systems or by replicating it in a free-standing repository. From a security and privacy perspective, the location and ownership of the data is the main issue to be taken into account. Therefore, we only differentiate between *free standing* and *provider-tethered* approaches, assuming that (1) free-standing approaches can always be integrated with provider systems to avoid manual re-entry of data and (2) multiple providers can always provide one PHR as thin extensions to their EHRs.

From the privacy point of view, free standing PHRs systems (e.g., MyHealtheVet and HealthVault) pose a more prominent threat than tethered approaches. Organizations providing free standig PHR systems, in fact, will have access to a sizable mount of personal health information of different patients, which can have economic value for insurance and pharmaceutical companies. Moreover, such organizations may not be subject to privacy and security regulations such as the HIPAA, which normally apply only to healthcare providers, e.g. hospitals and private clinics [3].

From the IT architecture point of view, PHR systems can be classified into *Web-based* and *offline* PHRs. Web-based PHR systems adopt an architecture storing health information in a database and making it available to patients through a Web application [15]. Offline PHRs use a physical support such as smart cards or USB keys [14] to store health information. Security and privacy of patient-owned health information in a PHR system have been addressed mostly from a technical perspective. As far as security is concerned, the literature suggests traditional solutions developed for Web application to strengthen the confidentiality, integrity, and availability of Web-based PHR systems [21]. Offline PHR systems, in this context, show the same security-related features of paper-based records. They reduce the risk of unauthorized access by maintaining health information in one physical location and by supporting enhanced authentication mechanisms. At the same time, however, they can be stolen, misplaced, or simply lost.

III. Security and Privacy related vulnerabilities of PHR systems

In this section we discuss the requirements for health information privacy that, according to our analysis, are not satisfied by current PHR systems. In the next section, we will demonstrate how such requirements could be satisfied, under specific circumstances, by our MyPHRMachines system, in the scenario of health information sharing for digital diagnosis services.

Information privacy can be defined as the ability of an individual or a group to seclude information selectively and thereby reveal or made available such information selectively. Privacy has become an important concern in the modern digital world, since the amount of information about a certain object that can be collected by either an individual or an institution has dramatically increased with the advent of the Internet [17]. An information privacy issue occurs when information is disclosed to or improperly accessed by a third-party without the consent of the owner of the information.

For several reasons, in the healthcare domain, patients may not want for their medical health records to be revealed to others. Healthcare records may affect the patients' insurance coverages or employment; patients may consider inappropriate or even embarrassing for others to know about their personal psychological conditions or treatments; they may not want for their medical records to be used by governmental agencies or third-party companies for commercial purposes as a source of profit; or, simply, they may not want others to know particular details of their own lives. Patient health information privacy is regulated by the law in most countries through the definition of physician-patient privileges, e.g. the HIPAA act in the United States.

PHR systems shift the ownership of patient health information from health institutions to patients [8], [18]. In this paper, we do not focus on the legal implications of such an ownership shift, but we rather focus on how the shift of the ownership of health information impacts the technical requirements of modern PHR systems. Specifically, the above discussion about introduces two requirements for PHR systems:

- **R1** *PHR systems should allow patients to selectively reveal information to other stakeholders, such as GPs, hospitals, insurers, or other interested parties*;
- **R2** *PHR systems should guarantee that, once shared with a stakeholder, health information of a patient cannot be improperly used by the stakeholder*.

The requirement R1 is not explicitly satisfied by current Web-based PHR systems. These systems, in fact, only allow patients to collect and store digitized health information, but they usually do not include any design mechanism to selectively delegate access. Only very few commercial systems implement simple role-based access control mechanisms on PHR data. PeopleChart [1], for instance, allows to separate private and public health information and it gives the opportunity to specify different roles, e.g. provider or caregiver, to access the information labelled as public.

The requirement R2 cannot be satisfied alone by the design of the PHR system, since it requires also the collaboration of the stakeholder with whom information is shared. In this paper, however, we will demonstrate that under specific circumstances, the design of the PHR system can force the satisfaction of such a requirement.

