The Comprehensive Health Information System: a Platform for Privacy-Aware and Social Health Monitoring
De Coi Juri Luca, Gwendal Deaunay, Adrien Martins Albino, Fabrice Muhlenbach, Pierre Maret, Guillaume Lopez, Ichiro Yamada

To cite this version:

HAL Id: hal-00725365
https://hal.archives-ouvertes.fr/hal-00725365
Submitted on 28 Aug 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
THE COMPREHENSIVE HEALTH INFORMATION SYSTEM: A PLATFORM FOR PRIVACY-AWARE AND SOCIAL HEALTH MONITORING

Juri Luca De Coi, Gwendal Delaunay, Adrien Martins Albino, Fabrice Muhlenbach, Pierre Maret
Université de Lyon
CNRS, UMR5516, Laboratoire Hubert Curien
Université de Saint-Etienne, Jean Monnet
18 Rue du Professeur Benoît Lauras, 42000 Saint-Etienne (F)

Guillaume Lopez, Ichiro Yamada
Graduate School of Frontier Sciences, The University of Tokyo
5-1-5 Kashiwanoha, Kashiwa-shi, Chiba-ken 277-8563 (J)

ABSTRACT
Rapid ageing of world population makes the traditional hospitalization therapy more and more expensive, new forms of medical care are hence required. At the same time, the medical community perceives the availability of vital data as an essential prerequisite to acquire objective information about patients. Finally, the patients themselves show willingness to provide such data, in order for their health conditions to be regularly monitored.

This paper presents the last developments of the Comprehensive Health Information System (CHISel’d), a platform for continuous health monitoring whose main goal is to show how information technologies can be exploited to provide a feasible alternative to the traditional hospitalization therapy. Remarkable features of CHISel’d are its capability to process patient data according to a dynamically evolving set of data mining techniques and to share them among stakeholders (doctors, researchers, friends…) according to patient-defined access policies and in the frame of e-communities. These features enable CHISel’d to address the needs perceived by patients and the medical community.

KEYWORDS
Health Information Monitoring, Healthcare Community, Information Processing Platform.

1. MOTIVATIONS
Nowadays, the world population over sixty years old reached 600 million people and about 900 million people suffer from chronic diseases (Kotani et al, 2007), resulting in an increase of medical expenditure. Rapid ageing of world population results in increasing medical expenses, up to the point that nations will not be able to afford them anymore. As an example, in 2007 the percentage of Japanese population over sixty-five years old exceeded 20% and it is expected to exceed 30% by the year 20301. As these data show, a different and more sustainable form of medical care than the traditional hospitalization therapy is required.

At the same time, the medical community sets a strong demand for patients’ vital data in daily life, since the information physicians collect during medical interviews is often not objective and the symptoms reported do not necessarily appear during a medical examination. Such data should both encompass long periods and be as detailed as to allow the identification of small-scale discontinuities. Moreover, both

---

2 Hospitalization therapy (a.k.a. inpatient care) is the care of patients whose condition requires admission to a hospital.
physical and psychological characteristics as well as information about the physical activity carried out by the patient should be tracked. If provided with such data, physicians could identify problems timely, devise countermeasures and establish programs tailored to the patient’s lifestyle and preferences. On the other hand, data collected during the treatment could be used both to check whether the patient indeed sticks to his/her program and to evaluate the effectiveness of the program itself.

Finally, awareness and interest in healthcare is growing among common people: for three years in a row, “my own health” (resp. “health of my family members”) obtained the second (resp. third) position in an opinion poll about Japanese people’s worries and anxieties. These results seem to indicate that common people would appreciate a steady, accurate and continuous monitoring of their health conditions, which in turn requires the systematic collection of daily-life vital data anywhere and at any time.

Since quite some time, information technologies have been successfully applied to the medical field: applications such as tele-radiology, tele-consultation and tele-surgery are being commonly employed to support wellness and independent living (Hung and Zhang, 2003). Healthcare sensors embedded in the environment can provide health information by continuously monitoring people’s activity at home (Stanford, 2002). Wearable healthcare sensors are non-intrusive wireless microsensors enclosed in items which can be worn (such as a ring—Asada et al, 2003): in comparison to common healthcare sensors, they do not require a closed environment and can sense vital signs directly (Kunze et al, 2002 and Sachpazidis, 2002). Moreover, since patients can wear different wearable healthcare sensors at the same time, these can be exploited in order to concurrently track health data of different kinds.

