

Project eSanté

*Architecture and Security of a
National eHealth Platform*

(Deliverable for WP 7, WP 8, WP 13)

Version 1.06
June 29 2011

Contact :

Dr. Stefan Benzschawel (stefan.benzschawel@tudor.lu)

Heiko Zimmermann (heiko.zimmermann@tudor.lu)

CRP Henri Tudor

c/o Centre de Ressources SANTEC

2a rue Kalchesbrück

L-1852 LUXEMBOURG

Phone: +352 42 59 91-250 Fax: +352 42 59 91-251

Web: www.tudor.lu

© SANTEC, published on <http://www.santec.lu/project/esante/efes/start>

The authors thank their partners from the Health Ministry of Luxembourg for their very helpful advises and for providing insights into organizational and legal aspects of the eHealth platform. Special thanks belong to René Krippes, Mike Schwebag and Jean-Charles Dron for many fruitful discussions and very helpful comments.

Special thanks belongs to Pierre Weimerskirch and his colleagues from the National Data Protection Commission of Luxembourg for very valuable hints and their explicit support for the concept of pseudonymization combined with a two step encryption.

DOCUMENT HISTORY

Version	Initiative	Object	Date	Redaction
0.1	SBE	Draft version with first TOC		SBE
0.2	SBE	Feedback of René to TOC and text snippets		SBE
0.3	FWI	Hints and Comments		SBE
0.4	SBE	Draft version chapter 1 - 6	27/09/10	SBE
0.5	HZI	Hints and Comments	29/09/10	SBE
0.6	URO	Review	29/09/10	SBE
0.7	URO	Review	04/10/10	SBE
0.8	RKR	Review, Corrections, Remarks	19/10/10	SBE
0.81	CPR	Review, Corrections on Consent Chapter	25/11/10	SBE
0.82	HZI, URO	Review and discussion, long term archiving	26/11/10	SBE
0.83	RKR	Improvements	08/12/10	SBE
0.84	SBE	Renew chapter	28/12/10	SBE
0.85	MDS/URO/ SBE	Cont. and corrections, solution outline	28/12/10	SBE
0.86	HZI/SBE	Integration of IHE Chapter	17/01/11	SBE
0.87	SBE	Cont. additional chapter, Pictures	18/02/11	SBE
0.88	MDS	Proof read and Comments of MDS accepted	10/03/11	SBE
0.89	HZI	Review of IHE chapter with references	10/03/11	SBE
0.90	MDS	Proof read cont.	11/03/11	SBE
0.91	SBE	Rework Intro, Solution Outline, related work, and other parts	15/03/11	SBE
0.92	MDS	Review of Intro	17/03/11	SBE
0.93	SBE	Rework until chapter 9	18/03/11	SBE
0.94	SBE	Review chapter 10,11,13,14,15	21/03/11	SBE

0.95	SBE	Feedback included of RKR, FWI, Feedback of JCD copied in (of pdf)	05/04/11	SBE
0.96	SBE	Comments worked in and open discussions to clear remaining points in a meeting	14/04/11	SBE
0.97	SBE	Explain some details for crypto key generation	15/04/11	SBE
0.98	SBE	Final Feedback of RKR, JCD, FWI, MDS included	06/05/11	SBE
1.00	SBE		06/05/11	SBE
1.01	SBE	Some optical corrections	10/05/11	SBE
1.02	RKR	Affiliation check as web-service (extended) new chapter 8 introduces for ext. services	13/05/11	SBE
1.03	RKR	Exec Summary, some clarifications in c. 5	17/05/11	SBE
1.04	RKR,JCD	Chapter 5 improvements	26/05/11	SBE
1.05	SBE	Version for publication	20/06/11	SBE
1.06	SBE	Minor corrections	29/06/11	SBE

CPR = Cédric Prusky	MDS = Marcos Da Silveira
FWI = François Wisniewski	RKR = René Krippes (MdSanté)
HZI = Heiko Zimmermann	SBE = Stefan Benzschawel
JCD = Jean-Charles Dron (HMS France)	URO = Uwe Roth

Table of Contents

1	Requirements for an eHealth Platform.....	11
2	Benefits and Risks.....	13
2.1	Benefits.....	13
2.2	Risks.....	14
3	Cornerstones of the eHealth Platform.....	15
3.1	Three Kinds of Information Exchange Methods.....	15
3.2	Content of the Records.....	17
3.3	Pragmatic Structure of the Health Records.....	18
3.4	First eHealth Services of the eHealth Platform.....	18
3.5	Cross Border Communication.....	21
3.6	Secondary Usage for Statistics.....	22
3.7	Open Questions.....	22
4	Technical Concepts of the eHealth Platform	24
4.1	Centralized or Decentralized Repositories.....	24
4.2	Centralized Registry – With or Without ?.....	24
4.3	IHE Cross Enterprise Document Sharing (XDS).....	25
4.4	Evaluation Criteria for Architectural Topology	27
5	Encryption and Pseudonymization meet IHE XDS.....	33
5.1	Separation, Encryption and Pseudonymization.....	33
5.2	Saving the Costs of a Cryptography PKI.....	36
5.3	Security Shell and Possible Extensions.....	37
5.4	Statistical Usage under Pseudonymization.....	38
5.5	The Minimal Constellation for Security.....	38
6	Related Work.....	39
6.1	Québec's Dossier de Santé.....	39
6.2	Belgian Approach with TTP.....	39
6.3	German Approach D2D.....	40
6.4	PIPE Project.....	41
6.5	Other Approaches described by T. Neubauer.....	41
7	Details of the eHealth Platform.....	42
7.1	IHE-XDS Adaptation	42
7.2	Pseudonymization.....	43
7.3	Primary Usage for Patient Care.....	44
7.4	Secondary Usage for Statistics.....	45
7.5	Personal Dashboard.....	46
7.6	The Enemy knows the System.....	46
7.7	Archiving and Data Aging of Encrypted Documents.....	46
7.8	Hiding Patient's Identity against the Web-Server.....	50
7.9	Certificates, User-Roles, and Role-Assigned Applications.....	50
7.10	PKI for Authentication, for Signature and for Cryptography.....	51
7.11	Alert Functions.....	52
8	Embedding External Services.....	52
8.1	Affiliation Checks and Insurance Information.....	52

8.2	Generic Services for User Authentication / Access Verification.....	53
8.3	Improvement of Demographic Data Quality	53
9	Data Protection Overview.....	54
9.1	Levels of Data Protection.....	54
9.2	Security Limitations.....	55
10	Consent Management	56
10.1	Media for Patient's IT-Consent Declaration.....	57
10.2	Opt-In or Opt-Out.....	58
11	Compilation and Composition of the Health Records.....	59
11.1	Compilation of Relevant Subsets in the Source Systems.....	59
11.2	Compiled Subsets to compose the Health Record.....	59
11.3	Automation and Fine Tuning of Health Record Composition.....	59
12	Storage Locations for Patient's Health Record.....	60
12.1	Paper-Based Health Records.....	60
12.2	Pocket-Based Electronic Health Records.....	61
12.3	Pocket-Based Structured Electronic Health Records.....	62
12.4	Platform-Based Structured Electronic Health Record.....	62
12.5	Patient's Choice ?.....	63
12.6	Technically Enabled Options.....	63
13	From Concept to Implementation.....	64
13.1	IHE Profiles that are useful to realize this Concept.....	64
13.2	IHE Platform XDS, XDR, NAV, XCA, XDS-I.....	64
13.3	IHE Security, Node Authentication & Authorization ATNA & CT.....	68
13.4	IHE User Identification & Authentication XUA, XUA++.....	69
13.5	IHE Patient Identification PIX, PAM, PDQ, XCPD.....	70
13.6	IHE Consent Management BPPC.....	72
13.7	IHE Content Profiles for Medical Documents XDS-MS.....	73
13.8	Closing the Gaps with Connectors.....	75
14	Iterative Evolution.....	75
15	Test-Bed Backlog for 1st and/or 2nd Iteration.....	76
16	Next Iterations for Research, Concept, and Test-Bed.....	78
17	References.....	82

Executive Summary

A national Electronic-Health-Record (EHR) platform aims to provide an environment where the most relevant medical information of patients can be safely shared and exchanged. The relevant information is compiled by the health professionals as users of the so called primary systems. These selections are composed into the national health record of each patient and/or send directly to an other involved health professional.

The benefits of sharing are the increase of quality of care activities, reduction of redundant exams, and better contextualization of patients' health evolution with the full range of information about a patient's health constitution. Next steps are the implementation of decision support systems that use the available information in the EHR of a patient to support healthcare activities. This is the way towards a personalized medicine.

Besides these direct personal advantages, statistical evaluations of a medical data can indirectly contribute to the improvement of care activities. The knowledge of systematic correlations can have an impact on the future treatment of patients.

Sharing implies the trust on each other and it rises the danger of misuse of shared information. The access to such data by private insurance companies, banks, employers, etc. is forbidden by the law. Illegal attacks to access the information have to be avoided.

Technically, the backbone of the system is described as follows:

- (1) Certificate-based user authentication for health professionals and patients.
- (2) Role-based user management with pre-registered users.
- (3) Separated storage of identification data and medical data by use of pseudonymization.
- (4) Encryption of clinical documents.
- (5) Individual access restrictions through IT-consent declarations.
- (6) Access to the logging information by the patient and automatic notifications to patients as psychological barrier against unjustified accesses done in the name of an “emergency situation”.
- (7) Non-disclosure guarantee with respect to administrators and intruders of the systems.

External services of other parties in the healthcare sector can reuse the user authentication and user management that is offered by the platform.

For statistical use, data protection is based on four pillars:

- (a) Stripped fragments of the medical reports are stored in a separated database.
- (b) Fragments are definitively without any person identifying data.
- (c) Same pseudonymization techniques are applied for the fragments.
- (d) Governmental IT-consents are mandatory for using the data fragments for statistical analysis.

The proposed concept for the eHealth platform is well founded on standards for document sharing and exchange in medical environments. The basic IT infrastructure consists of one central registry and multiple central and/or decentral repositories for the clinical documents. In order to assure a high level of data protection, the proposed concept includes some improvements with respect to the commonly usage of security standards. These improvements can be described as pragmatic combination of *document exchange and sharing standards* together with *pseudonymization, encryption, and electronic signatures*.

This document will be used as conceptual implementation guide for the national eHealth platform of Luxembourg. As immediate next step a prototype of the eHealth platform will be implemented mainly to validate the technical feasibility of the proposed improvements. After this proof of concept the implementation guide and the experience gained with the implementation of the prototype will be presented to potential industrial partners in order to implement the productive eHealth platform.

Disclaimer

The information contained in this document describes the current view of the issues discussed until the date of publication. The recommendations and proposal of concepts for a national eHealth platform have to be validated concerning security and technical feasibility. The authors propose the implementation of a prototype as proof of concept.

Introduction

Electronic Health Record Systems store very intimate and private, medical related information of patients. These information are the base for better diagnoses and better treatments. Today, this kind of information is already kept in isolated systems within the so called primary systems like Hospital Information Systems (HIS), Radiology Information Systems (RIS), Laboratory Information Systems (LIS), and General Practitioners' systems. The improvement is to build up a national eHealth *platform* for sharing and exchanging subparts of these information, i.e., the *relevant* health information of each patient.

The content of the patients' national EHR will be defined progressively and in collaboration with the health professionals. Each health professional decides what information he or she wants to sent to the EHR system, and which information is relevant only for their own HIS, RIS, LIS, etc. and will not be shared. Doctor's letters, referrals, medication prescriptions, etc. are today's paper-based examples, where a health professional gives a piece of information (related to one patient) to an other health professional.

An *electronic* sharing of information enables one step more: it includes not only the direct exchange of information, like today's paper-documents; electronic information sharing, based on an IT platform, allows that *relevant* health information is associated to a patient's national EHR and is accessible by other health professionals – immediately and for later use in the future.

The document frame for sharing and exchanging information *is* the electronic health record (EHR) and the IT infrastructure, that enables this technically *is* the EHR system.

With respect to data protection the needs for an EHR system are far beyond of those of isolated systems. This document describes a secure, national EHR system, proposed to be implemented in Luxembourg for the welfare and better treatment of people who live, work, or interact with the Luxembourgish health system.

Document Structure

The document proposes infrastructural and technical details of a national EHR system and its first applications. In more general statements the term *national EHR* is used, while closer to realization of the system the term *eHealth platform* is more appropriate. The focus of this document is on the architecture and its incorporated data protection. The described eHealth platform concept is generic in order to respect current and future political decisions about availability of information, patient IT-consent declarations, doctors rights, new laws, or other future needed adjustments.

The description in this document avoids technical details where possible. Some technical details are explained on a non-technical level because they are important to trust in the technical data protection mechanism. A simple “It is safe, trust us!” is not acceptable here.

The reader should be able to understand why this system is secure and where its limitations are. The document is divided into three parts: (1) general overview, (2) details, (3) implementation steps.

PART ONE is about the general requirements, the benefits, the risks, followed by a description of the cornerstones of an eHealth platform and the selection criteria for an architectural topology of the eHealth platform. Chapter 1 starts with the general requirements description based on a recent study of PriceWaterhouseCoopers.

Chapter 2 sketches the benefits and risks of such an enterprise; so the pros and cons of a national electronic health record. The pros overweight and the cons are mainly in data protection risks.

This leads to the central chapter 3 with the cornerstones of the proposed eHealth platform; the security aspect drives the whole architectural IT design. Three typical kinds of information exchange are explained, then the content structures of an electronic health record including a pragmatic first version, proposed for a first implementation. According to the PwC study the first planned services and features of the eHealth platform are mentioned. A discussion about cross border communication and possible statistical usage of the information follows. The chapter ends with a list of open questions to redirect the reader's attention on the different impacts of Electronic Health Record (EHR) when they come in place.

Chapter 4 raises the general question of centralized versus decentralized system architecture. Advantages and disadvantages are named, selection criteria concerning central or decentral structures are defined. Finally, the decision for a mixed approach with a central registry and central as well as decentral repositories is founded. The busy reader may skip the details of this selection chapter 4.

In PART TWO, Chapter 5 gives an overview of the planned eHealth platform, especially concerning about the security improvement of the existing standards for document exchange and document sharing. Step by step it motivates the topic and the arising tasks, and finally it leads to the proposed solution.

Chapter 6 describes some related work of other countries and initiatives.

Chapter 7 presents a deeper insight into the eHealth platform details. The foundation of the concept based on international standards is explained and the different layers of data protection that form the whole eHealth platform are described more in detail. Besides the basic concepts of pseudonymization and encryption, details about further practical topics are explained. These are archiving and data aging as well as secure accessibility over the web, ID-cards and user management, access logging and alert functions. A pragmatic usage of authentication and signature cards avoids the costly maintenance of a public key infrastructure (PKI) for encryption/decryption with public/private keys.

Chapter 9 gives an overview of the different build-in methods of data protection and ends with the known security limitations – technically, accidentally, and organizationally. The reader will understand the remaining risks and will be able to reflect the high data protection level that is possible with the proposed IT architecture.

Chapter 10 deals with IT-consent, i.e., the allowance of a patient for storing and processing

her/his health data on such an eHealth platform. IT-consent is different from the consent for medical treatments (MT-consent). Nevertheless, MT-consent can be declared as well in the IT system. As this topic is in research state, the chapter notes the current discussion and the first version of an implementation that respects that a patient must be able to declare her/his IT-consent.

Chapter 11 is about the compilation of relevant information in the source systems (hospital, laboratory, medical doctor's office, etc.). The source system is as well the starting point for enabling secondary usage of the data for statistics. Refocusing on the primary usage, i.e., the diagnoses and treatments, the different information pieces of the different source systems are composed into a national EHR of a patient.

Chapter 12 opens the horizon and discusses beyond the eHealth platform about possible storage locations for the health information of a patient. Paper based folders – composed by different doctors and compiled by the patients themselves – are the classical predecessor and in use all over the world. Pocket based electronic folders, i.e., PDF files on USB devices or DVD are a natural successors of the paper based folders. Security in case of lost is a critical topic in both cases, paper and USB or DVD. The preferred storage location is a well protected eHealth platform.

In PART THREE, the more technical chapter 13 describes how the eHealth platform concept will be implemented based on existing standards and well known methods. The so called Integrating the Healthcare Enterprise (IHE) profiles are introduced and it is described how they help to speed-up a realization that will be compatible with these standards. The busy reader may skip chapter 13.

Chapter 14 explains the iterative evolution of the eHealth platform. The aim of an iterative evolution is that a product is available very soon. First versions of the product are useful to gather experience, early versions already fulfill the requirements for productive usage. Next product iterations, respectively, enlarge the set of features offered by the eHealth platform.

Consequently chapter 15 proposes the so called backlog of a first and a second test-bed implementation, and chapter 16 outlines the possible features of the next iterations.