[1] http://www.peoplechart.com

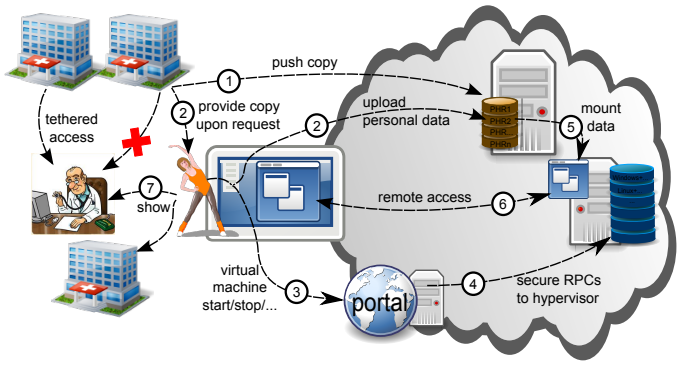

Fig. 1. Remote Virtual Machines supporting PHR management in the cloud.

IV. A Cloud-based PHR architecture

Subsection IV-A provides an introduction to the MyPHRMachines architecture. Subsection IV-A provides a more rigorous description of the key system interaction scenarios relevant for assuring the privacy of PHR data.

A. The architecture of MyPHRMachines

Fig. 1 visualizes the architecture of MyPHRMachines. The top-left end of the figure shows care organizations generating health-related data, which may become relevant also for other caregivers. In particular cases, the data is already offered by portals. For example, various hospitals provide portal-based EHR or PACS access to GPs in their region. The proposed architecture applies particularly to those domains where the data should transcend regional and temporal constraints: the idea is to copy patient data from isolated care systems to one or more central, trusted data repository. Data entry in the repository can be done directly when the care institutions push their data to the repository (cfr., the edge labeled as ① in Fig. 1), or indirectly if they provide a copy to the patient, who can then upload his/her data to the repository (cfr., the two edges labeled as ② in Fig. 1). Note that the repository only supports the uploading of raw data. More advanced repository functionality is conceivable but not needed here.

We model the repository functionality as two operations of the *Data Repository* entity in the conceptual model shown in Fig. 2. One operation stores *PHR data* while the other retrieves that. The fact that this data is raw is modeled by making it a subclass of entity *BLOB* (Binary Large OBject). Note that both operations should be parameterized with patient credentials but we abstract of such details since they are quite trivial and would clutter our diagrams.

Edge ③ in Fig. 1 models a patient who interacts with the MyPHRMachines portal. Via this portal, the patient can select, start and stop remote Virtual Machines (VMs) to which the PHR data will be mounted securely. By default, MyPHRMachines blocks all traffic from a virtual machine to the Internet; thus by default patients are assured that their sensitive data cannot be transferred to other Internet locations during the VM execution. Moreover, the remote procedure calls for starting and stopping remote virtual machines (cfr., ④ in Fig. 1) are assumed to leverage state-of-the-art security mechanisms to ensure that they cannot be executed maliciously by other Internet users. Edge ⑤ in Fig. 1 models that the PHR data of a patient is mounted to his/her virtual machine. The supportive infrastructure should ensure that data is only mounted to VMs that have been started by the data owner. Edge ⑥ in Fig. 1 models remote patient access to his/her virtual machine. Remote access should be ubiquitous: patients can use a desktop web browser, a smartphone application, or any other device with appropriate internet access and display capabilities. Virtual machines can contain advanced decision support software, specialized medical viewers, data transformation software, etc. The execution of application software is not constrained by the caregiver device capabilities, since it occurs remotely in virtual machines.