The goal of this paper is to demonstrate a different way information technologies can prove useful to the medical field: since 2007, the authors have been working on the Comprehensive Health Information System (CHISel’d), a platform collecting, storing and making practical use of health information (Faudot et al, 2010, Lopez et al, 2011 and Shuzo and Yamada, 2009). Remarkable features of CHISel’d are

1. it stores physiological and environmental data as provided by (wearable) healthcare sensors. By tracking daily-life health data over long periods of time, it enables the constant monitoring of patients’ health conditions
2. it allows to process such data in order to extract information any user can understand by means of data mining techniques (Han et al, 2011), thereby making them aware of patients’ health conditions and overall status
3. it allows patients to fully control the sharing of such information with appropriate stakeholders (doctors, researchers, friends…) as necessary, thereby enabling more effective and efficient health services

CHISel’d is a concrete answer to the research problems we listed at the beginning of this section. Firstly, it provides physicians with a wealth of objective information and hence enables them to precisely identify the patients’ profile. Secondly, it meets the desire of common people to be kept up-to-date with accurate information about their conditions and hence helps them to preserve their health and, in the end, improve their quality of life (QOL). Last but not least, by regularly monitoring patients’ conditions and providing them with up-to-date health information, CHISel’d fosters disease prevention as opposed to disease treatment, thereby (hopefully) producing a revolution in the patients’ way of thinking whose most visible effect should be a reduction of the overall medical expenses.

The remainder of this paper is organized as follows. Section 2 presents the requirements a platform like CHISel’d must fulfill as they are perceived by its end-users. Section 3 outlines the research issues we had to face when designing the platform itself. The architecture of CHISel’d is described in Section 4, whereas Section 5 accounts for related work. We conclude in Section 6 by summarizing our approach and presenting future work.

2. REQUIREMENTS

---

Whilst being one single system, CHISel’d is supposed to provide as many profiles as roles played by its users. Up to now, four roles have been identified: patient, researcher, doctor and friend. For each such role, this section describes the corresponding requirements which a platform like CHISel’d must fulfill. Further roles (e.g., medical institution or health business company) or specializations of existing ones (e.g., family member or neighbor as specializations of friend) will be added as needed. This will in turn require to: (i) identify the requirements for such roles; and (ii) extend CHISel’d accordingly.

Patients obviously play a central role within the CHISel’d platform, being the main reason which led to its development. In order to regularly monitor their health conditions, CHISel’d must: (i) be able to track their daily-life health data; and (ii) provide sufficient storage capabilities to keep them over long periods of time. On the other hand, patients must be granted the possibility to retrieve their own data. Since it is unlikely that common people will understand low-level data like the ones captured by (wearable) healthcare sensors, a suitable presentation facility must be provided. Finally, an easy-to-use interface making patients feel to be the main actors in their treatment would be a plus.

The role of researchers is to make sense out of health data, i.e., to develop algorithms which mine information out of raw data. Some algorithms will simply extract higher-level information out of lower-level one (e.g., overall body activity out of data about chest acceleration). Other algorithms will inspect health data to find out evidences of potential diseases. Such algorithms will play a central role in presenting health data to patients in a way they can understand, but they will be most useful to doctors as well: even assuming that doctors have the competences to extract information out of raw data themselves, the automation of this activity will let them spare time and focus on more important tasks. Obviously, researchers need to access the health data stored within CHISel’d in order to develop their algorithms. Moreover, a seamless integration of newly developed algorithms and already available ones would be desirable: ideally, researchers should only need to upload their algorithms’ implementation to CHISel’d in order to make them immediately available to doctors.

Not surprisingly, the role of doctors is to take care of patients by monitoring their health conditions. To accomplish this task, doctors are supposed to make use of the algorithms available within CHISel’d. More specifically, for each patient they must identify the most suitable algorithms according to his/her profile. The higher-level information they retrieve might then be shared with the patient himself/herself. CHISel’d must allow doctors and patients to communicate directly, so that the formers can provide advices to the latters. Finally, doctors must be able to alert appropriate actors in case they detect an abnormal condition in some patient’s real-time data.

The role friends play is a social one and is related to: (i) their physical proximity to patients; or (ii) their intention to follow patients’ health status. The physical proximity enables friends to spring to their aid in case of an emergency: for this reason, CHISel’d must be able to send them an alert, either if requested by a doctor or if some algorithm running on patients’ real-time data detected an abnormal condition. On the other hand, the ability to access patients’ health data (as far as allowed by patients themselves) enables friends to support them better (e.g., by helping them following their program). Communities of/around patients may then arise to foster motivation or reassure stakeholders.