PART ONE

1 Requirements for an eHealth Platform

PriceWaterhouseCoopers defines the requirements for an electronic health record system for Luxembourg in their eHealth Service Platform Study¹ [PwC-study-2010]. They wrote:

For the Platform, we have considered the following generic services:

- *Access management*
- *HC professional register and identification management*
- *Single sign-on*
- *Secure e-mail*
- *Consent management*
- *Master Patient Index and Alias creation to be run by a Trusted Third Party*
- *Centralised catalogues for prescribers*

Santec's proposal of an eHealth platform archives the following requirements:

Concerning the user management:

- All users are identified over a card based infrastructure. Only identity-verifying processes are allowed.
- Access rights of users to the eHealth platform are predefined over the registration of their identities in the user register of the eHealth platform.
- The eHealth platform services are available only according to the users' rights.

The core eHealth platform architecture fulfills the PwC named "Alias creation":

- Medical data are separated from person-identifying data by using a pseudonymization.
- The pseudonymization service is hosted by a Trusted Third Party (TTP).
- A separated organization – called the Agency – hosts the medical information under pseudonyms instead of real patient names.
- The patient's identity is pseudonymized via the TTP's pseudonymization service.
- If the medical information is guaranteed NOT to contain any person identifying information, the simple pseudonymization may be strong enough to ensure the data protection requirements. Here the information is separated in two parts: The identity information for the TTP and the medical information for the core medical system.
- If the medical information could in any way contain person identifying information somewhere else in the "free text" of the medical information, then the patient's identity data are extracted for the pseudonymization. The entire original document –

¹ *Luxembourg Ministry of Health , eHealth Service Platform Study , pre-final version*

with the patient's identity data inside – is encrypted. So the condition NOT to contain any person identifying information is fulfilled by encryption.

- Encrypted data and the decryption key are never available to the same person or same organization inside of the eHealth platform (e.g., only the authorized physician requesting the patient data will receive both, decryption key and encrypted data); the Agency hosts the encrypted medical data while the Trusted Third Party keeps the decryption keys, respectively. This “keying” service is the second task of the TTP besides the pseudonymization service. Alternatively a further organization can operate the keying service.

With regards to the user communication with the eHealth platform

- Information is provided and accessible via secured connections over the Internet.

The technical aspects of the eHealth platform are:

- The eHealth platform is based on the IHE² XDS profile³ with centralized registry and centralized as well as decentralized data repositories.
- This eHealth platform design is enriched by pseudonymization, encryption, transparent re-encryption, and identity guaranteeing signature features.
- The results of a query are readable only by the requester. If a server administrator illegally accesses the communication channel or uses long-term monitors to accumulate the transmitted information, he only gets encrypted data garbage.
- Pseudonyms of patients are never known outside of the eHealth platform. They are never visible for information provider nor for information consumer.

For patient consent and emergency access,

- The patient's declared consents for any access to his data is foreseen.
- In case of no consent for ordinary cases, an access only for emergency cases can be declared by the patient.
- An emergency access should trigger a notification to the patient himself, a friend or relative, and/or his family doctor to inform about the emergency data access.
- Each access to the system is logged. Each information delivery and each information request as well as each modification of the patient's consent declaration. The logs are stored encrypted to avoid illegal usage and modifications.

2 [see <http://www.ihe.net>:] “IHE is an initiative by healthcare professionals and industry to improve the way computer systems in healthcare share information. IHE promotes the coordinated use of established standards such as DICOM and HL7 to address specific clinical need in support of optimal patient care. Systems developed in accordance with IHE communicate with one another better, are easier to implement, and enable care providers to use information more effectively.”

3 [see <http://www.ihe.net/Profiles/index.cfm#IT>:] “Cross Enterprise Document Sharing (XDS) registers and shares electronic health record documents between healthcare enterprises, ranging from physician offices to clinics to acute care in-patient facilities.”

Analogous for any other access like statistics and scientific studies, a Governmental consent⁴ is foreseen, where statistic evaluations are possible on non-encrypted, pseudonymized, medical data. Governmental consent for any statistical accesses to pseudonymized data is mandatory.

Remark on *non-encrypted and pseudonymized information*: The eHealth platform offers this possibility, and the corresponding use case descriptions will determine in which situation this feature is useful and a lower level of data protection is acceptable.

With the requirements of the eHealth platform in mind, the next chapter highlights the benefits of an EHR. Then the risks associate to the implementation of an EHR are shown. This insight motivates the importance of data protection in this sensible sector.

2 Benefits and Risks

Compilations of medical records are the base for decision support systems (DSS). They support practitioners in their diagnosis of diseases and help them to select therapies or medical pathways by taking into account a patient's current symptoms and his available medical history. With the advances of the Internet enabled communication possibilities, medical data can be more easily shared between healthcare professionals.

But this context also has some critical risks, for instance, related to the confidentiality and reliability of medical data. In a beginning phase, an EHR can contain laboratory reports, X-ray and sonography reports, the current and prior medication, known allergies on medication, food and environment, then a collection of discharge letters from hospitals, and vaccination information. In a later phase it may be extended by declaration of living wills and the declaration of organs donation. On the way to personalized medicine the EHR may be extended by DNA information, genetic predispositions, and molecular profiling. At least then it is clear that those data are very confidential and need to be protected against misuse on highest possible level.

2.1 Benefits

It is not the aim of an eHealth platform to accumulate any kind of medical data for some unknown reasons. The outcome must highly improve the patient's care, reduce unnecessary exams (X-rays) and treatments, and as side effect it may also reduce the costs of the healthcare system.

The self-evident benefit of an EHR is the general availability of a patient's health information for all healthcare professionals who are involved in the treatment of this patient. Further, the compilation of a medical record is the necessary basis for decision support systems (DSS). Enabling the usage of decision support systems – based on the cumulated information – is an important and promising benefit of an EHR system. For instance, the patient's private disease

4 A patient's consent declares the allowance to access individual data while a governmental consent allows the usage of information items of a group of patients, i.e., for a special statistic evaluation. A governmental consent may be demanded by an institution for statistical analysis (e.g. CEPS in Luxembourg), approved by an ethic commission and checked by the data protection CNPD.

information can be combined with general treatment and diagnostic's knowledge to obtain a personalized best practice treatment. This is a good contribution compared to what we have today, where a single specialist, with his personal experience and knowledge and based on his notes, analyzes the current symptoms of a patient. With decision support systems, healthcare professionals have the possibility to combine their personal knowledge with the ones entered by other experts on the specific disease domain, to use data from several sources (e.g., labs, hospitals, etc.) and check the availability of limited resources necessary to perform the defined treatments. The expected benefits are the improvement of the quality of care and the best utilization of available resources. In a first step, the DDS are located inside the healthcare providers' IT systems. The required input data is selected and downloaded by the health professionals. Then it gets encrypted and – as it is in a usable format structure – it serves as input data for the local DSS process.

2.2 Risks

EHR information in wrong hands can be catastrophic for a person's future and social life. A person's application for a new life insurance, a house credit, a new job, etc. is in danger. All these critical changes in a person's life can be affected by information out of his EHR. Outer-European companies may offer these kinds of services to banks or private insurance companies. Is the person a “good risk” for an insurer or is his medical history and his genetic predisposition to risky for the good business? The list of risks should not be enlarged towards “Crime 2.0”, i.e., seeking for the best fitting organ donor and force him for kidney donation. Even if these services are illegal in Luxembourg, they may be offered from outside. Illegal information access means criminal administrator or unauthorized user access as well as attacks from outside. The last category may be called “legalized attacks”. This is the case when a government decides to change the laws with the aim to inspect medical information. This afterward legalized access was not intended at the time the information has been provided. It may also ignore a formerly declared and valid patient dissent.

Unauthorized user's actions are prevented by a role management (i.e. giving access rules to the user) together with consent management (i.e. access allowance declared by the patient for each information item).

Criminal administrator access is the danger from inside. The service provider has to trust on his staff and additionally has to avoid accidents if this trust is misused.

Against server attacks from outside a system can be protected according to the state-of-the-art firewalls. If the firewalls are broken, the further attack prevention can be treated in the same category of criminal administrator attacks.

Legalized “attacks” can be prevented, if the patient himself gets a cryptographic key for his information, and it is guaranteed that neither a copy of this key exists, nor an decrypted copy of the information itself.

3 Cornerstones of the eHealth Platform

The outlook to increase the quality for diagnoses and treatments for every patient is very promising. A remaining job is to minimize the risks of data theft and illegal accesses by establishing a very high level of data protection. Besides the technical usage of encrypted communication channels, encrypted databases, or modern encryption algorithms, the security aspects depend on the architecture of the eHealth platform itself. The chapters later in this paper describe in more detail the key topics, addressed with the term “build-in security”.

This chapter will introduce the main aspects of the eHealth platform: First it will sketch the types of information exchange methods used in medical context and considered in this document. Then the first services offered via the eHealth platform are sketched. Cross border communication needs are discussed, and the chapter ends with a list of important questions.

3.1 Three Kinds of Information Exchange Methods

The eHealth platform is designed for information exchange. Depending on the type of information exchange method adopted, there are different solutions to assure the data protection requirements. The three different kinds of communications considered are (1) Provider to well known, intended receiver(s), (2) Provider to unknown⁵ receiver, and (3) Provider to multiple unknown receivers. They are described more in detail below.

1. P2R, Provider P to (known) Receiver R

The data provider knows the data receiver. For example a laboratory sends a report directly to the family doctor who has prescribed the examination. In this case the laboratory result is signed by the biologist in charge, encrypted and then sent to the family doctor. The report can only be decrypted by the family doctor. The whole process could be based on a public key *cryptography* infrastructure as it is used in secure mail systems.

If the other kinds of communication are provided by an eHealth platform (see below), the expensive maintenance of a public key infrastructure for encrypted sending can be avoided. Later it is shown how to base a temporary public/private encryption on an existing *authentication* PKI. In short, if you trust the authentication PKI (e.g., Luxtrust), then a registered receiver can generate a temporary public/private key-pair for encryption and send the temporary public key together with his request when pulling-off the reports out of his folder on an eHealth platform. On the sending side, the data provider encrypts the information with the public key of the eHealth platform, which can be temporary as well. Later more details, but now let's come back to the second kind of communication:

2. P2UR, Provider P to Unknown Receiver UR

The data provider does not know which receiver will finally receive the information. Because of the free choice of the patient, the data provider is not allowed, in some cases, to address an information directly to another health professional. For example: The family doctor sends the patient to a specialist. The patient has the free choice and can not be forced to visit the

5 The unknown receiver may be a specialist that the patient wants to consult. But at the time of providing the report the patient is not sure whether he really wants to consult the specialist. In contrast, if the patient is sure and asks his doctor to address the specialist as well, then the specialist is an intended receiver of case (1).

specialist that the family doctor selects. Family doctor and patient may agree on one, but at the end the patient can re-consider his decision and choose another specialist. If the family doctor has sent already the data to the first specialist, there are two problems: a) the data need to be sent again to the patient selected specialist, and b) the family doctor had broken his professional secret and disclosed information to a non-involved third person.

A more usual example is the free choice of the pharmacist. The family doctor should not send the electronic prescription to a (doctor selected) pharmacist. According to the law of many European countries, the patient is free to select the pharmacy. The electronic workflows must respect this law. The eHealth platform solves it by implementing a P2UR communication method: The data provider provides the information to the eHealth platform and the later selected receiver pulls this information down from the eHealth platform. What happens to data security and encryption? – The provider (family doctor) encrypts the data for “the eHealth platform.” As soon as the later selected receiver (pharmacist) pulls down the information, the eHealth platform re-encrypts the information for the now known receiver (this pharmacist). An important question is: “Is the information disclosed for the time interval of re-encryption?” – Yes, it is. To solve this leak, a secure solution is provided and described in detail later in this document.

3. P2MUR, Provider P to Multiple Unknown Receivers MUR

This is the classical use case for shared electronic records. One data provider produces information for an undetermined number of unknown receivers. In P2R and P2UR the information is stored temporary in the system until the receiver pulls it down; afterward the information can be deleted or marked as processed. For P2MUR the receiver(s) are unknown at the time when the information is provided. Additionally, the information has to stay in the eHealth platform even after it has been pulled down by one or more receivers. This is the basis to build a health record: Because of a large set of different types of information stored for one patient, a structure to organize all information is needed. This structure is called the Electronic Health Record (EHR). Remark: It has to be discussed, independently of the adopted data exchange method, how long the information is kept in the EHR. Concerning data security and encryption, P2MUR and P2UR can be treated very similar. This will be described later more in detail.

The decision which exchange method will fit best to which use case is part of each use case descriptions. For examples:

- Prescription should use P2UR because of the free patient's choice.
- Referral letter should use P2MUR if the use case description states that a referral letter has to be part of the health record.
- A discharge letter from hospital to family doctor can be P2R.
- If the family doctor then decides that the letter is relevant for the EHR, he may add some comments and provide it via P2MUR to the eHealth platform.

3.2 Content of the Records

Building an eHealth platform implies to define an structure of data that is useful for medical purposes, others secondary or tertiary uses can also be done, but it would be better if they were clearly defined before. According to the utilization of the data several medical data structures are proposed. The most frequently refereed are the Electronic Medical Record (EMR) and the Electronic Health Record (EHR). Dave Garets and Mike Davis [GaDa2006] define:

“Electronic Medical Record: [...] is used by healthcare practitioners to document, monitor, and manage health care delivery within a care delivery organization (CDO). The data in the EMR is the legal record of what happened to the patient during their encounter at the CDO and is owned by the CDO.”

“Electronic Health Record: A subset of each care delivery organization's EMR, presently assumed to be summaries like ASTM's Continuity of Care Record (CCR) or HL7's Continuity of Care Document (CCD), is owned by the patient and has patient input and access that spans episodes of care across multiple CDOs within a community, region, or state (or in some countries, the entire country).”

The choice which data structure is necessary depends on different criteria. [Haas2005] proposes five criteria to support the decision process: purpose, subject area, level of digitalization, disease relation, and moderation. These criteria are detail hereafter and compared with the needs of Luxembourg in square brackets [].

(1) The purpose can be primary (patient care and treatment), secondary (invoicing reasons), tertiary (statistical usage).

[Luxembourg: The patient's care and treatment is in focus, statistical usage and invoicing reasons should be possible later on but are not in focus.]

(2) The subject area covers two aspects, the institutional and the medical results case leading to four options: (a) one case in one institution, (b) an integrated patient centered record of all his treatments in one institution, (c) same as b but over all involved institution, (d) same as c and extended with paramedical information and personal notes of the patient himself.

[Luxembourg selects the complete version (d).]

(3) The level of digitalization and standardization is given in three steps:

(a) patient demographic data with a reference to paper based record, (b) medical record that have been scanned into a system, and (c) where all content is electronically structured and part wise formalized.

[Luxembourg: at the beginning all kind of electronic documents are considered, even scanned paper or PDFs files. Nevertheless the aim is a high level of digitalization on HL7 CDA (clinical document architecture) level 3. Decision support systems are best based on structured information.]

(4) The disease relation says whether the record only contains information for one

disease or for all health information of this patient.

[Luxembourg: The records cover all areas, which does not mean all information; it should contain useful extractions of the CDOs' medical records. The eHealth platform will be general enough to respect changing political decisions restricting the content of the records.]

(5) *The moderation of the record can be done by an institution (GP or hospital) or the patient.*

[Luxembourg: The proposed eHealth platform is enabled for both options. Political decisions of the future can be parametrized on demand.]

According to these observations, the recommended data structure to be use in Luxembourg is the EHR. The informational content of the EHR is more detailed explained in another document [EFES-2010-WP10] and will not be in the scope of this document.

3.3 Pragmatic Structure of the Health Records

For pragmatic reasons it is assumed that the incoming documents are stored in a kind of log-book. The log-book is part of the health record. The only sorting criteria of the entries in the log-book is the time of creation or the time of arrival.

A more elaborated structure of the health record has different sections like “case related section”, “radiology section”, “laboratory section”, “current medication”, etc. Each entry in such a section has some extra explanations of a doctor and may be linked to documents in the log-book. Not every chronological log-book entry has to be referenced by such a section link. Some of them, like temporary laboratory reports, are never referenced out of the laboratory section of the health record. The moderator(s) of the health record decide(s) whether a log-book entry should be referenced in one or more of the structuring sections.

Most of those moderator actions should be done automatically by a rule-based record structuring tool, or at least they are automatically proposed for the record moderator to confirm. Every enhancement in structuring the information improves this automation. It is also possible that the patient himself organizes his records – at least inside of dedicated folders.

The next section is about the first services of the eHealth platform.