Edge ⑦ in Fig. 1 models that patients can also show their PHR to caregivers, e.g. their GP. In this case, caregivers receive a unique URL that gives browser access to a running VM as long as the patient keeps it running. In principle, the patient can grant access to a specific VM where only selected PHR data have been mounted. This guarantees that patients can selectively reveal information to other stakeholders. Patients can expect the caregiver to consider his/her PHR data since no untrusted software needs to be executed on the caregiver's machine. The quality of care provided by caregivers increases, since (1) they can take into account the full patient history, and (2) they can get service-based access to decision support tools that their organization has not invested in.

MyPHRMachines does not enable caregivers to download data from a shared virtual machine. This ensures that once patients shut down the virtual machine, caregivers no longer have access to the PHR data. As soon as data would be offered for download, this could no longer be ensured formally. This control feature is a unique architectural strength from a patient privacy point of view. Additionally, the remote virtual machines can monitor the complete caregiver session. This enables real-time or a posteriori checks against malicious usage patterns. In this context, the proposed architecture respects the caregiver privacy since only the interactions with software within the remote virtual machines can be monitored.

B. System Interaction Models

The UML sequence diagram shown in Fig. 3 models the system behavior when a patient starts a new virtual machine: after retrieving PHR data from a trusted PHR data repository, the system retrieves an image of the virtual machine. Virtual machine images are stored in a so-called *VM Image Library*. We model this library as a separate entity since it enables scenarios with multiple distinct image libraries (e.g., one per insurance company, country, etc.). Once MyPHRMachines has retrieved both the patient data and the virtual machine image, it can start a virtual machine with the right data and the right software. The architecture preserves patient privacy since maintenance of images in the library can be performed also without mounting PHR data.

The UML sequence diagram shown in Fig. 4 models the

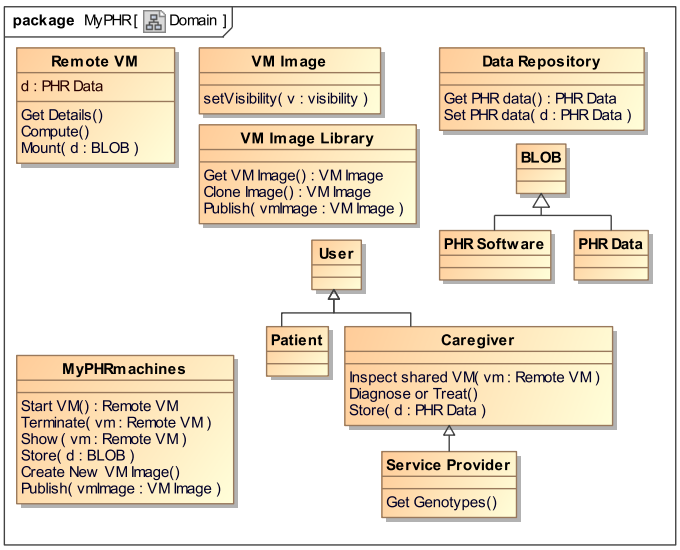

Fig. 2. Conceptual Model for PHRs in the Cloud.

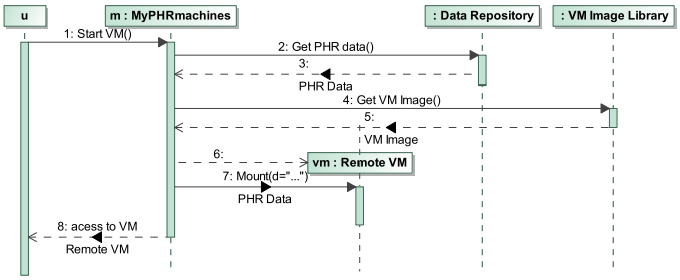

Fig. 3. Starting a new virtual machine session in MyPHRMachines.