3. CHALLENGES

The requirements presented in Section 2 fit nicely into four categories.

1. Tracking of patients’ health data
2. Storage of health data and algorithms mining information out of them
3. Access to health data
4. Communication-related requirements

In order to obtain a picture as accurate as possible of patients’ health conditions, all of physical (e.g., activity and behavior), psychological (e.g., stress and emotions) and context data (e.g., atmospheric pressure, outer

---

4 By patients we do not necessarily mean ill people but possibly healthy people willing to monitor their health conditions.
temperature) should be tracked. As described in Section 1, wearable healthcare sensors are the least-intrusive technology available to date in order to sense vital signs. As a consequence, a platform like CHISel’ed can be expected to provide as many wearable healthcare sensors as needed to sense all of the health data mentioned above.

Storage of health data and storage of algorithms pose different problems. In the first case, memory is the main concern: regularly monitoring patients’ health conditions produces large amounts of data and the ability to keep them over long periods of time makes the availability of huge amounts of memory essential. In the second case, memory is not expected to represent a big issue, not even for a large number of algorithms: whenever untrusted code has to be loaded and executed on a system, the main concern is rather security. Suitable countermeasures should then be taken in order to avoid malicious code to harm the system.

The common label “access to health data” refers to a set of different issues. First of all, being health data sensitive information, the privacy of their owners must be taken into account when evaluating access requests: as mentioned in Section 2, the approach most respectful of patients’ privacy would probably be to enable them to decide themselves who can access their data. Even independently from privacy concerns, it is not meaningful allowing all users to access all data, since not all users can understand all of them: different access policies should hence be defined according to the role played by the requesting user, so that users can only access data whose granularity level is sufficiently high with respect to the role they play. Finally, “access to data” refers to the format the accessed data are presented to the requesting user: researchers might only need to access (files containing) raw data through a software API, whereas other users might need to access higher-level information (such as the patients’ daily level of activity) through a user-friendly presentation facility.

The last set of requirements is related to communication. As outlined in Section 2, CHISel’ed must support both user-triggered communication (e.g., message exchange) and reactive communication. By the latter we mean information exchange which takes place automatically whenever a given event occurs, the prototypical example being the notification of patients’ friends when the formers could be in danger, as detected by some algorithm running on their real-time data. Finally, the ability of doctors to specify which algorithm applies to which patient can be considered a communication-related requirement as well, since the doctor (implicitly) communicates to the patient which higher-level information the system will handle for that specific patient, thereby enabling him/her to define corresponding access policies.

4. ARCHITECTURE

The four sets of requirements described in Section 3 correspond to the main components of the CHISel’ed architecture. Beside the components themselves (shown in the middle column), Fig. 1a attempts to depict the main interactions both among them and with system users according to their role. Interactions in which users act as information providers (resp. consumers) are shown in the left (resp. right) column.

In a bottom-up fashion, patients provide health data to (wearable) healthcare sensors which transmit them to the storage facility. Algorithms developed by researchers in order to mine information out of health data are stored as well. These algorithms can be retrieved by doctors who will possibly assign them to their patients according to their profile. Doctors, as well as friends, patients and researchers, can retrieve (sufficiently high-level) health data by means of suitable interfaces and according to patient-defined access policies. Finally, communication among users can occur through the corresponding component of the CHISel’ed platform and the platform itself can contact users upon occurrence of specific events. For generality, CHISel’ed: (i) allows communication between any pair of users; and (ii) can issue notifications to any user; although e.g., friend-researcher communication is unlikely to occur, as are notifications to researchers.

The remainder of this section is devoted to the description of the main components of the CHISel’ed architecture as well as their interactions. Since the Caption (resp. Storage) component has been already described in (Lopez et al, 2011–resp. Faudot et al, 2010), in this paper we will focus on the Data Access and Communication components (described in Sections 4.1 and 4.2 respectively).

4.1 The Data Access Component
As described in Section 3, not all users of CHISel’d should be allowed to access all data it stores, since: (i) being health data sensitive information, the privacy of their owners must be taken into account; and (ii) not all users can understand all of them. For this reason: (i) different access policies should be defined according to the requesting user; and (ii) the approach most respectful of patients’ privacy would probably be to enable them to decide and control themselves who can access their data.

As for the first point, the need of automatizing policy enforcement is well known in the research community, as witnessed by the IEEE International Symposium (formerly Workshop) on Policies for Distributed Systems and Networks which is being held since 1999. Policy languages have been proposed as a solution for such a need: they are special-purpose programming languages which allow defining policies in a formal way. Upon definition, formal policies can be enforced by providing them as input to a compatible policy engine.