3.4 First eHealth Services of the eHealth Platform

PriceWaterhouseCoopers proposes a list of services for a first implementation:

“The Platform may host the following value-added services:

- *Electronic Health Record (EHR) consisting of*
 - *eSanté-CARA [authors' remark: the sharing and exchange of imaging related reports]*
 - *eSanté-LABO [authors' remark: the sharing and exchange of laboratory reports]*
 - *a Personal Health Record (PHR)*
 - *Medication Dispense*

- *a Medical Summary sub-service*
- *Hospital discharge letters*
- *Cancer oriented medical record (COMR)*
- *Results server for prescribers of exams providing access through a work list to ordered results*
- *Other important documentation yet to be identified and implemented after 2015*
- *Affiliation Control*
- *Electronic Prescription (ePrescription)*
- *Decision support service (DSS) ”*

[PwC-study-2010]

The short and medium term services riding on the Luxembourgian eHealth platform are:

eSanté-CARA :

- Provide image diagnoses information;
- Display image diagnoses information;
- Provide illustrative images;
- Display illustrative images;
- Download one, more or all radiology reports for one prescriber
- (later) Provide significative images
- (later) Display significative images
-

eSanté-LABO

- Provide laboratory result data;
- Display laboratory reports of one order of one patient;
- Download one, more or all laboratory reports of one prescriber;
- (later) Display laboratory reports of multiple orders of one patient;
- (later) Display graphical history overview of laboratory results of one patient;

Personal Health Record (PHR)

- Provide patient's personal remarks concerning blood pressure, remarks on medication reactions, or any other information he wants to provide.
- Display patient's personal remarks;

Medication Dispense

- The current and prior medication “folder” includes the dispensation of medicaments on demand of a doctor and dispensed by a pharmacist, as well as the OTC (over-the-counter) medication which the patient has bought by himself.
- The pharmacist and/or the patient can insert this information.

Medical Summary sub-service

- Provide Patient Summary with relevant hints on allergies and diseases, as well as references to relevant discharge letters or other documents;
- Download Patient Summary;
- Display Patient Summary;

Hospital discharge letters

- Provide a hospital discharge letter for a patient;
- Download hospital discharge letter;
- Send hospital discharge letter;
- Display hospital discharge letter;

Referral letters

- Provide a referral letter for a patient;
- Download referral letter;
- Send referral letter;
- Display referral letter;

Cancer oriented medical record (COMR)

- Cancer oriented medical records are foreseen as case related records. The restriction of the general EHR towards cancer orientation would implement this option.

Results server for prescribers of exams providing access through a work list

- A medical professional can download one, many or all examination results and reports that are provided for her/him over the P2R communications. This is similar to pop off a mail account.

Other important documentation yet to be identified and implemented after 2015

- Install / Update Medication Databases including available rules for medication interaction checks;
- Display Medication Database Information via search criteria;
- Provide vaccination and organ donation declaration;
- Update vaccination and organ donation;
- Display vaccination and organ donation;

Affiliation Control

- The check whether a patient has an insurance affiliation is important for indirect reimbursement. The check is offered on demand and as automated service when querying documents. In the latter case, a missing or unconfirmed affiliation is indicated together with the retrieved documents.

Electronic Prescription (ePrescription)

- Provide/deliver a medication prescription;
- Download a medication prescription; (i.e.; start dispensing process);
- Provide/deliver prescription of laboratory, treatments, physiotherapy, etc;
- Download prescription of laboratory, treatments, physiotherapy, etc;
- Display prescription of laboratory, treatments, physiotherapy, etc;
- Confirm fulfillment of prescription

Decision support service (DSS)

- The eHealth platform is a mandatory precondition for DSS.

Patient consent declaration [extra]

- Declare patient consent onto the access of his medical data from a GP side;
- Declare patient consent onto the access of his medical data by patient via web;
- Display positive patient consent declarations for authenticated user;
- Declare “living will” from a GP side;
- Declare “living will” over patient web-application;
- Display “living will”; (consent per default to every care provider or nobody?)
- Declare organ donation from a GP side;
- Declare organ donation from patient web-application;
- Display organ donation; (consent per default to every care provider?)

More details concerning the possible structure and content of electronic health records can be found in the deliverable of EFES WP10 [EFES-2010-WP10].

3.5 Cross Border Communication

Different national or regional EHR systems should be enabled for cross boarder interaction. The European projects epSOS and CALLIOPE have been set up to support the cross border communication. For Luxembourg these activities are very important because of the geographical situation and the small number of inhabitants. Medical services of neighbor countries are used quite frequently. Commuters and tourists use health services in Luxembourg too.

Technically, a guarantee for later inter-connectivity is hard to give. But, the design of the eHealth platform cares about this requirements and is based on international standards. An important extension proposed in this eHealth platform that can be the origin of some interoperability problems is the use of pseudonymization on person identifying data, which leads to a higher level of data protection, but is not supported yet by international standards of medical domain.

In epSOS, some efforts have been done to specify an European connector that will provide interoperability between national eHealth platforms. Even if technically the cross border exchange of medical data becomes possible the problem may occur in legal level. What

happens if the security level established in the foreign country is lower than the security level in Luxembourg? In the Luxembourgian eHealth platform, (encrypted) medical data and person identifying data are separated. If the eHealth platform of a receiving country does not support this separation, the foreign country connector (which will be part of the Luxembourgian eHealth platform) has to join the person identifying data with the (encrypted) medical data and construct a document format that is compatible with the European cross-platform interface. Then, one single administrator will have access to patient (encrypted) data and identity. One part of the security strategy presented in this document is broken. The encryption may be kept even on European level, but the pseudonymization may be lost. The patient must be aware of this lack of security when declaring his IT-consent for cross boarder communication.

3.6 Secondary Usage for Statistics

The optimization of the patients' care is the primary usage of every eHealth platform. Statistics are called secondary or tertiary usage. Every additional usage of patients' data implies an additional risk for data misuse. Statistical usage does not support a patient's care at a first glance. The question why to risk a data protection leak is reasonable. On the other side, the outcome of statistical evaluations may improve the medical knowledge for specific diseases. And this feedback improves the care of patients.

Section 5.4 shows that statistical usage can be enabled without opening a wide leak in data protection, and section 7.4 explains the details.

3.7 Open Questions

Questions are still open and have to wait for political decision even on EU level. The proposed eHealth platform is generic enough and prepared to cover several situations. A frequent change of the legislative environment can be expected, so that adaptations and fine tuning concerning the eHealth platform behavior will occur. Having this in mind, the system layout was defined.

Most of the questions below are taken out of publications of the German data protection specialists Mann and Engels [Mann2008].

- Who compiles a selection of information of the EMRs?
- Which structure and which content is appropriate?
- Who is allowed to access those informations?
- Under which circumstances? (emergency, with consent, ...)
- Is the patient allowed to access the EHR without a doctor?
- Should the patient define the access rights?
- Who is responsible for the composed EHR?
- Who decides what is useful information?
- If the patient deletes parts of the record the medical content has been changed. Who is afterwards responsible with respect to medical completeness?

- Who is responsible with respect to data protection laws in case of a data accident?
- Who can change the contents in case of obviously wrong information?
- Who should lock informations that are detected as wrong?
- Who decides that an information is outdated?
- What happens if the patient want to keep this outdated information?
- The deleted part has already been overtaken into a primary system, what to do?
- How long should overtaken data reside in a primary system?
- Should a primary system view the data or import and store them?
- The context of a document has been deleted, what to do now?
- Is an update notification service necessary where the eHealth platform notifies all prior information consumers about the updates or corrections?
- Is every practitioner required to read all information for medical reasons?
- Is a practitioner responsible for a wrong treatment, just because he did not read all and every information that was provided in a very large EHR of a patient? Wrong treatment despite of clear warnings “deep” inside the record documentation?
- How many has the practitioner to read for legal aspects to avoid compensations?
- Who is allowed to delete what – or how “deep”?
- Who is liable for which part of the data processing?
- Who is liable for the whole record? (secrets carrier)
- Who is liable in case the patient has ordered to delete something?
- Who guarantees which level of trust for the contents? To whom?

Mann states that if no one is liable for anything in context with the records and the record system, the necessary establishment of trust and acceptance is questionable?

Depending on the later political decision the systems should be parametrized in the way that those decisions are implemented by parameter-settings. On one side the system can only go live when all decisions have been taken. On the other side a “playground” is needed to get conform with the questions in awareness of a system. Some of these queries are already addressed within the current projects in laboratory and radiology. And first answers have been found already in the corresponding work groups. The solution for this is the conceptualization and creation of a “testbed” system, present this to the politics, discuss, modify, adapt, and then find answers to the remaining questions. As the projects in laboratory and radiology show that already the discussion towards realization of use cases – with the eHealth platform design in mind – shows up some consistent answers.

4 Technical Concepts of the eHealth Platform

The design of an eHealth platform raises the general question whether the system should follow a central approach or should it be a decentral one. This chapter gives arguments for both variants and find an appropriate solution, which is as well a standardized one.

4.1 Centralized or Decentralized Repositories

In the usual definitions, a centralized repository means that all information of the whole system is central at one location. Decentralized repositories means that every information stays in the location of its origin. Central solutions have the trend to be very attractive for hacking attacks. On the other hand central solution can invest more in security than multiple decentralized servers in the doctors' offices can invest.

For hospitals this is different again. The choice of a centralized storage implies always copies of the information over the net into the central data repository. For large amounts of data decentralized solutions avoid the net traffic at data providing time. Nevertheless these data will be copied out of the institutions' primary system into a – let's call it – staging area. The staging area presents the selected information to the outside world. Normally such staging areas are operated in a so called demilitarized zone (DMZ) with lesser access restrictions and a more open firewall than it is for the original primary system. Decentralized solutions generate the network traffic only at access time. Centralized repositories generates the network traffic at the time of providing the information to the eHealth platform as well as at the time it is accessed later.

Any massive data access (e.g., for statistics) on decentralized information tends to be very expensive in network traffic. That means the application's waiting time increases massively for the users up to the point where it is impossible to use such an application.

4.2 Centralized Registry – With or Without ?

The next question concerns about the searching and finding of information: a centralized registry is an index for any information that is stored somewhere in the systems. For the registry it is independent whether the information is stored in a central repository or in the staging areas outside in the hospital DMZ's. The registry keeps an index with meta data for all information that is registered and accessible in the system. The registry knows where this information resides: in the central repository, in a staging area of a hospital, or even (just to bring in the idea here) in a national backup PACS as currently discussed by the hospitals in Luxembourg.

Systems without a central registry are also possible for the cost of expensive broadcast queries. A system without a central registry forces any query to be sent to each decentralized node in the system. An example shows the difference: The family doctor Dr. Help is searching an X-ray of the lung of his patient Mr. Bee.

In the case with a **central registry**, Dr. Help accesses the central registry of the eHealth platform. (a) The right Mr. Bee has to be searched, (b) the consent of Mr. Bee to access his X-ray has to be checked. (c) If there is a registered X-ray of Mr. Bee the IT-system of Dr. Help gets the storage location. This may be a central repository or a staging area of an

hospital. In any case Dr. Help's system knows where the storage location is. Now (d) the IT-system of Dr. Help can ask for a copy of the picture.

In case **without a central registry**, Dr. Help's system has to ask every known remote system or staging area the above actions (a) to (d). And if a remote system is off line, Dr. Help's system may assume that this remote system do not have any information about Mr. Bee's lung.

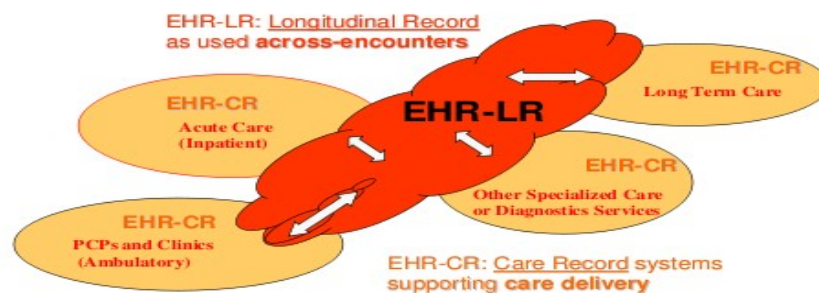
A centralized registry or index seems to be the only acceptable solution with respect to reliability, performance, matching complexity, and network load.

For repositories the question of decentralized or centralized depends on the application, the amount of data, the trade-off between data providing and data downloading, and on security aspects. The security aspects we will care later on. Now, ignoring the security aspects for a moment, the general architecture enables both options – centralized and decentralized – for the repository, while the registry is a centralized one.

The next section describes infrastructure options in detail by referring to the IHE XDS documentation.

4.3 IHE Cross Enterprise Document Sharing (XDS)

The figure below shows the so called longitudinal record, an EHR consisting of subsets of the information of different care records. The following figures are copies of the original IHE documentation. To stay as close as possible to IHE the figures are not changed but explained in the context of this paper. The EHR care records (EHR-CR), which are used in the figures, represent the information based of the care institutions. The red painted EHR-LR Longitudinal Record is the national EHR.



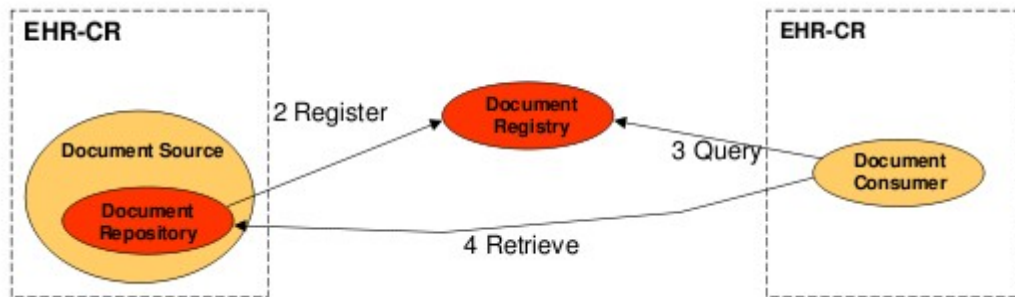
IHE Figure: Care Records and EHR

The suggested implementation strategies are nearly self explaining. IHE lets much freedom in implementing this XDS profile. The main use case actors are the Document Registry and the Document Repository. Four implementation possibilities are shown below. We use them also as different options for a system architecture. To stay in-line and consistent with the IHE nomenclature the used figures are copied out of [IHE-profiles].

4.3.1 Decentralized Document Repositories and Central Registry

The first figure shows a central Document Registry where care delivery organizations register

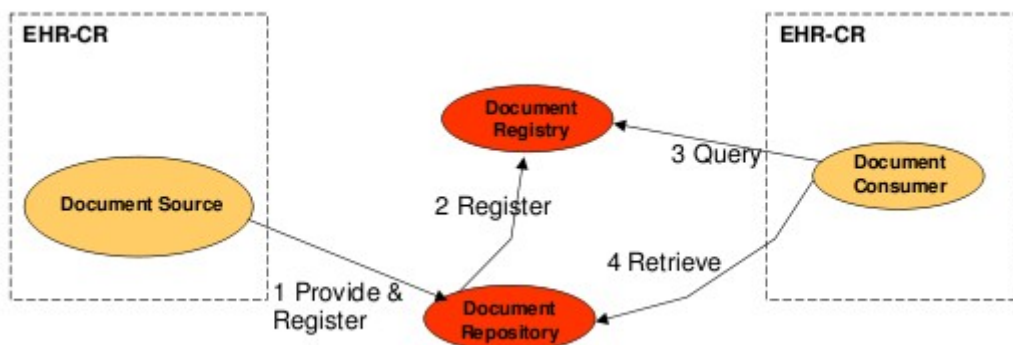
their provided documents. The documents themselves resides in the domain of the Document Producer. A Document Consumer queries the Document Registry and then retrieves the desired document from the Document Repository at producers side. If multiple producers are involved, and every producer acts in the described way we call this a Decentralized Document Repository with a Central Registry.



IHE Figure: Decentralized Repository, Central Registry

4.3.2 Central Document Repository with independent Central Registry

Here the Document Producer provides the document to a central repository while the central repository further registers the document in the central Document Registry. Like in the case above the query goes to the Central Document Registry which tells the location to the consumer. The consumer then accesses the document at the Document Repository. This approach is called a Centralized Document Repository with independent Central Document Registry.

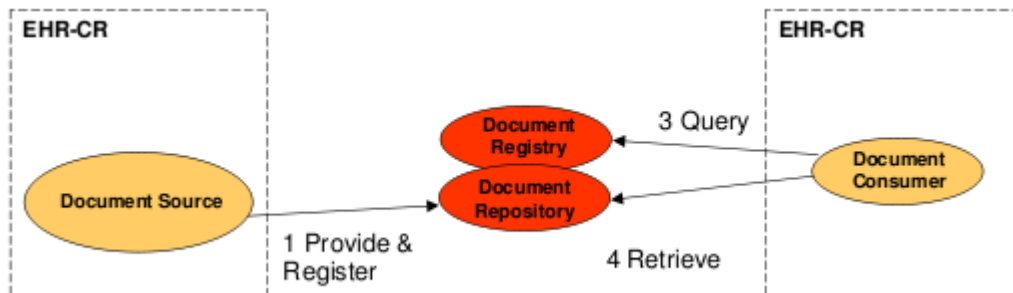


IHE Figure: Central Repository, Independent Central Registry

4.3.3 Central Repository with associated Central Registry

The next figure shows the central Registry and the Central Repository sharing the same physical location. This is equivalent to the case above with the difference that the Registry and the Repository are on the same server location. Therefore the registry is called associated,

and the first two steps are summarized to one single step. i.e. the registration is implicit, and the number 2 falls together with number 1 in the figure.