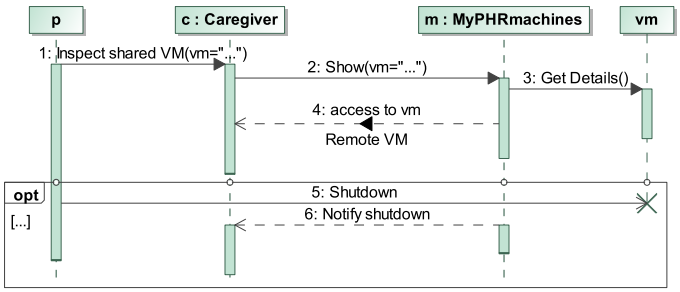

Fig. 4. Sharing a virtual machine session in MyPHRMachines.

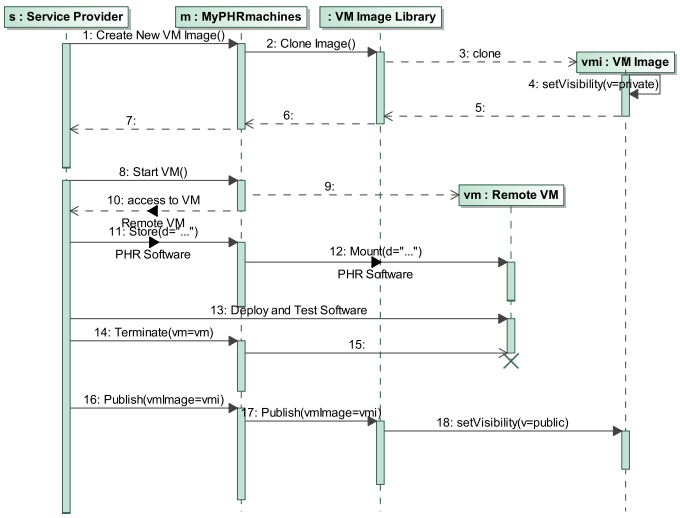

Fig. 5. Deploying a new software service in MyPHRMachines.

patient-oriented scenario of sharing a virtual machine session with a caregiver. The diagram models that a patient passes some virtual machine identifier to the caregiver. In the operational MyPHRMachines prototype, this identifier consists of a long, cryptic string identifier. The caregiver can use this identifier to access the virtual machine even without having a system account, i.e. he/she does not have to login. This enables true 1-click access to the shared virtual machine session. The downside of this solution is that without additional security measures, malicious Internet users could intercept the access delegation message to get access to the VM session. Since caregivers tend to have secure messaging tools in place, we do not consider this as a major threat.

The *opt* block in Figure 4 models optional behavior. It represents the explicit shutdown of a session by a patient. The cross on the dashed *vm* line denotes that such a shutdown requests effectively destroys the VM session. For the sake of clarity, the supportive system should then notify the caregiver who was previously given access to the session (cfr., message *6* in Figure 4).

The crux of many cloud-based application architectures lies in the management of access capabilities of third-party applications. Gordon et al. take the developer API for building services on the Facebook platform as an example [7]. Facebook can be seen as a conventional Web application that does not implement separation between execution containers and personalized mount points for data. Therefore, third party applications can access and possibly log or abuse user data in unpredictable ways, i.e. violating the privacy requirement **R2** identified in Section III. Gordon et al. advocate that (1) APIs need to be minimal for applications on such platforms, and (2) the application platform's user interface should give users an effective view upon which of their data is used when by third party applications. In MyPHRMachines we adopt a different architectural solution to preserve requirement R2. Specifically, we argue that the application platform, i.e. the PHR System, should enable unrestricted analyses on user data as long as input data and computational results are confined in a trusted sandbox, i.e. the the VM running the third-party application software where PHR data can be selectively mounted by the patient.

Application developers of MyPHRMachines applications do not need access to a central database containing patient data. Instead, they contribute software that processes data that is contained in a remote virtual machine. While developing the application, developers will access VM sessions to which their own test PHR data can be mounted. Only when patients access the VM, their actual PHR data will be mounted. Since, by default, VMs are internet-connectionless, this data cannot be

abused by the third-party service provider.