Formal policies yield many advantages in comparison to natural-language ones. To start with, their semantics is unambiguous: they can hence be shared among interested parties without the risk of misunderstandings. This goal is usually accomplished by grounding policy languages in some mathematical formalism: as a consequence, the same formalism can be exploited in order to infer implicit knowledge out of formal policies (e.g., whether they are consistent or too restrictive). On the practical side, the choice of uncoupling policy definition and enforcement (as opposed to building a specific policy once and for all into the policy engine) is a winning one, since a modification in the policy does not require a completely new engine but only to provide the current one with a different input. Finally, although policy languages are indeed programming languages, they are typically designed as declarative ones, i.e., they usually belong to a class of programming languages which is claimed to be easier to learn than alternative ones: as a consequence, writing formal policies may not only be a task for computer scientists.

A number of policy languages have been defined so far. According to comparisons available in the literature (De Coi and Olmedilla, 2008 and Duma et al, 2007), PROTUNE (Bonatti et al, 2010) turns out to be one of the most complete ones available to date and as such it was an obvious candidate for the CHISel’d platform.

As for the second point, although formal policies are not as hard to write as usual computer programs, patients cannot be expected to define them themselves. Fig. 1b shows the editor CHISel’d provides to patients and other trustworthy users in order to define policies: combo boxes list all users and roles available in the platform. By selecting the appropriate (user, role) pairs and clicking on the Associate

---

5 http://www.policy-workshop.org/
Contact/Role button, roles can be assigned to users. New roles can be created by clicking on the Create role button: the user will be presented with a screen enabling him/her to specify a name for the role as well as to select: (i) which data the role members will be able to access; and (ii) at which granularity level such data should be accessed.

The solution provided by CHISel’d is intuitive and does not almost require any learning step. Also, the great expressiveness of the PROTUNE policy language overcomes conventional role-based policies by enabling the definition of access rules which depend not only on the role of the requester but also on the granularity level of the data to be accessed and on context information, as shown in Fig. 2.

Figure 2. PROTUNE policies allow to specify which requester is allowed to access which data according to which context

4.2 The Communication Component

As described in Section 3, CHISel’d should support both user-triggered and reactive communication, i.e., the exchange of information among its users should be initiated either by the users themselves or by the system upon occurrence of a given event.

User-triggered communication is natively provided by a broad spectrum of technologies which have been developed during the last years and are commonly referred to under the shared label “social software”. Social software encompasses a number of technologies which allow one-to-one (e.g., instant messaging applications), one-to-many (e.g., blogs) and many-to-many communication (e.g., online chat technologies, Internet forums, wikis). Social networking and online community services exploit many such technologies in order to provide their users with an integrated experience.

Social networking and online community services foster the creation of virtual communities, i.e., social networks of individuals who interact through specific media in order to pursue common interests or goals. Virtual communities can give users a feeling of membership and belonging and promote the creation of trust relationships among them. As for the medical domain, virtual communities can enhance patients’ mood, behavior and willingness to follow a program, thereby improving their QOL (Battles and Wiener, 2002 and Eysenbach, 2003).

Since social networking services: (i) employ technologies which enable user-triggered communication; and (ii) can positively impact their users’ QOL; it was an obvious decision designing the communication component of CHISel’d as an enhanced social networking service. When it came to the point of choosing the Social Network Engine to use as a basis for it, extensibility was considered the most relevant factor because of the modifications we planned to do in order to configure the chosen engine toward our needs. According
to this criterion, Elgg\footnote{http://www.elgg.org/} has been regarded as the open-source Social Network Engine best suiting our requirements.

The most remarkable modifications we made to Elgg are the addition of: (i) plugins to visualize health data in a user-friendly way; and (ii) a mechanism enabling the engine to contact users upon occurrence of a given event. As for the first point, we developed a set of plugins which, if installed on an Elgg instance, integrate health data into its user interface. Data are shown to a user: (i) only if s/he can access them; and (ii) in a user-friendly way, i.e., not as a list of values but in a graphical form. For instance, heart activity is displayed as an ECG, whereas the percentage of time spent by the patient laying, sitting or standing is visualized as a pie chart. As for the second point, we added Elgg a mechanism enabling it to send users messages upon occurrence of a given event: whenever some algorithm running on patients’ real-time data detects an abnormal condition, such mechanism can be exploited in order to automatically notify patients’ friends according to the scenario described in Section 2.