IHE Figure: Central Repository, Associated Central Registry

4.3.4 Consumer Centric / Push Everything to Consumer

Here the Document Producer “provides & registers” the documents directly into the domain of the Document Consumer. It is comparable with the case above with the bigger difference that the consumer hosts in fact the eHealth platform.



IHE Figure: Consumer Centric

Based on this examples of system design concerning registry and repository and registry and multiple repositories, the next paragraph describes the criteria to select an appropriate architecture. It should be mentioned, that even with a central registry, the information could be mirrored to several system, located in different locations around the country.

4.4 Evaluation Criteria for Architectural Topology

The previous section shows the implementation options for the IHE XDS profile. They all share the idea of a central registry. The location of the repositories – that means the locations where the clinical documents, the referral letters, the X-rays, etc. are stored – differ from one implementation option to the other.

This section lists selection criteria for finding the appropriate system architecture. Each criteria is described with this intention. A short statement at the end of the description shows the implications of that criteria for the eHealth platform design.

The busy reader may skip this argumentation chapter and continue with PART TWO.

4.4.1 Evaluation Criteria: Applications using the eHealth Platform

Every application needs a special view on the information that the eHealth platform provides. If the internal organization of information in the system matches the needs of an application well, the application can be implemented more efficiently than in cases where the information is organized with other aspects in mind. This may be compared with a plain Excel sheet: If the sheet is sorted patient wise, a patient centered application will be efficient. If the sheet is sorted by diagnoses and treatments, the name of the prescriber, the involved hospitals, the age of the patients, the disease relation, or whatever, the appropriate application eventually needs a resorting or pivoting of the sheet. The evaluation is possible – but maybe slower in performance and complexer in programming.

The information of an Electronic Health Record is stored in the eHealth platform's databases “patient-centered” because this is the main focus of an EHR. Nevertheless, some very useful applications will provide a doctors work-list view in the way of: “Give me all open laboratory examinations where I am the prescriber.” This is not a patient-centered application. The realizations needs a regrouping of the information which may result in a small performance disadvantage. Database technologies, like additional indexes on the secondary usage of information may circumvent this disadvantages later.

Applications are about how to retrieve (back end) and how to present (front end) the information that is stored inside the system. Use case descriptions bridge the gap towards the details of eHealth platform applications. Some work groups are already elaborating these details for laboratory and radiology domains in Luxembourg.

Summary: The “application” criteria alone does not imply a special structure of the system. Every stored information is accessible and presentable.

4.4.2 Evaluation Criteria: Haas' Five Dimensions

The Five Dimensions of Haas are already mentioned. Here is the question weather the Luxembourgian answers to the dimensions imply advantages for a special architecture. That means: What do the dimensions (1) purpose, (2) subject area, (3) level of digitalization, (4) disease relation, (5) moderation – imply for the eHealth platform topology?

An integrated patient centered record with the primary aim of patient care and treatment must be available on demand. All relevant information must be accessible without the risk that some repository server are not online at the time the information is needed. If the availability is assured, then decentralized repositories are acceptable as well.

The information should be centralized patient-wise. Even a personalized memory (USB

device) in patient's hands fulfills this criteria. Same works with a personalized access token (like an ID card) to a personal information storage, which resides on a server somewhere and which is accessible and on-line 24 hours on 7 days a week – like the USB stick is.

Summary: The Luxembourgian selection concerning Haas' five dimensions may show a trend towards a centralized storage of important patient information. Moderation of the record by patient and doctor indicates a web hosted system.

4.4.3 Evaluation Criteria: Technical Reliability and Availability

The whole system needs to operate in a technical reliable way. With availability the over all availability of the whole system, respectively the availability of all of its components is addressed. The complexity should be reduced to the necessary minimum. As a central registry is mandatory for efficient finding of information in the eHealth platform, an associated central repository on the same hardware may decrease the complexity of the hardware and net infrastructure. A redundant hot standby solution for the whole eHealth platform is less complex if only one machine has to be redundant.

Summary: Concerning technical reliability a central repository is preferred.

4.4.4 Evaluation Criteria: Reliability with respect to Content

Primary systems of the providers (hospitals, laboratories, ...) are the sources of information. Delivered information which gets detected as wrong has to be updated, deleted, or renewed as soon as possible. If outdated information is stored in a centralized repository, the update of the information may take longer time than having the information on producers side in the “staging area” repositories. A consumer will search for the information in the central registry. As result the consumer gets a location where the wished content can be found. If this location is a staging area it can be updated quicker.

Summary: Concerning the reliability of the content of an information, the decentralized “staging areas” as repositories may fit better.

4.4.5 Evaluation Criteria: Security and Vulnerability

A centralized system may be in focus for attacks from outside and inside. It can be protected by state-of-the-art firewall techniques and a well trained IT-team. In case of a data accident, all information of the system are disclosed. On the other side, the information nodes of decentralized systems can hardly be protected all in the same way compared to one central high security solution. But in case of a data accident only a part of the information is disclosed with a decentralized solution.

Summary: Concerning security and vulnerability, decentralized locations of repositories

decrease the attraction for get-it-all attacks. But decentralized locations are harder to protect: tie here.

Remark: In this document an extended approach towards a pseudonymized information storage is shown. The information items are separated in a way that even if one part is disclosed the information is useless for the attacker.

4.4.6 Evaluation Criteria: Patient Consent Declaration is Mandatory

The patient must have the possibility to declare and define his consent concerning any information access to his EHR. If this declaration has to be possible via Internet, this criteria gives an indication for a web-based system. Consent management will be part of the meta data of each information item that is registered in the central registry of the eHealth platform.

Summary: Patient consent declaration via Internet implies a web-based system. It needs a central registry. For the repositories there are no favorites.

4.4.7 Evaluation Criteria: Total Cost of Ownership

The costs (buying or leasing) are those for the servers, the network, management to negotiate contracts, management for setting up a support, management of teams, and not to forget costs on end-user side. The results that Gartner publish for ERP systems [Gartner2003] may as well hold for an eHealth platform. “The Simple Answer: Decentralized Solutions Have Higher TCO”

Summary: The more decentral the more expensive.

4.4.8 Evaluation Criteria: Response-Time

A good response-time of the system is mandatory for the acceptance of the users. On the one side a system where the services are divided on different sub-services and those are spread over different servers, this kind of decentralization may increase the response-time. On the other side redundant structures may result in a higher decentralization but with the aim of shorten the overall response-time.

Summary: Response-time does not vote for or against decentral structures.

4.4.9 Evaluation Criteria: Data Exchange with Foreign EHRs

In context of cross-border applications in the spirit of epSOS and CALLIOPE the compatibility with other regional and national EHR projects has to be assured. If the proposed national eHealth platform has a higher data protection, it is still an open question whether the security level can be downsized for the foreign communication. As the proposed data protection here includes the pseudonymization layer – and as such differs from other countries solutions – this criteria has no implication on the eHealth platform layout. A special patient consent declaration and a government given default value for this consent is necessary

for cross-border communication. The patient explicitly has to allow this type of communication with reduced data protection. Additionally it is clear, that an access from foreign EHRs never allow the access of the entire set of data, but only to a prepared subset, which is especially tagged with the “foreign access” consent.

Summary: Concerning the eHealth platform this criteria does not favor a special topology.

4.4.10 Evaluation Criteria at a Glance

As summary of the above discussion the evaluation criteria are shown in a table.

Criteria	eHealth Platform Topology Indication	Central Repository	Decentral Repository
Applications Using the eHealth Platform	No indication for special topology.	0	0
Haas' Five Dimensions selected for Luxembourg	Trend towards centralized repository, web-based.	+	0
Technical Reliability / Availability	Central registry is preferred.	+	-
Reliability with Respect to Content	Trend towards decentralized repositories.	0	0
Security and Vulnerability	Tie: decentralized locations are harder to protect. Central locations are in focus for attacks.	0	0
Patient Consent Declaration is Mandatory	Web based, no indication for repository.	0	0
Total Cost of Ownership	The more decentral the more expensive	+	-
Response-time	Response-time does not vote for or against decentral structures.	0	0
Exchange with foreign EHRs	Implies a minimum of data protection.	0	0

This chapter is about the selection criteria for a system architecture. For the question whether a central repository or multiple decentralized repositories are better some criteria are listed. There are three groups: (1) neutral in this question, (2) for a central repository, or (3) for a more elaborated topology with central and decentralized repositories. Concerning the registry, there is no decentralized alternative.

For the repository some use cases may benefit from a central repository while other may benefit from a decentralized repository. The architecture proposal in this document allows both: central and decentral repositories and let is to the use case to decide which one appropriate. The use case designer may notice the results and decide for the best fitting approach, respectively. Fortunately the same options are given by the IHE XDS profile. As conclusion the IHE proposed solution is an appropriate basis for an eHealth platform. It is extended by some security features. The main components of the proposed system architecture are:

- A central registry
- One or multiple central repositories (according with IHE XDS)
- One or multiple decentral repositories (according with IHE XDS)
- A central pseudonymization service operated by a Trusted Third Party Provider
- A central re-encryption service operated by the Trusted Third Party Provider

Details of the pseudonymization and encryption are explained in the next chapters.

PART TWO

5 Encryption and Pseudonymization meet IHE XDS

This chapter motivates towards a secure solution for a national eHealth platform. Step by step it describes the arising open tasks, while the respectively following section solves these tasks. The different techniques themselves and their contribution to the final solution are introduced and explained in a certain order. Combination and interaction of these techniques are important.

5.1 Separation, Encryption and Pseudonymization

A patient's record contains multiple medical documents of different medical domains. A first approach can be the partitioning of the health record over different repositories, according to the different medical domains, i.e., for each medical domain an other repository. An attack will only disclose medical documents of the involved repository.

A better approach for information separation is encryption. Encrypted documents are stored in a repository and the decryption keys are kept elsewhere – separated!

Both approaches can be combined.

Remark: The incorporated security concept of the architecture for the eHealth platform does not depend on point-to-point channel encryption techniques (https, ssl, etc.) as it has its own user-to-user encryption based on certificates. Nevertheless, existing networks or server connections with implemented point-to-point channel encryption can continue to be used.

Advantages of Decentralized Systems

The security advantage of decentral systems is that when one location is attacked and cracked, the attacker only gets a part of the information. For example patient records may be partitioned by clinical domains, and each decentral server carries information of one medical domains only. If such a decentralized server is hacked, not the whole patient record is disclosed. But nevertheless, depending on the disclosed domains the negative implications for the patient are between grave and very serious. The security problem is:

(1) Parts of the eHealth records are disclosed.

Separation of Encrypted Documents and Decryption Keys

The idea is to partition the information in a way that one single part is useless or unreadable without knowing the other part. The reader may note that this is a step further than the usual decentralized storage with respect to medical domains for example. Instead of domain partitioning over multiple decentralized servers, the basic version of this concept is operated by two separated servers, hosted at two separated organizations. One organization cares about the encrypted medical documents and the second organization stores the decryption keys for the documents. It is clear that an attack to one of the servers will not disclose anything. As

well it is clear that encrypted documents can be exchanged without a danger of disclosure⁶. As long as a receiver – legal or illegal – does not have the decryption key, the information is valueless. The outstanding task resp. the *problem* is:

(1) We need a secure transportation method for the decryption keys.

Transportation of Decryption Keys by the Patient

A solution to transport the decryption key (halfway) secure toward the intended receiver is the patient. He can carry the decryption key. The information provider (doctor, hospital, etc.) hand over to the patient the decryption key and provides the encrypted information to an eHealth platform. The patient hands over the decryption key to a next health professional. Decryption keys and server locations are printed as bar code or stored on the patient's health card. A clear advantage is that the IT-consent to access the information by the next health professional is implicitly declared by handing over the media with the decryption key. The *Remaining Problems* are:

(1) The media may have been forgotten, or the media has been lost.

(2) In case of an electronic media the patient may have forgotten the access code (PIN), or in case of printed media the bar code may be unreadable because the paper has been folded and then transported in the patients trouser pocket.

Transportation of Decryption Keys by using a Public Key Infrastructure (PKI)

An other solution is a public key infrastructure (PKI) for cryptography. The information provider (a doctor) knows the intended receiver and encrypts the medical information with the public key of the receiver (an other doctor). Only this receiver can decrypt the medical information. The disadvantage is that many workflows for medical documents do not allow a direct addressing of the receiver, because the patient has the free choice of pharmacy, laboratory, hospital, specialist, etc. The patient can decide about the next steps for his health – explicitly after his current consultation. So the medical information of the current consultation has to be provided without knowing the next receiver(s). This trend disqualifies the direct addressing of receivers. The *Remaining Problem* is:

(1) This workflow does not respect the free choice of health professionals by the patient.

Both approaches “transport of decryption key by patient” or “PKI encryption” do not fulfill the requirements of a national eHealth platform as permanent storage of medical information. A patients record contains a set of medical documents. On the one hand a single decryption key for the whole record violates most data protection rules. The disclose of such a master key would allow the access for each health professional that has been involved once, even years ago. An involved health professional may additionally forward this master decryption key of the whole record to other people – without any control by the patient. On the other hand the patient can not keep all corresponding decryption keys.

⁶ If we assume that the cryptography algorithms are secure.

Decryption Keys are stored at a Trusted Third Party (TTP)

The intended receiver(s) may be unknown at the time when the medical information is provided. This assumption respects the patient's free choice. If so, the public key of an unknown receiver can not be used. Instead the public key of a so called *Trusted Third Party* (TTP) is used for encryption. Later, when a potential receiver requests information, the receiver is known! Before delivering, the TTP will re-encrypt the requested documents with the public key of the (now) known receiver. The remaining problems are:

- (1) How can a health professional search for the patient's records on encrypted documents?
- (2) The administrator/intruder of the TTP can access every medical information during the re-encryption.

Enabling Search on the eHealth Platform

The search for patients' data is not possible on encrypted data on the platform. A solution is an appropriate preparation of the information on provider side. On provider's side the patient identifying data and some useful meta-data are copied-out of the information before the encryption of the medical report. An annotation of the readable patient identifying data to the encrypted medical report enables a health professional to search for information about the patient.

Unfortunately this enables the search possibility on the register as well for administrators/intruders. A first attack to get the medical information of a dedicated person (e.g., a famous politician) is to search for this patient. Then the encrypted medical information can be stolen and the criminals can try to decrypt this data-block "at home". So, we have solved the search task but we got a new problem. The remaining problems are:

- (1) New: Search is enabled for hackers.
- (2) Remaining: The TTP administrator/intruder can access everything during re-encryption.

Pseudonymization and Two-Stroke Encryption

If *pseudonymization* and a special *two-stroke encryption* technique are combined in a certain way, then the two problems of above are solved finally: hackers can not search immediately for a dedicated patient, and the TTP administrator/intruder can not access any information during re-encryption.

For a real *pseudonymization*, we propose an organizational separation of patient identifying data and medical data. One organization holds the patient identifying data together with a pseudonym, and a second organization holds the encrypted medical data associated to the pseudonyms.

The TTP gets a second job besides the decryption key storage, namely the *pseudonymization service*. It generates and provides pseudonyms for patients. The organization that stores the (encrypted) medical data under these pseudonyms is called the *Pseudonymized Medical Information Provider* (PMIP). With this separation into TTP and PMIP the eHealth platform consists of two independently operated system parts. Now the hacker-search problem is more difficult for the hacker. A direct attack to search for the records of a dedicated patient at the PMIP is useless. First the hacker needs to know the pseudonym of the patient, which means

he needs to start first with an attack against the TTP, then an independent attack against the PMIP to find the encrypted data-block, and finally a brute force attack to decrypt the encrypted medical information.

With the existence of two parties, TTP and PMIP, the disclosure of the medical information during re-encryption is avoided. Secure re-encryption is possible with a two-stroke encryption and two-stroke decryption. The trick is to have only one stroke in the re-encryption phase at the TTP.

Data providing

A *data provider* (e.g., the laboratory IT side) generates a new symmetric encryption key for each medical information-item it wants to provide to the eHealth platform.

- Stroke 1: Data provider encrypts the information with the generated symmetric key.
- Stroke 2: Data provider encrypts the symmetric key with the public key of the TTP.

The encrypted symmetric key and the encrypted document is provided as a tuple to the PMIP. The private key to decrypt the symmetric key can be the TTP's private key.

Data retrieval

During the *retrieve process*, the re-encryption is protected against an administrator or intruder. The requester splits the query – patient identifying data to the TTP and rest of the query to the PMIP. PMIP receives the pseudonyms from TTP, and the retrieval can start. For re-encryption PMIP sends the encrypted symmetric key and the public key of the requesting receiver to the TTP. (Remind: PMIP does not send the encrypted medical document to the TTP.) Then the TTP decrypts the symmetric key with its own private key and re-encrypts the symmetric key with the public key of the receiver. TTP sends back to PMIP the re-encrypted symmetric key. Finally, PMIP provides the (still) encrypted document itself together with the re-encrypted symmetric key of the document to the receiver. On receiver side, the two-stroke decryption is done transparent for the user:

- Stroke 1: The legal receiver decrypts the symmetric key.
- Stroke 2: With the symmetric key he decrypts the medical information

There are no remaining tasks or problems.