The UML sequence diagram shown in Fig. 5 models the different interactions that occur in the development of a new MyPHRMachines service. First, the service provider requests a clone of a virtual machine image. The MyPHRMachines platform answers this requests by delegating the request to one of its VM image libraries. Initially, a cloned VM image is made private, which means that only its service provider can access it (cfr., message *4* in Figure 5). Message *8* in Figure 5 represents a new session request by the service provider for its private VM image. Message *11* represents the uploading of service binaries to the remote VM. MyPHRMachines does not offer an advanced API for this purpose: the conceptual model shown in Fig. 2 clarifies that the *Store* operation takes a raw BLOB as an argument. After uploading service binaries (and potentially also some test data), a service developer tests his new service in the remote VM (cfr., message *13* in Figure 5). Once all tests have passed, the developer termintates his VM session (message *14*) and "publishes" the image. MyPHRMachines then sets the visibility of the image to *public* (message *18*). This makes the service available to a particular target audience (e.g., citizens of a country, customers of an ensurance company, etc.). Then, patients can access the service using the scenario that was discussed in the context of Fig. 3.

As outlined before, MyPHRMachines blocks traffic from VMs to the internet. Additionally, traffic needs to be monitored both for real-time, automated verification purposes as well as for a posteriori, manual inspections by patients and platform administrators. Firewall policies should be set at the MyPHRMachines platform level instead of inside individual virtual machines since otherwise these can be overridden by malicious expert VM users. The concrete firewall policies, e.g., restricting access to particular domains, disabling outgoing traffic, etc., are outside the scope of this paper.

V. Health Information Sharing for Genomic Data Analysis

We have successfully evaluated the proposed architecture and its privacy-related concerns on a complex use case. More specifically, we have used the MyPHRMachines platform to realize the use case of privacy-compliant personalized medicine in the cloud. Subsection V-A describes the specifics of this use case and subsection V-B discussed the implementation of the use case as well as the open challenges for deploying the related prototype in production.

A. Use Case Description

Ginsburg et al. describe their vision of personalized medicine as follows: *"tailored care is given for every individual based on their specific, molecular disease will become the standard of care. In the prototypical office visit of 2015, the physician will examine a patient's genetic profile (stored on CD ROMs or equivalent), lifestyle, and results from objective molecular screening and monitoring tests. Algorithms, derived from previous research efforts, will be used to compute the likelihood that a patient develops a host of chronic diseases."* [6]. In this paper, we do not focus on the algorithms that are needed to realized this vision. Instead, we show why MyPHRMachines should be used instead of CD ROMs to realize the above vision in a privacy compliant manner.

In order to benefit from personalized medicine, a patient needs to get a digital representation of his/her genetic profile. This involves a one-time analogue to digital conversion of his or her DNA sequence, i.e. the DNA sequencing [11]. The cost of this process is dropping at such a dramatic rate (cfr., [20]) that we can soon[2] expect such a service to be freely available to citizens of developed countries.

A more durable area of value innovation will consist of software services that give personalized medical advice based on a genetic string. Patients may access some of these services themselves while a large share of other services will require physician interpretation [12]. Biomedical research institutes could provide the software services on a pay per use or flat fee basis while specialized clinics could empower their caregivers with licenses for such services. Insurers may also have special coverage schemes for using such services, balancing the cost of the services against their expected health returns.

Clearly, a patient's genomic data is quite privacy sensitive, as it may reveal information, e.g. on past clinical or psychological conditions, that could have a negative influence on one's career, mortgage negotiations, social relations, etc. Although MyPHRMachines cannot ensure that genome data is used ethically by the organization performing the initial DNA sequencing, it can be used to protect patient privacy in the context of software services giving personalized medical advice based on the DNA sequence.

As a validation of our theoretical arguments, we have implemented the above use case in MyPHRMachines. The interested reader is encouraged to explore the use case by following the instructions available at https://sites.google.com/site/myphrmachines/demo-phr. We created inside MyPHRMachines an account for a dummy patient. For this patient, we have uploaded a fully sequenced DNA publicly available on the Internet. Moreover, we deployed to MyPHRMachines a virtual machine image containing the Promethease software, a freely available personalized medicine package performing analysis on sequenced DNA, e.g. calculating probabilities of developing specific sorts of disease in a lifetime.