5. RELATED WORK

To the best of our knowledge, an approach as broad and generic as the one pursued by CHISel\’d appears to be unique. On the one hand, a number of platforms make use of social software in order to build virtual communities of healthcare stakeholders. However, such platforms focus on the community aspect and do not provide any means to track patients’ health data and provide them to researchers. On the other hand, some systems supporting remote healthcare monitoring have been proposed. However, such systems focus on health data collection and do not provide an overall framework within which captured data can be processed. Moreover, such data are not used in their social dimension, i.e., they cannot be shared within a virtual community and privacy issues appear to be neglected.

As for the first set of platforms, social software in the healthcare domain may serve different purposes according to the point of view one wants to take. For healthcare professionals, it is a means to disseminate peer-to-peer knowledge and highlight individuals as well as institutions. For patients, it offers the chance to connect with others suffering of similar diseases and possibly to gain motivation and encouragement through others’ support. Social software itself (namely, a wiki) is exploited to provide the probably most comprehensive list of healthcare-related social software available to date: the goal of the \textit{Pharma and Healthcare Social Media Wiki}\footnote{http://www.doseofdigital.com/healthcare-pharma-social-media-wiki/} is “to house every Social Media project that has been created by […] healthcare companies”. Listed social software is classified according to the intended users (patients, healthcare professionals or marketing professionals), the employed technology (blog or wiki) or the hosting platform (Facebook\textsuperscript{8}, YouTube\textsuperscript{9}, MySpace\textsuperscript{10} or Twitter\textsuperscript{11}).

As for the second set of systems, a good overview thereof is provided by (Anliker et al, 2004): as the authors point out, most available systems target a specific environment and can hardly be generalized to support different ones. An approach not described in (Anliker et al, 2004) is presented in (Blount et al, 2007): a major drawback of this proposal is its usage of HTTP to transfer health data, which makes regular health monitoring and reactive communication unpractical. We conclude by mentioning parallel work which is being carried out by a partner of ours: (Rikitake et al, 2009) describes a system which exploits wearable healthcare sensors and mobile networks to support continuous health data capture and storage as well as reactive communication. The technology it is based upon (namely, Next-generation network/IP Multimedia Subsystem) enables it to overcome the shortcomings of (Blount et al, 2007) and provides native support for AAA (Authentication, Authorization and Accounting), QoS (Quality of Service) and event notification.
6. CONCLUSIONS AND FURTHER WORK

Rapid ageing of world population makes the traditional hospitalization therapy more and more expensive, new forms of medical care are hence required. At the same time, the medical community perceives the availability of vital data as an essential prerequisite to acquire objective information about patients. Finally, the patients themselves show willingness to provide such data, in order for their health conditions to be regularly monitored.

This paper presented the last developments of the Comprehensive Health Information System (CHISel’d), a platform for continuous health monitoring whose main goal is to show how information technologies can be exploited to provide a feasible alternative to the traditional hospitalization therapy. Remarkable features of CHISel’d are its capability to process patient data according to a dynamically evolving set of data mining techniques and to share them among stakeholders (doctors, researchers, friends…) according to patient-defined access policies and in the frame of e-communities.

Not surprisingly for a long-term project far from its end, the current version of the CHISel’d platform can be improved and extended in many ways. A first line of development concerns the roles which can be played by system users. As described in Section 2, CHISel’d currently supports only four roles: patient, researcher, doctor and friend. However, evidences are showing that the integration of further roles (e.g., medical institution or health business company) or specializations of existing ones (e.g., family member or neighbor as specializations of friend) would be beneficial. However, this can be considered a minor issue since Section 4.1 already showed that the CHISel’d technology makes the creation of new roles straightforward.

A second line of development has been mentioned in Section 4.1: the policy editor CHISel’d provides to patients and other trustworthy users is more constraining than it ought to be. We are currently considering replacing it with an interface based on controlled natural languages. Whilst not being much harder to use than the current one, such an interface would enable to fully exploit the expressiveness of the PROTUNE policy language. Preliminary results (De Coi et al, 2009) are encouraging and we plan to pursue this approach further.

Finally, Section 4.2 pointed out that doctors should be able to specify which algorithm applies to which patient. CHISel’d does not provide yet for such a possibility, which is hence regarded as further work.

ACKNOWLEDGEMENT

This work has been supported by the Core Research for Evolutinal Science & Technology (CREST), Japan Science and Technology Corporation (JST), by the Centre National pour la Recherche Scientifique (CNRS), France and the Région Rhône-Alpes, France.

REFERENCES

Book

Journal


Conference paper or contributed volume