5.2 Saving the Costs of a Cryptography PKI

A PKI infrastructure means that public/private key pairs are created, maintained, updated, revoked, etc. by a trustful organization. If a provider delivers information to the platform, the medical data is encrypted with a generated symmetric key and then the symmetric key itself is encrypted with the public key of the TTP. If the platform delivers information to a receiver, the encrypted medical information is forwarded untouched by the PMIP. Solely the encrypted symmetric key is re-encrypted by the TTP with the public key of the (now known) receiver. Only the intended receiver knows his private key and can decrypt the symmetric key. With the symmetric key he can decrypt the medical information. Here is the question whether we

can save the costs for a PKI for cryptography.

A PKI infrastructure for authentication and signatures is already in place (e.g., Luxtrust) and part of the security shell that prevents unauthorized access to the platform. Based on this existing PKI, a temporary cryptography key-pairs can be generated for the information exchange. That means, the system trusts on the given authentication and signature features of the certificate creator.

Data providing

An information provider (laboratory, hospital, etc.) asks the TTP for its current public key. This request is signed by the information provider (signature PKI used) and the information provider is logged in properly (authentication PKI used). The TTP will provide a public key, and the information provider will use this instead of the TTP public key of a cryptography PKI. The delivered public key is signed by the TTP (again signature PKI is used). The TTP keeps track of time and corresponding public/private key pairs. The rest stays the same.

Data retrieval

For retrieval, an information requester sends its temporary public key to the PMIP. PMIP uses this public key instead of a PKI provided public key of the receiver. The rest stays the same. Again the query of the requester is signed (signature PKI) and the receiver is logged in properly (authentication PKI).

With this slight modification, the encryption/decryption is solved without maintaining a costly cryptography PKI. It is based on an existing user-authentication and signature PKI.

5.3 Security Shell and Possible Extensions

The eHealth platform is protected against unauthorized access by multiple security levels. The initial login is done with a personal certificate, for example with an ID-card. A user&role directory guards the legal access rights to the system. Requests of potential receivers have to be electronically signed by the requesters. In fact they send their public keys signed and together with the requests. For user authentication and signature a certificate based solution is in place.

In future it may turn out that two organizations (TTP, PMIP) are not trustful enough, or that databases can get stolen and joined later. Therefore two improvements should be mentioned below.

Scheduled Pseudonym Exchange

An even higher security level for data protection can be reached by *scheduled pseudonym exchange*. The pseudonyms will be exchanged on a regular basis each hour or if necessary in shorter intervals. The stolen mapping table only works if the second administrator has stolen the medical databases during the same time interval. If this extension gets necessary further elaboration concerning the switching time has to be done.

Multiple Pseudonymization

To further enlarge the trust level, *multiple pseudonymization* steps are possible. The first pseudonymization service maps real identities to pseudonyms. The second pseudonymization service maps the first pseudonym to a second pseudonym. The n-th pseudonymization service maps the (n-1)-th pseudonym to an n-th pseudonym. Each pseudonymization mapping must be hosted by an independent trusted N-th party. A combination of scheduled pseudonym exchange and multiple pseudonymization with different pseudonym exchange intervals of the different levels is possible.

5.4 Statistical Usage under Pseudonymization

There is always a trade-off between statistical evaluations and the aims of a very high level of data protection. With the primary usage, the care in mind, statistical usage is acceptable because the outcome influences the medical knowledge and as such implies a better quality of care in future.

For qualified statistical usage, the relationship of medical information with a named patient is not necessary. Essential is the knowledge that different medical information of different sources belong to the same patient. Not the patient's identity but the fact that it is the same patient is relevant. This requirement is fulfilled when the corresponding pseudonyms of one patient can be grouped and their corresponding medical data can be used coherently for the statistics. A violation of patients' privacy protection has to be excluded as far as possible. This is the trade-off that is mentioned above.

For the primary usage (i.e., care) the documents are encrypted as whole for PMIP storage. The pseudonymization and its mapping by the TTP enables the search for a patient. This constellation guarantees a high data protection against misuse of administrators or intruders with administrator rights. The two stepped encryption of the documents ensures a secure re-encryption for a legal receiver.

To enable statistical usage the whole document has to be stripped-off from any patient identifying data. To be clear: for care usage the document as whole is encrypted, and for statistical usage a duplicate of the document is stripped-off from any person identifying data. The pseudonymization step is the same for both. Later, for care usage, the TTP knows all pseudonyms of a patient. And for statistical usage, the TTP will inform that different information fragments belong to the same patient. Needless to mention that this is without disclosing the patient's identity to the statistical requester.

It should be mentioned, that this system constellation even enables real-time statistics. If a statistical evaluation is covered by a governmental consent, the corresponding data fragments are collected in real time. In case of epidemics or other sudden health related events real-time statistics can be very helpful.

5.5 The Minimal Constellation for Security

The architecture combined in the proposed way – without scheduled pseudonym exchange and without multi pseudonymization – is the minimal starting constellation to establish an appropriate data protection. Concerning the option of additional statistical usage, the data

provider has to guarantee the correct erasing of any person identifying data. For care usage this guarantee is given by the encryption of the whole document. Any modification can lead to a security leak and need to be deeply analyzed before it can be recommended.

6 Related Work

Some references to related projects in other countries shows that the proposed eHealth platform concept is not new or unique. Most of the related work eHealth platforms focus on the communication of information from a provider to an unknown receiver. The real receiver turns out later when the patient decides which health professional he likes to choose for his further treatment, or which pharmacy he may choose for dispensing the medication. Our approach focus on the EHR system, i.e., on the storage of information that can be used later by every professional who is allowed to use it.

6.1 Québec's Dossier de Santé

In a project in the region of Québec [Quebec2011] the data security explanations of the “dossier de santé” are described: *“The information accessed through your EHR is spread out, in **encrypted form**, in databases that are located in different areas of Québec. The same goes for information on your identity: the information that makes it possible to identify you is kept separate from the information on your health. This makes it possible to keep the information incomprehensible, should an unauthorised person try to intercept it. Only when an authorised health care provider consults your EHR does all of the information “come together” for display on his or her screen in a readable format.”*

This approach seems to be very similar. As the publication is very new, we could not find enough information to verify.

6.2 Belgian Approach with TTP

Robben describes an eHealth platform that uses a TTP as depot for cryptography keys. Again a main difference to our approach is that a doctor can not search for a patient. This is not intended in the use cases there. Remind: In our approach, we take a copy-out of the patient-identifying data and the document's meta data to enable different search possibilities. In the Belgian model this search functionalities are not in focus. The secure transport of a medical information is in focus there. The access rights to that information is given with a token that may be transported by the patient physically or electronically. The workflow is described with a good figure in [BE-GOV2010] and can be summarized as follow:

- The data provider asks the TTP for a symmetric key.
- With this key the provider encrypts the whole medical document.
- The encrypted document is transferred to the so called message depot (comparable with our PMIP repository).
- For retrieving documents, a potential receiver justifies his access rights.
- Then he receives the symmetric key from the TTP's key depot, encrypted with his public key.
- From the message depot the receiver gets the encrypted medical document. Up to now

he has no information what the document is about.

- The receiver then decrypts the symmetric key by using his private key.
- With the gained symmetric key he decrypts the medical information.

In summary, the workflow is very similar to our proposition. It describes a secure transportation for the case that the receiver was unknown at the time when the document has been provided. Compared to our approach, in addition to this secure transport, we allow a search for a patient with his identifying data against the TTP. Furthermore a search is possible for special documents of that patient – by using the provided meta information of the encrypted document. These search functionalities are essential for our understanding of an EHR.

6.3 German Approach D2D

The Doctor-to-Doctor (D2D) communication service is a successfully running eHealth platform for health professionals' communication in Germany [D2D-Padok]. D2D is based on the three kinds of communications that we introduced in section 3.1 as P2R, P2UR, P2MUR. Honestly, we were inspired and influenced of this classification and overtook it. The P2R (provider to receiver) way is realized by encrypted mail communication, where the public key of the intended receiver is used for encryption. D2D calls this “directed” communication. D2D maintains a cryptography PKI for its users.

If the receiver is unknown at the time when the medical document is created, the system on provider's side generates a key, encrypts the medical document with that key - in a special patented way - and prints a part of that key on a paper-ticket which will be handed over physically to the patient. D2D calls this the “non-directed” communication. The patients physically selects the receiver, i.e., an other health professional, and hands over his ticket with the encryption key. With this key the selected receiver can query for the document on a D2D server, receive the document and decrypt it. Instead of non-directed communication, we prefer the name P2UR for provider to unknown receiver(s).

The interesting communication is about the creation of a health record on the eHealth platform. At D2D there is one document provider and multiple unknown document receivers. Any participating health professional can create a so called D2D net-folder for the patient. He can depot documents in this folder. The documents are encrypted with a combination of a D2D server key and a key that is printed on the ticket for the patient. In that way the document in the net-folder is protected against illegal administrator accesses or intruders. Potential receivers have to justify themselves as legal receivers with showing some credentials from the handed-over paper-ticket. The medical information is re-encrypted for the receiver. To decrypt the medical information the receiver needs his private key and the key-part of the paper-ticket. The paper-ticket can be replaced by an electronic stored information on a patient's electronic card. A point to mention is that the hand-over of the token is in fact the declared IT-consent of the patient.

Compared to our approach, the key on the ticket is an outsourcing of the Trusted Third Party's key service. The patient overtakes this part. A disadvantage can be that if once a

doctor has a copy of the ticket's key, he can access the net-folder any time he wants. He can as well share this information with an other doctor without the patient's IT-consent.

The D2D approach is different from ours because the intention of D2D focuses more on the transformation of medical information from “one provider” to an “unknown receiver”, P2UR.

The EHR functionality is based on the ticket system. An independent search for a patient is not possible, nor the search for a special document via meta-data. Compared to our approach, we are able to search for patients with their identifying data against the pseudonymization service of the TTP. Furthermore the professional users are able to select special medical documents (X-ray reports, laboratory reports, medication information, etc.) because of the open search-able meta data associated with the encrypted documents in the PMIP.

6.4 PIPE Project

Thomas Neubauer and Mathias Kolb are working on the PIPE project; PIPE stands for Pseudonymization of Information for Privacy in e-Health. In their paper [Neubauer2009] they refer to different approaches for pseudonymization in the eHealth context. For PIPE they mention that PIPE is “... a new architecture that provides the following contributions compared to other methodologies: (i) authorization of health care providers or relatives to access defined medical data on encryption level, (ii) secure fall-back mechanism, in case the security token is lost or worn out, (iii) data storage without the possibility of data profiling, and (iv) secondary use without establishing a link between the data and the owner.”

The medical data are stored under a pseudonym and the secret knowledge which patient really belongs to the pseudonym is stored on the smart card of the patient. Compared to our approach, a search for patients is not in the focus. The patient hands over his smart card and with this “consent declaration” the workflow starts.

6.5 Other Approaches described by T. Neubauer

As related work to the PIPE project Neubauer and Kolb present a related work to the PIPE project. Peterson's Approach [Peterson2003], the Electronic Health Card of Germany [Fraunhofer2005], Thielscher's Approach [Thielscher2005], an approach of Slamanig and Stingl [Stingl2007], and finally Pommerening's Approaches [Pommerening2004] are discussed in [Neubauer2009].

Advantages and disadvantages are mentioned. The criteria in their evaluation are user authentication, data ownership, limited access, protection against unauthorized access, notice about uses of patients data, access and copy down possibilities of data. Furthermore from a technical viewpoint the fall-back mechanisms, secondary usage, emergency access, insider abuse, and database modification possibilities are studied.

In any approach the patient has the secret knowledge to access his data, his smart card is necessary because it contains a token or the private key. For example in Germany

the Health Card of the patient is used to decrypt a symmetric key, while in a second step the data can be decrypted with that symmetric key.

General search facilities for patients or search facilities over meta-data are not mentioned for non of these approaches. For us, this indicates how dangerous it would be if the TTP answers for identity queries, where only parts of the identifying data (i.e., Carla B*) are given by the requester. The requester should be forced to specify exactly the patient-identifying data; without proving this knowledge, a requester will not get an answer from the eHealth platform.

7 Details of the eHealth Platform

This chapter extends the previous chapter 5 with some technical details on dedicated topics.

7.1 IHE-XDS Adaptation

The eHealth platform is composed of two main, operative separated parts: the Trusted Third Party Provider (TTP) and the Pseudonymized Medical Information Provider (PMIP). The TTP keeps the mapping table of patients' identifying data to their pseudonyms, while the PMIP contains the medical information, encrypted and associated to patients' pseudonyms.

The main component inside the PMIP is a so called *Central Medical Registry* (CMReg). Each document provided to the eHealth platform has to be registered in the CMReg. CMReg is the central place where a searcher of information starts his search. The storage location of the corresponding documents are *registered* in the CMReg. Possible storage locations for documents are (1) the Centralized Data Repository, (2) one of the Decentral Data Repositories of the information providers, (3) a national back-up storage system for medical data, or (4) any other location that can be made accessible to the requester.

Technically all data repositories are treated in the same way. The CMReg keeps the meta data of each information item, including the storage location of the (encrypted) data itself. An example shows the use of pseudonymization:

Data set:

- patient: Mr. Antony Bee, 131, Rue de Strasbourg, Luxembourg, SSNo. 19650304280
- document-type: Laboratory Report
- execution date: 2011-03-18-08:00
- prescriber: Dr. med. Wasp, User-ID 13483
- location of information: platform-central01://444.333.222.111.000.ABC
- storage format: encrypted

Confidentiality of data is not guaranteed against the administrator or an intruder because the identity of the patient is disclosed here together with his medical information. As mentioned in the previous chapters the information kept in an EHR is very private and confidential. The CMReg can be attacked from external or by an internal administrator. The medical records and the

corresponding person data can be stolen. If there is time enough the administrator may produce copies – even of the data in the decentral repositories. An illegal usage of the data is (partly) avoided by pseudonymization of the patients' identities.

7.2 Pseudonymization

Considering again the above Laboratory report data. The medical part is now stored under a pseudonym at the PMIP:

- pseudonym: **5413562340535634557**
- document-type: Laboratory Report
- execution date: 2011-03-18-08:00
- prescriber: Dr. med. Wasp, User-ID 13483
- location of information: platform-central01://444.333.222.111.000.ABC
- storage format: encrypted

The mapping of the pseudonym to the real identity is stored at the TTP:

- patient: Mr. Antony Bee, 131, Rue de Strasbourg, Luxembourg, SSNo. 19650304280
- pseudonym: **5413562340535634557**

Each information part (the TTP's part and the PMIP's part) is managed on separate servers, by separate organizations, and by different administrators. It has to be assured that the operational business of both organization is never outsourced to the same IT service provider. The preparation of the information for TTP and PMIP is done already on information delivery side. But the pseudonym stays a secret between the TTP and the PMIP. It will never be visible for a data delivering party nor for a data querying party.

The question is how to hide the pseudonym from outside? The solution is, when providing data to the system, the data delivering party (laboratory, hospital, etc.) creates a data package for the TTP and a data package for the PMIP. The TTP data package contains the person identifying data and a join-token; the PMIP data package contains the same generated join-token together with the medical data. Having the person identifying data, the TTP assigns a pseudonym to this person identifying data and waits with pseudonym and join-token for the following PMIP request. The PMIP asks the TTP with the join-token for the pseudonym belonging to that join-token. The TTP returns the pseudonym to the PMIP and the PMIP registers the medical data associated to that pseudonym.

When querying data from the system, it works analogously. The requesting party (a general practitioner for example) splits up the query stream into two parts. The TTP receives the person identifying fragment together with a join-token; the PMIP receives the medical fragment of the query together with the same join-token. The PMIP asks the TTP for the corresponding pseudonym(s) by presenting the join-token. The TTP replies a list of pseudonyms to the PMIP.

Remark 1: Why is it *a list of* pseudonyms in the retrieve case? – For each different institution the patient gets a new pseudonym in the TTP. The reason is that the patient's name, address, or social security number often needs to be corrected in the primary system. The only stable information in the assigned patient identifier (PID) of the provider's IT system when a patient is encountered there. If the pseudonym inside the TTP is stored in relation to the local PID of the delivering institute, an update message for the person identifying data always contains the local PID. Such an update of patient identifying data (name, address, etc.) can be communicated to the TTP without involving the PMIP. As those updates are expected very often, the pseudonym in the TTP is stored in relation with the local PID of the delivering institute. This implies that the TTP keeps multiple pseudonyms for each patient – i.e., one for each data providing IT system. In the requesting case the patient gets identified at the TTP by his person identifying data, not by a local PID. The TTP provides the list of pseudonyms for that patient. The PMIP executes the queries for the list of pseudonyms – which all focus the same patient.