B. Architecture Revisited

Figure 6 shows a model of the expected system behavior for the genome analysis use case. The *neg* block models that the genome data should not be stored by the service provider that performs the analog to digital data acquisition step. The model shows that the digital data is sent back to the user (cfr., message *3*), who stores it via MyPHRMachines to the secure data repository (cfr., messages *4* and *5*). Such services providers could also store the data directly to such a repository as soon as (inter-)national standards and services

[2]http://geniachip.com/ aims at the 100USD barrier within the decade.

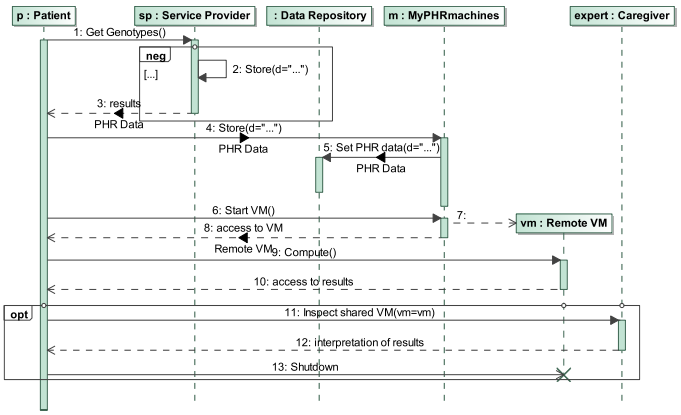

Fig. 6. Accessing a genome analysis service in MyPHRmachines.

consolidate in this context. Steps *6* to *13* are not specific to the genome use case: they simply represent the generic VM start and share operations that we have discussed already in the context of Figures 3 and 4 respectively. Note that in the current prototype of MyPHRmachines, a patient's complete PHR will be mounted to each new VM session. For this use case, that means that besides the sequenced DNA of our example patient, all his other data (e.g., a personal log of blood test results) will be mounted into the VM executing the Promethease software. Although VM users can unmount this data before sharing the session to a caregiver, we argue that a production system should be more user friendly in this context. Future extensions of our prototype will allow patients to select via a simple web interface which type of PHR data to mount in a specific VM.

Going back to privacy-related requirements of Section III, the architectural features implemented by MyPHRMachines in the above use case clearly satisfy requirement **R2**. The sequenced DNA, i.e. the PHR data, can be accessed by the VM implementing the Promethease software only within a virtual sandbox environment without Internet connection. This prevents the third-party service developer to download the PHR data and enables real-time or a posteriori checks of malicious usage patterns within the sandbox.

The satisfaction of requirement **R1** is also guaranteed by the possibility to unmount (subsets of) the available PHR data in the required VMs. The requirement can be supported better by enabling patients to select before VM startup which PHR data subset should be mounted in the first place. Since this paper focuses on the system architecture, the improvement of the user interface does not represent a fundamental issue but it is the object of ongoing implementation work nonetheless since we aim to use the prototype for dissemination purposes.

VI. Conclusions and Outlook

This paper investigates privacy-related concerns in the management of PHR data. Our analysis shows that a novel PHR system architecture is desirable since privacy-related requirements can only be partially satisfied by commercial PHR systems. We also demonstrate that the architecture can be implemented. In particular, we discuss how a complex privacy-sensitive use case (in the domain of personalized medicine) has been realized using our prototypical system implementation.

Future work will concern the refinement of the so-called MyPHRMachines prototype, which we use as a vehicle to validate and disseminate our results. This ranges from very simple features, such as the selective mount of PHR data, to more complex functionality, such as the run-time or a posteriori checks of third-party PHR data malicious usage patterns within the VM. We are also investigating the possibility of applying MyPHRMachines at a care institution requiring patient health information from several providers.