Remark 2: Why not using the local PID instead of a pseudonym? – A good argument for simplification is to use directly the local PID instead of a pseudonym. As every hospital uses independent local PIDs, and if a bad administrator gets a stolen copy of the PMIP he only retrieves the information out of the eHealth platform that he has in any way in his hospital system. Unfortunately it is not only the hospital's IT administrator who knows the local PIDs. Local PIDs are used as well on every label inside the hospital, at each patients record. Even the patient's bed is sometimes labeled with his local PID. It is recommended to use the described join-token mechanism. Then the pseudonym is a secret between TTP and PMIP.

7.3 Primary Usage for Patient Care

Each patient related document (CDA, PDF, X-ray, etc.) contains person identifying data inside of the document. The running example may illustrate the situation. The pseudonymization is done, and the registry CMReg, which is a component of PMIP, keeps the information:

- pseudonym: 5413562340535634557
- document-type: Laboratory Report
- execution date: 2011-03-18-08:00
- prescriber: Dr. med. Wasp, User-ID 13483
- **location of information:** platform-central01://444.333.222.111.000.ABC
- storage format: encrypted

The named location refers to a data repository, where the document is stored. This can be a central repository or a decentral repository in the staging area of an hospital's DMZ for example. The stored document itself can be encrypted or open. For patient care usage, i.e., the primary usage of the eHealth platform, the document is encrypted in any case.

7.4 Secondary Usage for Statistics

A parallel storage of the same information is recommended, if needed and if legally covered:

- First, in *encrypted form* for patient care (see subsection above), i.e., the original document including its signature.
- Second, a copy in the *identity-stripped form* for statistical usage. It is recommended to encrypt this fragment with the public key of a dedicated statistics user.

A further precondition for statistical usage is that the data items are evaluable. For CDA level 1 i.e., free text, a statistical evaluation is a research approach of data mining techniques is not easy applicable. CDA in level 2 can be seen as text but already structured with some main headlines. CDA in level 3 hold the classical precondition for statistical evaluations. For the example use above, the pseudonymization is done in the same way and the registry CMReg keeps the information:

- pseudonym: 5413562340535634557
- document-type: Laboratory Report
- execution date: 2011-03-18-08:00
- prescriber: Dr. med. Wasp, User-ID 13483
- **location of information:** platform-central01://444.333.222.111.000.**Statistics**
- storage format: **stripped-CDA**

Statistical evaluations without involvement of the TTP are limited. If further statistical research has to be done, and personalized data like age and sex are necessary, special exceptional queries must be allowed explicitly to retrieve the list of corresponding pseudonyms from the TTP. The allowance is a kind of governmental consent to access those data.

For example we can query an average blood value of all patients in the medical database. If we need this average for all female patients with the age interval of 60 to 80 years living in the Luxembourg city, we need a governmental consent to ask the TTP first for list of the corresponding pseudonyms. With this list we can extract the medical results sets and assign new pseudonyms to the results – in case a further research on the results is also covered by the governmental consent. In anyway an allowance for research with person related data is necessary.

Statistical usage and care usage are seldom mixed-up. (An exception is described in the next subsection.) Technically seen, the PMIP registry holds a database table where one column has the pseudonyms. Then, for any care related application, associated to the pseudonyms are pointers to the fully encrypted documents. And, for any statistical application, associated to the pseudonyms are pointers to the stripped-off data fragments. The sketched database table below shows that a common usage (care and statistics) of the pseudonyms is possible in the PMIP without decreasing the existing level of data privacy.

pseudonym	only-for-care	only-for-statistics
Ght677uIoP:&d	Ref-to-encrypted-document 4711	Ref-to-stripped-fragment 9243
Jdwed669iu;otZ	Ref-to-encrypted-document 1344	Ref-to-stripped-fragment 3567

On the one side, statistical applications only uses the information “*this is the same patient*” while on the other side care related applications are in contact with the TTP and “*know the patients' identities*” behind the pseudonyms.

Remark: A sub-optimal alternative for statistical usage is to declare a special health professional role for statistics. This statistics health professional has access rights to all data of all patients. An application driven under the rights of this role can access every data, decrypt them, and prepare them for further statistical usage. That means the erasing of person identifying data is postponed. – This alternative can not be recommended, because it opens an uncontrolled access to any information in the eHealth platform for a dedicated role or user.

7.5 Personal Dashboard

A personal dashboard is the exception for not intermixing statistical data and personal health care data. Here a patient may compare his values with the average values of other patients. In some cases this can motivate a patient to continue with his efforts in losing weight or stop smoking or drinking. In how far this applications are useful or contradict a professional treatment has to be cleared with the health professional community. Technically it is possible.

7.6 The Enemy knows the System

In the year 1883 Kerckhoff states a principle about secret algorithms versus public known algorithms: “Kerckhoffs' principle (also called Kerckhoffs' assumption, axiom or law) was stated by Auguste Kerckhoffs in the 19th century: a cryptosystem should be secure even if everything about the system, except the key, is public knowledge. Kerckhoffs' principle was reformulated (perhaps independently) by Claude Shannon as "The enemy knows the system." In that form, it is called Shannon's maxim. In contrast to "security through obscurity" it is widely embraced by cryptographers.” [Kerckhoffs1883]. In that spirit the description of the proposed eHealth platform, the used algorithms, and the protection mechanisms are available for discussion. The authors of this document encourage the readers for any feedback concerning security and possible improvements of the proposed data protection.

7.7 Archiving and Data Aging of Encrypted Documents

Encrypted documents are in danger of getting disclosed in the future because (1) the used cryptographic algorithms turned out to be corrupted or (2) the power of computers increase in a way that the key length of today's encryption is not large enough, or (3) keys are corrupted due to an accident or theft. Refer back to chapter 3.1 the communication P2R (provider to

receiver) and P2UR (provider to unknown receiver) are not involved in this problem because the data can be deleted after delivering them to the receiver. The P2MUR (provider to Multiple unknown Receivers) is the candidate for data aging problems. Those data are kept over years or even over decades in the eHealth platform.

7.7.1 Example

Imagine, we are in the year 2030. 15 years ago, a document had been provided to the eHealth platform for the patient Shania Bee's. The document provider – a former clinic that had stopped its services years before – had encrypted the document *doc* with a generated symmetrical key *symkey* which results in *symkey(doc)*. The *symkey* itself had been encrypted with the public server key of the TTP which results in *pubkeyttp(symkey)*. The pair

$$\langle | \text{pubkeyttp}(\text{symkey}), \text{symkey}(\text{doc}) | \rangle$$

has been provided to the PMIP. Both key lengths are now too short and in danger of getting cracked. The computer power has increased unexpectedly fast. Additionally, the Saint-Bankruptcy-Clinic does not exist any more.

7.7.2 Data Aging of Public/Private Keys

Concerning the *pubkeyttp*, there is a simple solution. The TTP has already a re-encryption feature which is used by the PMIP for re-encrypting the symmetric document-keys for a requesting receiver. This existing feature can be used for re-newing the *symkey* encryption on TTP side:

$$\text{pubkeyttp}(\text{symkey}) \rightarrow \text{symkey} \rightarrow \text{pubkeyttp}^*(\text{symkey})$$

Instead of re-encrypting for a receiver, the TTP re-encrypts with its own new public key. At PMIP side the pair is now:

$$\langle | \text{pubkeyttp}^*(\text{symkey}), \text{symkey}(\text{doc}) | \rangle$$

Remains the problem with the short key lengths of the *symkey*, where this *symkey* had been generated 15 years ago by the former Saint-Bankruptcy-Clinic.

7.7.3 Data Aging of Symmetric Keys

The case where the symmetric key lengths get outdated is more complex to handle. As the re-encryption task for the asymmetric keys only discloses the symmetric key on TTP side, and the TTP never has access to the symmetric encrypted document, the re-encryption of the asymmetric keys is without additional danger. Not so for the symmetric keys.

Discussion

The re-encryption of the *symkey(doc)* towards a *symkey*(doc)* is dangerous, because the party

who should execute this has to disclose the document. Disclosing the document is directly a violation of the professional secret.

TTP should not re-encrypt the Document

The TTP is a bad candidate to re-encrypt the encrypted document $symkey(doc)$. The symmetric key $symkey$ and the document doc must never be together in one administrator's realm.

Origin Data Deliverer should not re-encrypt the Document

In the example, the former Saint-Bankruptcy-Clinic had delivered the document 15 years ago. The successor organization – i.e. the Saint-HealthAgain-Clinic has the same data in the internal IT systems. So the Saint-HealthAgain-Clinic may be asked for renewing the encryption. – But this is not practicable!

Symmetric Encryption Cascade is a Solution

Instead of re-encryption of the symmetric encrypted document, a cascade of symmetric encryptions is a better solution. In average expectation the out-dating of symmetric algorithms may occur every 25 years. For a 99 year lifetime this means a cascade of maximum 3 - 4 extra symmetric keys. For documents older than 25 years, one extra decryption, and for documents older than 50 years, two extra decryptions are necessary.

After the first 25 years – assuming the law allows or foresees the storage of documents for that long periods – the information item

$$\langle | pubkeyttp(symkey), symkey(doc) | \rangle$$

is getting updated towards

$$\langle | [pubkeyttp*(symkey*), pubkeyttp(symkey)], symkey*(symkey(doc)) | \rangle$$

Again with some words: For each encrypted document $symkey(doc)$ the PMIP generates a new $symkey^*$ and over-encrypts the encrypted document towards $symkey*(symkey(doc))$. In order of a proper decryption, the PMIP then asymmetrically encrypt the new $symkey^*$ with the (potentially new) public key $pubkeyttp^*$ of the TTP. The PMIP then deletes the generated new $symkey^*$. The cascade of symmetric encryptions is noted in the first argument of the notation above; this is now a list of $pubkey$ encrypted $symkeys$, respectively. Please notice the first argument above. After 50 years this information item extend towards:

$$\langle | [pubkeyttp**(symkey**), pubkeyttp*(symkey*), pubkeyttp(symkey)], symkey**(symkey*(symkey(doc))) | \rangle.$$

Re-Encryption of the Encryption Cascade for a Receiver

For completeness it should be mentioned that the re-encryption process for a legal receiver (a doctor asking for the document) handles as well the encryption cascade. In our example after 74 years the TTP is asked by PMIP for re-encryption for the receiver *rcv*. TTP will exchange:

$[pubkey_{tpp}^{**}(symkey^{**}), pubkey_{tpp}(symkey^*), pubkey_{tpp}(symkey)]$.

towards:

$[pubkey_{rcv}^{**}(symkey^{**}), pubkey_{rcv}^{**}(symkey^*), pubkey_{rcv}^{**}(symkey)]$.

Please notice, that for all *symkey*'s the newest public key of the querying receiver *pubkey_{rcv}^{**}* is used. Please notice as well that this *pubkey_{rcv}^{**}* is a temporary key, because we do not support an extra cryptography PKI. The cascading list of *symkey^{**}*, *symkey^{*}* and *symkey* is used by the receiver to decrypt the document's encryption cascade:

$symkey^{**}(symkey^*(symkey(doc)))$.

The cascaded decryption is fully automated. Performance is not critical. Remember that the over-encryption was necessary because the computers have evolved that far, so that a brute-force attack to the old symmetrical keys has been expected and their length was seen as too short.

Should PMIP perform the Encryption Cascade?

After generating the new *symkey^{*}*, PMIP use it to over-encrypt the document. Afterward, PMIP has to destroy the *symkey^{*}*. If not, and an evil PMIP administrator keeps a copy, then after the next 10 years when the old encryption is assumed to be in danger of getting cracked, the evil administrator can first apply the *symkey^{*}* on *symkey^{*}(symkey(doc))* to get *symkey(doc)*. With his “new” computer and a brute-force attack he can disclose the document.

Even if the over-encryption is done by a Fourth Trusted Party, the risk remains. The evil administrator may keep copies of *symkey(doc)* that are vulnerable with his “new” computer and a brute-force attack. – After 15 to 25 years.

Remaining Risk, Destroy Information or Renew Encryption

This remaining risk is a reason to destroy the documents in the eHealth platform – per law – after approximately 15 to 20 years. Then there is no remaining risk due to data aging. If a disease needs a lifelong recording, a trusted medical specialist may download the whole record and provide it afterwards anew with a new encryption. But before elaborating a technical or even the sketched workflow solution, the political will for lifelong preservation of data should be validated.

Backups can be Secure

For Backups the situation is similar, but here a “strong” solution could be practicable. The backups are encrypted with keys of the same size as the file itself, that should be encrypted. Really(!), both, the key and the encrypted file have the same size. The XOR⁷ of key and encrypted file results in the original information. If the key is truly random and the key is separately stored – for example key at the TTP and encrypted file at the PMIP – backups are secure as long as not both administrators are evil and work together.

7.8 Hiding Patient's Identity against the Web-Server

An requester (doctor, patient) can access the eHealth platform via a web-server. An illegal acting administrator of the eHealth platform's web-server may install a logging, catching the requests containing patient names and catching the (encrypted) results for those patients that pass the web-server as well in the other direction.

A first prevention against illegal logging is to hide the person identifying data from the web-server. Therefore the web-client sends the patient identifying data over the web-server to the TTP. As only the TTP needs to know this information, it is encrypted with the TTP's public key. The TTP has one additional step to do: it has to decrypt the person identifying data. Then it continues by looking-up and providing the pseudonyms and then waiting for the PMIP request for the pseudonym-list.

Remark: On the technical implementation side, patient identifying data have to be displayed in the client's application but not rendered in the same web-session on the web-server together with the medical results. A technical solution may be the separation on two web-servers as well as an applet, which hides the person identifying data against the web-session.

7.9 Certificates, User-Roles, and Role-Assigned Applications

The previous sections sketched the protection of information against unauthorized administrator or intruder access. End-to-end encryption on application level, and the special case of intermediate storage for unknown receivers has been described. Pseudonymization in combination with encryption is used for medical information inside. Besides these levels of information protection, a user access to the eHealth platform is secured with a personal ID-card or an other certificate (e.g., a Luxtrust card).

The login is done with a personal certificate, where the certificate delivering organization guarantees for the correct match of the certificate owner and the corresponding human individual. For example: The stored data of Dr Wasp's on the card do really match with the identity of the human individual Dr Wasp.

Nevertheless, the certificate delivering organization can not guarantee that the certificate is used only by that

human individual. The human individual has pencil-signed an agreement to keep the password secret and not to give his certificate to an other person. Nothing more can be guaranteed.

7 The boolean XOR operation results in a boolean 1 if both input values differ, else the result is a boolean 0.

A user&role directory guards the access to the system. The identity data of the certificate has to be registered within the eHealth platform together with the user-roles. The role of a user can be doctor, dentist, physical therapist, pharmacist, patient, etc., where multiple roles are possible for one user. Further institutional certificates and their corresponding roles can be defined. Finally a set of (web-)services is assigned to each role.

The user logs in with his certificate; then the system checks the user's role(s). According to the assigned role(s), the corresponding (web-)services are offered. The offer of (web-)services can be realized as web-application or as build-in feature in a local software, that uses the offered service(s).

Depending on the use case specifications, personal certificates may be used for signing documents and institutional certificates may be used to secure the data transport. Concerning login workflows, the later policy (use case for login) may accept only personal certificates or – in a more open way – it accepts institutional certificates where a strict logging of the users of the connected institutes' software is guaranteed.

7.10 PKI for Authentication, for Signature and for Cryptography

The user login and authentication is done with a certificate. The offers of the certificate provider (e.g., Luxtrust) are often restricted to authentication and signature. Signature keys should never be used for cryptography! The reason is obvious: The key pairs (private/public) could in fact be used for both, signature and cryptography. But the infrastructure behind cryptography and behind signature is totally different. For signature keys a backup infrastructure is forbidden because it contradicts the non-repudiation of signatures. But why is a backup of a signature private key not necessary? – If the signature certificate (e.g., a card) is lost, a new signature certificate will be valid for the next time interval. The formerly (old) signed documents still have a valid signature. Compared to real life: a person ink-signs with the right hand. Unfortunately the person has a fracture of the right arm. In real life he ink-signs with the left hand, starting from now on. Back to electronic business, the user loose his signature card. He gets a new one which is valid, starting from now on. The old one is revoked. Clear, signature infrastructure works without backups.

Cryptography on the other side needs a backup strategy for the private keys. If a key is lost, and a backup is missing, all encrypted documents will be lost. A PKI build for signature should never be used for cryptography because of the missing backup. Vice versa, a key infrastructure build for cryptography should never be used for signature because of the loss of the non-repudiation. The user could argue that he has not signed a document, but someone else may have – with the stolen backup key. This is much worsen than saying someone else has decrypted an information with a stolen backup key.

In section 5.2 is already explained how temporary public/private key-pairs can be generated on the basis of an existing authentication and signature PKI.

7.11 Alert Functions

Alert functions and a yearly report for curious accesses set up a mental barrier for everybody who may misuse the system for unauthorized data retrieval. The interested doctor in the neighborhood, the relative who is working in a hospital, etc. Depending on the individual configuration for each patient, automatic alert functions are switched on. They inform the patient and (for example) his family doctor on any curious data access to the patient's record. In especial, an emergency access is reported. This information is important also in case of a legal access. Then the family doctor and maybe a friend of the patient is informed that he or she is in hospital in an emergency case – or a data misuse is detected! This secondary usage of the alert enhances the misuse protection.

The preferred way of information about alerts can be configured by the patient. This may be email, SMS, or a paper letter at the end of the year. Default configuration should be a yearly report.

Additionally to the subscription of alert information, the patient should be allowed to actively watch the current logging history of his records via a web-application. Subscription and active watching of the logging activities have to be enabled without disclosure of the patients identity to an evil administrator.

Non-reputation of audit information can be assured if besides the data delivery operations as well every data request has to be signed by the requester.

8 Embedding External Services

A formal centralization of information, i.e., a single point of access for important health related information automatically implies the adaptation or routing through options of other existing central services of the same domain. Two first examples are explained in more details below. Affiliation information and an improvement of the demographic data for in-house systems. Additionally, a generic service could export the *user authentication services* for other external web-services.

8.1 Affiliation Checks and Insurance Information

The embedded external service delivers the information whether a patient is affiliated with an insurance or not. Health care delivery organizations are often interested in this financial information. The affiliation information is routed through by the eHealth platform to the health professional. Therefore the eHealth platform uses a web-service which will be provided by the insurance organization (CNS). The forward through approach has the advantage that the accessed information is always up-to-date.

Technical there are two options which part of the eHealth platform should use this web-service. (1) The Trusted Third Party TTP that deals with patient identifying data, uses the web-service that reflects the affiliation information, or (2) the affiliation checking web-service is called by the Pseudonymized Medical Information Provider PMIP. As short description of both options shows how to handle. In any case the CNS will offer a web-service.

TTP version

For the TTP version the CNS web-service expects as input the matricule and/or some other patient identifying data. The web-service will answer with the information that this patient is affiliated or not. TTP calls the web-service of the CNS and gets as answer the affiliation information. Together with the requested pseudonym-list the TTP delivers this up-to-date affiliation information to the PMIP. PMIP build in the affiliation information into the current web-application. If the web-application is only an affiliation check, then the web-application in PMIP simply replies this information to the user.

PMIP Version

Here the PMIP will use the web-service of CNS. But it is slightly more difficult. PMIP does not know the patient identifying data. So, PMIP will generate a session token and hands over this to the CNS web-service as well as to the TTP. The CNS web-service will then ask the TTP with the session token for an associated matricule. Having the matricule the CNS web-service can retrieve the affiliation information and respond this in a final step to the PMIP – as answer to a given session token!

It is highly recommended that the session token is an other than the one used by the health professional for the initial query. As well it is highly recommended that the pseudonyms are never disclosed towards CNS. Remind that a patient has several pseudonyms. The ping-pong with session token, matricule, affiliation information has to be repeated for each pseudonym of the patient.

In the authors opinion the PMIP version has a higher risk factor for data protection, because the TTP responder service which is needed by the CNS web-service can be attacked. The TTP version is straight forward to implement.

8.2 Generic Services for User Authentication / Access Verification

A simple check of the patients affiliation can as well make use a more generic service. The proposed eHealth platform has an internal service for user authentication. The platform manages a directory service for pre-registered users, i.e., health professionals and patients with their corresponding roles. During the authentication process the certificates are used to identify the user. If other application provider like to offer health related services, they may be interested in using the user authentication service of the eHealth platform. Technically the authentication process returns SAML tokens (Security Assertion Markup Language) that are used internally for accessing the TTP and PMIP, and which can be exported for other external web-services.

If the CNS offers a web-service for affiliation checks and this web-service would accept SAML tokens generated by the eHealth platform, the affiliation check service can be outsourced back to the CNS as it is.

8.3 Improvement of Demographic Data Quality

Care delivering organization communicate with the TTP about patient identifying data. They can mistype addresses or names, names may change because of child adoptions or marriage, a patient can move, etc. As result, organizations have clearing departments to find out that a

patient is double or multiple inserted in their systems. The common function in those systems is a so called patient-data-merge. Two patient data sets are discovered as belonging to one person. They are merged in the way that at the end one of the duplicates is erased while the other one overtakes the information formerly stored at the duplicate. This kind of data clearing should be done at the places where the information is collected and where the patient is physically available.

With a central EHR the number of such duplicates will increase because more than one delivering organization is involved, and patient with complex names are often wrong spelled – in more than one institution.

A future clearing support can be established if the TTP is allowed to consolidate a national person register. The TTP can check with new and old addresses as well as name misspellings and name changes. In the case of a possible duplicate the TTP can inform the data providing organization that two person's data sets *may refer* to the same person. With this hint the clearing department of the organization can initialize a validation of this hint, and in the positive case the clearing department performs a local patient-data-merge and communicates this merge to the TTP via a special patient-data-merge message. Such a message is typical between inhouse systems. A standardized HL7 message is available. This is the starting point for improvement of the demographic data quality inhouse for every participating organization as well as for the TTP database itself.

The legal question is, whether the TTP is allowed to access a national person register and to inform the data delivery organizations (hospitals, laboratories, etc.) about potential duplicates.

9 Data Protection Overview

Data security is a primary aim in this concept of an eHealth platform. This chapter first summarize the levels of different data protection barriers and then it names the remaining weakness and security leaks.

9.1 Levels of Data Protection

The different levels are:

Pseudonymization

The separation of person-identifying information at the TTP and the medical information at the PMIP protects the stored information against a first attack with a patient's name.

Key Re-Encryption

The Key Re-Encryption protects the information during the preparation, i.e., the re-encryption for a requesting receiver. A secure re-encryption is possible because the documents are encrypted with symmetric keys, while the symmetric keys themselves are encrypted asymmetric. The processing node for re-encryption (TTP) is never aware of the encrypted document. Instead it re-encrypts symmetric keys.

IT-consent

IT-consent declarations protect each patient's medical information according to his private wishes, i.e., excluding his neighbor who is working as nurse in an hospital.

Logging and Alert Functions

Logging and alert functions are a psychological barrier that can avoid unauthorized access. Informed patients will ask for legal consequences in case of recorded data protection violations.

Security Token Service

The Security Token Service of the User Management protects the whole eHealth platform from non authorized access.

Timestamp Service

A timestamp service protects against manipulations of information backwards in the past.

Public Key Infrastructure for Signatures

A Public Key Infrastructure (PKI) for signatures guarantees that information is provided by registered professionals, and that requests are coming from registered professionals.

Certificate Based Authentication

Certificate based authentication for login protects the access to the eHealth platform's services with the help of an existing Certification Authority. The token carrying the certificate may be a card, a USB stick or any other appropriate device.

9.2 Security Limitations

Nevertheless, there are remaining weak points in data protection. They are minimized from a technical viewpoint as well as from the organizational viewpoint (TTP and PMIP separation). The weak points are:

- A user of the eHealth platform may misuse the given access rights. The access certificate and the passwords/PINs can be stolen or handed over to a other person. The implemented logging and alert mechanism is a psychological barrier, but – first – the access has occurred then. Prevention can be established by a high sanctioning law. A necessary precondition for sanctioning measures is that it is doubtlessly clear who has accessed illegally. Therefore accesses have to be proofed definitively by the logging. Each requester must sign his request before the eHealth platform will answer, and the signed request appears in the audit protocols.
- An artificially identity can be created. The barrier is very high, but not impossible. A governmental organization can create an artificially identity for a fictive person. This person applies for a Health Professional Card (e.g., a Luxtrust certificate that is as

well registered at the eHealth platform) to gain access to the eHealth platform. The fictive created health professional need to be registered as well for the user management and the Security Token Service. With this fictive health professional identity, access to the health records of patients is possible. It can be avoided by establishing a white-list behavior for the IT-consent management. The fictive health professional is not in the allowed white-list. When using white lists, even emergency accesses have to be declared there.

- In case of a two-organizational TTP / PMIP separation, two administrators working illegally together, or both databases can be stolen by intruders. This risk can be minimized further by multilevel pseudonymization or regular pseudonym exchange.
- Master key weakness. For each document there is a new generated symmetric key with which the medical information item is encrypted. The symmetric key itself is encrypted with the public key of the Trusted Third Party. If the private key of the TTP gets disclosed and the database of the PMIP gets stolen all documents are disclosed. To avoid this, the master key gets renewed periodically – daily, hourly, or in the extreme case, the TTP generates a new asymmetric key pair each time a document provider asks the TTP for the TTP's public key. It should be noted that the access to the eHealth platform is still protected, even when the master key gets disclosed. And a re-keying with new master keys is provided with the re-keying for legal receivers. An equivalent situation is already solved in section 7.7.2: Data Aging of Public/Private Keys. An other approach is the usage of a hardware security module (HSM) for the private master key. The workflow has to be described including a backup: HSM solutions destroy their keys in case of an attack. Without a backup all information is lost in that case.⁸

Extensions of the services of the eHealth platform bring the danger of security violations. Each extension has to be designed very carefully. Some examples are given in the later chapters.

10 Consent Management

Consent is the patient's permission or denial to an “act”. The consent may cover:

- (a) explanation, information and agreement or denial to medical treatments (MT) or
- (b) the information technological (IT) storage and usage of the treatment related health data.

Both are different topics. In this document the focus is on the IT-consent. For distinction, the medical treatment consent is explained very short. Then the focus returns to the IT-consent.

Medical Treatment (MT) Consent

Consent on medical treatment (MT) should be asked during a discussion between the patient with the healthcare professional, who is directly responsible for patient's treatment. This could be a nurse who is arranging a blood test, a general practitioner who is prescribing new

⁸ Some people in Germany how to do so.

medication, or a surgeon who is planning operation. The notation of the MT-consent, i.e., the patient's will after an informative discussion, can be documented in an IT system. MT-consent is a further application for the eHealth platform.

The following focuses on the consent concerning the information technological (IT) storage and usage of the related health data. In contrast to MT-consent, IT-consent is a basic feature of the eHealth platform – belonging to the data access policy. IT-consent is denoted in the eHealth platform itself as a guardian for all stored information.

Information Technology (IT) Consent

IT-consent is the will declaration of a person, expressing who can create, access, modify, delete that person's data. The IT application for the declaration of this IT-consent is called *IT-consent management*. This should not be mixed up with MT-consent management. Both consents can be declared on paper or in electronic form.

Patient data should not be stored or processed on the eHealth platform without the patient's will, so without his IT-consent. Today, patient's data are stored within primary systems of hospital, laboratories, doctors' systems, pharmacists' systems, etc. The patient implicitly agrees, or he has declared his will on paper and ink-signed it. This allows the storage of patient-related information in the health professional's IT system. Processing and storing medical data on a national eHealth platform is a large step beyond – and this is the topic here.

In the following sections the term “consent” refers to “IT-consent” if nothing else is mentioned.

10.1 Media for Patient's IT-Consent Declaration

The media to declare the consent are: (1.) paper form, (2.) computerized form filled out in doctors cabinet, (3.) computerized form filled out at the patients own computer via a web-application.

A) **Paper based:** The simplest way of a patient consent declaration is a paper based form. The patient⁹ is asked to fill out and sign his declaration with a pencil. The paper form is scanned in, stored as “electronic” patient consent, and the doctor or his assistant fills in the information of the paper form into the IT system. The doctor assures that he fills out the system data as the patient told him with the paper form. As proof the scanned paper-document is attached.

B) **Electronic with Print-Sign-Scan:** The next step towards “electronic” is when the former paper based form is provided electronically. In this case the patient is asked to fill out directly the electronic form using an IT structure (tablet-PC). In absence of a valid electronic signature, the patient signature has to be done on a printout of the form. Again, a scan of the

9 For elderly people, handicapped people, children, etc. an other person can assist the patient for expressing his consent.

final ink-signed document has to be attached to the electronic form. The assistant does not need to type in the information, because the patient did it already.

C) Electronic with valid Electronic Signature: The ink-signature is replaced by an electronic signature and the print-sign-scan sequence can be avoided. This is the more flexible approach for declaring the IT-consent. It enables as well a web-based IT-consent declaration from the patients point of view.

This classification outlines in parallel the evolution steps towards a full electronic IT-consent management. The next sections concerns on the default settings for consent and the types of consents that may be declared by the patient.

10.2 Opt-In or Opt-Out

Depending on the political decision whether all people, i.e. all possible patients, are per default within the eHealth platform or not, the first level of the patient consent declaration is to agree or to disagree to the default. If all people are per default inside, this is called opt-out; it means that everybody has the option to step out if he do not like the default of being in. The other choice of a default is opt-in. This means the patient actively has to choose the option of being “in”.

The eHealth platform acts independent of this political decision. Technically the yes-or-no consent can be realized on the level of the TTP's pseudonymization service: A data provider sends patient identifying data (together with the local patient ID) to the TTP, and in parallel the encrypted medical information to the PMIP. Only if the TTP knows the patient, it provides a pseudonym to the PMIP for storing the information. If the PMIP does not receive a pseudonym, the medical data can not be stored in the eHealth platform.¹⁰

The remaining question is how a patient is inserted in the TTP. Therefore the TTP offers an interface to insert patient-data. This interface is used by the TTP providing organization. – If opt-in is preferred, every person can apply for participation. If opt-out is preferred, for everybody who does not disagree explicitly, his data are inserted at the TTP. Combined with the political decision opt-in or opt-out, the handling procedure has to be defined. For example the applications (for opt-in) or disagreements (for opt-out) can be collected over the public health insurance. The current Luxembourgish law tends to opt-out. The European directive 95/46/EC [EU-Dir95_46_EC] recommends only explicit consent in its Article 8.2.(a) followed by the exceptions for health relevant data with Article 8.3. More details on this can be found in [Hohmann2010]. Nevertheless, from the technically viewpoint of this document, both, opt-in and opt-out, is configurable. A more elaborated consent declaration will be possible with rule based consent declarations. As well the consent declaration on base of each information item will be possible in the future. The handling, i.e., how a patient can declare his IT-consent in a general understandable way is topic of current research.

¹⁰ Instead of storage in the eHealth platform, the result can be transported in a P2R communication into the folder of the known receiver – i.e. a prescriber of a laboratory report.

11 Compilation and Composition of the Health Records

Compilation is the selection of relevant information inside of source systems by the owner of these information. Composition is the structure giving work to compose the different result-subsets of multiple compilations.

11.1 Compilation of Relevant Subsets in the Source Systems

Compilation is the summary-creating work of a domain specialist. The specialist uses his local documentation system, a primary system like HIS (hospital information system), RIS (radiology information system), LIS (laboratory information system), a medical specialist's documentation system, etc., and he decides which subset of these medical information is important enough to become an input for the national EHR of a patient. The informative content of the selected subset may be compared with that of a doctor's letter. The difference is that – now – the information is addressed to a broader group of (potentially unknown) receivers instead of addressing it to one or more known colleagues or hand it over to the patient.

Concerning secondary usage, i.e., statistics, the data provider may as well prepare special data sets for other purposes. The original information can be prepared for statistical usage (fragments), before it gets encrypted and provided to the eHealth platform.

11.2 Compiled Subsets to compose the Health Record

Composition is the work of assembling data to build-up the national EHR for one patient, with all pieces of information that were *compiled* before by the different specialists. Today this work is often done by family doctors, who usually receives the results of different hospitals, laboratories, radiologists or other specialists, and accumulates them in their own information *tool*. This information tool can be computer-based, paper-based, or a mix of both.

11.3 Automation and Fine Tuning of Health Record Composition

In section 3.3 a pragmatic structure of the EHR is already mentioned. All documents of a patient's EHR firstly appear in a chronological sorted log-book. The reader may notice that some explanations of the previous chapters are repeated here for convenience. At that point the selective compilation has been done by each specialist in the primary systems, respectively. The composition for the national EHR is outstanding. The chronological log-book is, as already mentioned, a first rudimentary composition strategy.

Now, the designated moderator of the EHR decides whether a chronological inserted log-book entry should be referenced in one or more of the *structuring* sections, which may be radiology, laboratory, medication, discharge-letters, etc.

Some of the moderator's actions can be done automatically by a rule-based record structuring tool. It is also thinkable that well informed patients organize subsets of their own EHR. As further alternative several doctors may act as co-moderators, overlap within their composition-work on the patient's EHR and finally use the EHR as (online) information interchange medium.

12 Storage Locations for Patient's Health Record

The question for the best storage location for the composed EHR is not technical. From technical side, several options can be offered. The existence of alternatives influences the overall acceptance of the system. The selection of an appropriate storage location is a individual decision of the patient; it is established by subjective impressions, mainly concerning the degree of data protection. Four possibilities are described below: (1) paper-based health records, (2) pocket-based electronic health records, (3) pocket-based structured electronic health records, and (4) platform-based structured electronic health records.