References

[1] N. Archer, U. Fevrier-Thomas, C. Lokker, K. A. McKibbon, and S. E. Straus. Personal health records: a scoping review. *JAMIA*, 18:515–522, July 2011.

[2] J. Collmann and T. Cooper. Breaching the security of the Kaiser Permanente Internet patient portal: the organizational foundations of information security. *JAMIA*, 14(2):239–243, 2007.

[3] A. Datta, N. Dave, J. Mitchell, H. Nissenbaum, and D. S. . Privacy challenges in patient-centric health information systems. In *Usenix Workshop on Health Security and Privacy*, 2010.

[4] D. Detmer, M. Bloomrosen, B. Raymond, and P. Tang. Integrated personal health records: Transformative tools for consumer-centric care. *BMC Medical Informatics and Decision Making*, 8(1):45, 2008.

[5] G. Eysenbach. Medicine 2.0: Social networking, collaboration, participation, apomediation, and openness. *J Med Internet Res*, 10:e22, July 2008.

[6] G. S. Ginsburg and J. J. McCarthy. Personalized medicine: revolutionizing drug discovery and patient care. *Trends in Biotechnology*, 19(12):491 – 496, 2001.

[7] G. Hull, H. Lipford, and C. Latulipe. Contextual gaps: privacy issues on facebook. *Ethics and Information Technology*, 13:289–302, 2011.

[8] D. Kaelber and E. C. Pan. The value of personal health record (PHR) systems. In *AMIA Annu Symp Proc*, pages 343–347, 2008.

[9] D. C. Kaelber, A. K. Jha, D. Johnston, B. Middleton, and D. W. Bates. Viewpoint paper: A research agenda for personal health records (PHRs). *JAMIA*, 15(6):729–736, 2008.

[10] J. S. Kahn, V. Aulakh, and A. Bosworth. What it takes: characteristics of the ideal personal health record. *Health Aff*, 28:369–376, Mar. 2009.

[11] J. Kuczynski et al. Experimental and analytical tools for studying the human microbiome. *Nature Reviews Genetics*, 13(1):47–58, Dec. 2011.

[12] S. Murphy. In need of a reality check. *Nature Biotech*, 27:422, May 2009.

[13] C. Pagliari, D. Detmer, and P. Singleton. Potential of electronic personal health records. *BMJ*, 335(7615):330–333, 8 2007.

[14] V. Rybynok, P. Kyriacou, J. Binnersley, and A. Woodcock. MyCare Card development: Portable GUI framework for the personal electronic health record device. *IEEE Transactions on Information Technology in Biomedicine*, 15(1):66–73, 2011.

[15] D. F. Sittig. Personal health records on the internet: a snapshot of the pioneers at the end of the 20th century. *Int. J. of Medical Informatics*, 65(1):1 – 6, 2002.

[16] A. Srinivasan. Keeping online personal records private: Security and privacy considerations for web-based PHR systems. *Journal of AHIMA*, 77:62–63, Mar. 2006.

[17] Students of the Interdisciplinary Law and Technology Workshop. Privacy in the digital environment. Technical report, The Haifa Center of Law and Technology Publication Series n. 7, 2005.

[18] P. Tang, J. S. Ash, D. W. Bates, J. M. Overhage, and D. Z. Sands. Personal health records: Definitions, benefits, and strategies for overcoming barriers to adoption. *JAMIA*, 13:121–126, Mar. 2006.

[19] P. Van Gorp and M. Comuzzi. MyPHRMachines: Lifelong Personal Health Records in the cloud. In *Proc. 25th IEEE Int. Symposium on Computer-based Medical Systems*. forthcoming, 2012. forthcoming.

[20] K. Wetterstrand. DNA sequencing costs - data from the NHGRI large-scale genome sequencing program, Jan. 2012.

[21] K. Win, W. Susilo, and Y. Mu. Personal health record systems and their security protection. *J Med Syst*, 30:309–315, 2006.