12.1 Paper-Based Health Records

Today, a lot of people have health records with them. These are compositions of different information of different medical domains. For example, a 85 year old women has a paper folder containing the last laboratory reports, the last discharge letters of hospitals, the last reports of the oncologist and the cardiologist; a list with the current medication and allergies. Maybe this is completed with a medical status overview in form of a (unaddressed) referral letter written by her family doctor. She is prepared for the case of an emergency happens and she is sent to the hospital, in the worst case on a Saturday evening. Her paper based health record is accessible for healthcare professionals in the hospital – if she has the folder with her. The author of the composition (i.e., the paper health record) is the family doctor. The lady is the formal owner of a copy.

Some **advantages** are:

- The record is a copy of the up to date status of the patients health situation.
- It is authored by the family doctor, who has the best overview.
- It is usable without technical barriers.
- For accessibility, copies can be stored at different locations: car, house, relatives, etc.

Some **disadvantages** are:

- Accessibility not guaranteed. (For example: in case of an accident)
- Loss of document is possible, which implies a problem of disclosure and data protection.
- Paper print-outs of illustrative pictures are in bad quality.
- Data mining in form of general decision support is not possible.
- On demand overview reports can only be generated at the family doctor's side.
- General statistics are not possible.

Concerning **data protection**:

- Safe storage locations of the paper and the copies protect. But this contradicts the availability advantage.
- A systematical illegal computer search is not possible: The “database” is one paper.
- Disclosure problems in case of document lost can be prevented by black-painting

parts of the identifying data. But this implies a loss of credibility of the information.

The consequent next step is the storage of the composed health record on a pocket based electronic device.

12.2 Pocket-Based Electronic Health Records

Some disadvantages of paper based records are eliminated when moving to electronic storage devices. For example, a USB storage or a DVD can be used. Even a patient's health card can store small amounts of data. A combination of the health card with links to a DVD is also thinkable, but from a technical view point, it is error prone and too unhandy.

Neutral aspects (not better and not worser) compared to paper based are:

- Copies can still be stored on different locations for better accessibility.
- Accessibility / availability is still not guaranteed.
- The loss of the medium implies a problem of disclosure and data protection.
- Data mining in form of general decision support is still not possible.
- On demand overview reports are still not possible, at least until the data structures are not unified and not standardized.
- General statistics are still not possible.
- Safe storage locations of the storage devices and of copies protect. Again this contradicts the availability.
- A systematical illegal computer search is still not possible.

Concerning **additional advantages** to paper based:

- Picture material can be provided in good quality.

As **new disadvantages** compared to paper based:

- A new technical barrier occurs: a computer with open USB slot or a DVD drive is necessary to visualize the electronic content. USB ports are often blocked on workstations in a network and DVD drives are slow and unhandy, if available at all.
- The format of the files need to be compatible with the software available in the care point. At least a PDF format could be preferred.

Concerning **data protection**:

- Disclosure problems in case of lost can be prevented by modifying the parts of the identifying data – the analog procedure to black-painting on paper documents. Unfortunately for electronic documents this destroys any credibility of the information.

12.3 Pocket-Based Structured Electronic Health Records

Standardized data structures of the electronic health record enhances the usability of pocket based electronic media. Standardizing brings some advantages compared to the first pocket based version. The relevant data or the patient-identifying data are encrypted. The key to access the data is stored as well on the USB device. But the key is password protected.

Then additional **advantages** are,

- The loss of the medium does not disclose the information directly.
- Data mining in form of general decision support is possible, because the software works on standardized information on the USB-sticks or the DVDs.
- Overview reports may be generated on demand, if the syntactical structure of the reports on the storage devices (USB-stick, DVD) are standardized.

New **disadvantages** correspond with the advantages of data security:

- an emergency access is not possible, if the patient can not type his password.

The next evolution step is storing standardized electronic health records on an eHealth platform.

12.4 Platform-Based Structured Electronic Health Record

With the step towards an eHealth platform as storage location of the health records, availability, accessibility, and decision support features are enabled. For security, the access to the eHealth platform is protected by certificate based authorization (e.g., Luxtrust), the eHealth platform has a user management system as well as the platform's IT-consent management which implements patient's will with respect to defined access rules.

As additional **advantages**,

- Backups are done reliable by the system administrators.
- Disclosure problems in case of lost of a device are no longer relevant.
- Accessibility in form of 24/7 availability is guaranteed.
- Picture material can be provided in best quality. Restrictions are the bandwidths.
- Data Mining in form of general decision support is enabled.
- The generation of overview reports based on a set of single reports is supported.
- General statistics are possible on non encrypted data or meta data.

New **disadvantages**,

- Unfortunately, now systematical illegal computer search is possible: Here the data base contains a large amount, potentially the EHRs of all patients of the country.

Concerning **data security**,

- Protection against risks and solutions for these remaining disadvantage are described in the previous chapters. To remind: Santec proposes a separation of medical and

patient identifying data, pseudonymization in combination with encryption, an elaborated user management, an elaborated consent management, and a certificate based access policy.

12.5 Patient's Choice ?

Acceptation of electronic health records should not be forced by laws. A better way may be to convince by showing and explaining the advantages regarding the patient's personal health benefits. Technically the core part of the investment is a secure eHealth platform like it is proposed in this document. The author of this document proposes that for the start, each patient can select between

- a paper based health record
- a structured pocket-based electronic health record on a USB device
- a health record in the eHealth platform

Important to note that a decision for paper-based or pocket-based USB is a good preparation to move later to the platform option. In any case the original work of compilation and composition remains inside the primary system, where it has been prepared by a specialist. It is preferable that the patients decide immediately for the storage on the platform. This allows the usage of the health records for decision support systems as well as for statistical purposes. The patient consent is mandatory for each of these use cases.

12.6 Technically Enabled Options

The author of this document suggests, that a technical solution should provide all three options: (1) paper, (2) USB storage devices, and (3) eHealth platform. Once a patient has decided for the paper print version, he will upgrade to USB and later on to the eHealth platform solution. An upgrade means that the family doctor has printed the paper folder and after a while the patients want to have the same information as an electronic copy on a USB stick. Upgrade from paper to USB. Remind: The original data are on the system of his doctor. As better option: the patient starts immediately with the USB or even better with the platform solution. The advantages of the platform solution are obvious, but as first step a trust in the security has to be established. The proposed solution should be published and discussed on detailed level. Data security specialists should be asked to give their valuable feedback. At the end the eHealth platform should be secure and stable from the technical viewpoint and implemented by trustful IT industries.

In parallel the political viewpoint has to be discussed. Even if the risk of unauthorized accesses is very low, political clarifications are necessary to increase the acceptance factor and trust, so that data are not used and data will never be used for other purposes, even not by governmental organizations.

PART THREE

13 From Concept to Implementation

The previous chapters define the architectural concept of the eHealth platform and the functional needs as an outlook. This chapter deals with the realization, and in special with the realization of a first version of a test-bed implementation. The goal is still to be as close as possible with the established IT standards in the healthcare environment. A deeper look on existing – so called – IHE profiles will help to save time, reuse existing software parts and prevent from reinventing the wheel.

The methodology is: we first describe what we need and what we want, and then we look which subparts are already available as ready-to-use standard components or standard concepts. IHE serves with a solid basis for this work.

A last subsection describes how the gap between existing primary systems and the eHealth platform components can be bridged first.

13.1 IHE Profiles that are useful to realize this Concept

Here we look for IHE profiles that may be useful to fulfill our needs. We will start with the eHealth platform's kernel profile, the XDS, XDR, NAV, and XCA. XDS is in line with the proposed eHealth platform architecture, XDR is as well in the context of reliable document exchange, NAV and XCA are supporting profiles that seems to be very helpful. Further security and authentication profiles ATNA and XUA and profiles for patient identification PIX, patient administration management PAM, as well as the patient consent profile BPPC will follow. Descriptions of profiles concerning the structure of medical documents will end this part.

A discussion about usability, respectively about some minor necessary extensions of those existing IHE profiles concludes this chapter.

13.2 IHE Platform XDS, XDR, NAV, XCA, XDS-I

The XDS IHE integration profile introduces actors and transactions for the exchange of electronic health records between healthcare enterprises. Instead of a direct exchange between two communication peers, the distribution of electronic health records using XDS foresees first the registration and storage of the document and as an other transaction the retrieval of the document from the Document Consumers side. Currently there are two XDS profile definitions available, the XDS.a profile which is deprecated and the newer XDS.b profile.

The XDS.b profile demands the use of Web Services for communication. The content and structure is based on the ebXML (Enterprise-Business XML) standards. Based on ebXML, which defines the messages and the architecture for the registration and storage of documents, the XDS profile definition is open to register and store different kinds of

documents. There is no content verification done by the XDS Repository.

The XDS specification is divided into two categories, the XDS content and the XDS infrastructure category. As mentioned before, the infrastructure specification for XDS.b is based on the use of the Web-Services and ebXML Registry/Repository Version 3.0 standards, which are consistent with the current developments and widely accepted as best practices in the industry. It supports the Message Transfer Optimization (MTOM) based provide, register and retrieve transactions. The content category is separately published and defines document use cases and document content. For individual transactions, document content can be based on other standards e.g., HL7 CDA (Clinical Document Architecture), DICOM (Digital Imaging and Communications in Medicine), or ASTM (American Society for Testing Materials) CCR (Continuity of Care Record).

Regarding the intended structure of the eHealth platform, the usage of IHE-XDS profile, supports the definition of our centralized document Registry with decentralized and/or centralized repositories. XDS defines the Actors: Document Consumer, Document Provider, Document Registry, Document Repository, Patient Identity Source.

Using the IHE-XDS terminology of an Affinity Domain, we can imagine the national eHealth platform as one Affinity Domain.

Communication and exchange of documents in IHE-XDS are usually described without the definition of a receiver. The Document Provider provides and registers a document via the “Provide-and-Register” transactions, which moves the document to the Repository, while the Repository is responsible for the further registration of the metadata of the document at the Registry. Some metadata are the document creator, document provider, date of creation, and the Repository location of the document itself. At any time later a Document Consumer who is allowed to access a certain document can query the Registry to search for a particular document and retrieve the document from the corresponding Repository.

In some circumstances it is very useful that the Document Provider can send documents directly to known Document Consumers. Therefore he must know the communication parameters of the intended Document Consumers. The IHE-XDR (Cross-Enterprise Document Reliable Interchange) profile is such a profile where a Document Provider directly choose possible Document Consumers and send the document to them. Usually the XDR profile is used in environments where no XDS Infrastructure with Registry and Repository exists. Then the Document Provider in XDR may use protocols based on HTTP (SOAP) or SMTP for document transmission. The XDR Profile could be used together with XDS, to provide the ability to share documents between enterprises and to be able to send documents to one or more defined Document Receivers directly. To accomplish this, the Document Source Actor must be a grouped actor and support both profiles. This could be established by implementing the extensions for example in the content profile of the “Provide and Register Document Set-b” transaction, which can contain for XDR a list of intended recipients.

Instead of directly address to a list of intended recipients, a notification process can be used to inform a possible Document Receiver in a XDS environment (e.g. the patient) of the arrival of new documents or other events. The IHE-NAV (Notification of Document Availability) profile defines transactions for notifications between communication partners

inside an Affinity-Domain or between two Affinity-Domains. The profile defines the two actors Notification Sender and Notification Receiver which should use the Simple Mail Transfer Protocol (SMTP). A common use case for NAV is the response to an asynchronous request for records, a notification for new documents, a notification for document changes or a notification for incoming referrals.

Important to mention that the profile defines, that private or person identifying patient data must not be part of the notification message. Instead pseudonyms/tokens (not our TTP pseudonyms) should be used. By using a pseudonym, the data consumer is able to get access to the document mentioned in the notification. The access will be granted if current security checks are fulfilled (authentication, authorization). After the access has been granted, the document consumer uses the informations related to the document, from the notification, and can retrieve the document.

Another interesting profile regarding notifications in XDS is the profile which introduces the publisher/subscriber model for notification exchange. The IHE-DSUB (Document Metadata Subscription) profile introduces this messaging pattern. The profile describes four actors which are involved in the transactions, the Document Metadata Subscriber, which subscribes for the interest of notifications at the Document Metadata Notification Broker actor. The Document Metadata Publisher actors should be able to emit events, for which subscriptions may exist, to the Document Metadata Notification Broker using the Document Metadata Publish transaction. The Notification Broker is then able to notify the registered recipients (Document Metadata Notification Recipient actors).

A special enhancement of the XDS profile for the radiology domain is defined by the IHE XDS-I profile. It extends the XDS profile mentioned before with transactions for supporting the exchange and access of imaging documents. Similar to the XDS and XDS.b specification, there exists a XDS-I and XDS-I.b specification, which is the newer specification based on the usage of Web services as service endpoints. Although it is currently in IHE Trial-Implementation state, which means that it is intended for final release, it is the more interesting profile for sharing radiology images and reports in context of the eSanté platform, as it is expected to reach final state soon and differs only in two transactions and the transport protocol in comparison to the XDS-I profile.

The Document Registry, Document Repository and Document Consumer actors are similar to those of XDS.b. This means that the usage of XDS-I.b is based on the existence of an XDS.b environment. XDS-I introduces two new actors, the Imaging Document Source and the Imaging Document Consumer, which is grouped with the Document Consumer actor of XDS.b. This means that a Data consumer actor can be a grouped actor for the Document Consumer and Imaging Document Consumer actor meaning that the Data Consumer actor could be capable of the transactions of both actors.

The so called significant radiology images have large file sizes with all implications for storage and transfer. Those pictures are already stored in the PACS of the local radiology department. Instead of storing copies of the significant images inside the document repositories of the eHealth platform, XDS-I offers the option to store a file with references to images, while the images themselves are in the local PACS. The files containing the

references are called DICOM manifest files. In the eHealth platform the manifest files are treated like other documents, i.e., they are encrypted stored in the repository. In contrast to the significant images, the illustrative key images can be part of the (normal) medical reports. The assignment of the medical report and the corresponding manifest file is stored in the registry.

Concerning IHE-XCA: When we take a deeper look on the possible realization, we are dealing with local healthcare enterprises e.g. hospitals, laboratories and other institutions which need to participate and being connected to the eHealth platform. In an optimal scenario, all participating enterprises communicating with our Affinity Domain are fully compatible participants and are able to implement the IHE actors which are needed for cross enterprise document sharing. In reality we have to face many problems and incompatibilities, unfortunately avoiding an easy integration for communication.

At least in the early stages of the eHealth platform, the communication systems must be connected using a certain kind of gateway or connector, which ensures the compatibility of the connecting facilities with the eHealth platform. This connector could be, as defined in the IHE-XCA (Cross Community Access) profile as a gateway, translate internal messages and transactions to the eHealth platform compliant messages and transactions. As more functionalities are implemented in the interface specification by the services on the facilities side, the smarter and smaller the implementation of the connector could be. According to the XCA profile, they define two actors, the Initiating Gateway and the Responding Gateway, which are the communication peers during a communication process from one community to the other community. A community by means of this profile, is “A coupling of facilities/enterprises that have agreed to work together using a common set of policies for the purpose of sharing clinical information” [IHE ITI XCA2010]. Each community must be identifiable by a global unique identifier, called homeCommunityId. So for example an Affinity-Domain which is defined as part of the XDS profile could be seen as a community by the means of XCA. The XCA profile defines some grouping rules for grouping Document Consumer/Provider with the Initiating Gateway as so called Grouped Actors.

Corresponding to our eHealth platform concept and the definition of a Affinity Domain which includes all participants of the eHealth platform, the XCA communities in our environment are the healthcare providers e.g., hospitals/institutions with their own internal communication infrastructure. So the connector could be their way to communicate with the eHealth platform. The homeCommunityId of the XCA profile corresponds to our institution-key or identifier which we will need as well, as a part of the de-identification process. Our Gateway/Connector should be a Grouped Actor supporting XDS Document Consumer and Document Provider transactions as well as the Cross-Gateway Query and Cross-Gateway Retrieve operations. Thinking about a connector in our eHealth platform definition, we need some kind of another Grouped Actor, where the Gateway e.g., on hospitals side is grouped with a Document Consumer and Document Provider actor. This depends on the applications and their communication skills concerning IHE transactions. The more of the transactions are already implemented by the software inside the institutions, the less converting functions have to be implemented inside the connector.

As the XCA profile is not yet finally defined, it is in the so called trial implementation state, there are some open questions concerning unique patient identification, vocabulary differences across communities, error handling and others. Unfortunately XCA does not match exactly with the needs we have, however it is an interesting profile concerning connectivity to the eHealth platform. During a first implementation trial we will summarize our experience and probably we can support with some ideas for enhancement towards IHE.

13.3 IHE Security, Node Authentication & Authorization ATNA & CT

Access to data in a system which provides medical data must be secured, so that only identified actors are allowed to access such data. As the defined eHealth platform consists of a plenty of services and different facilities are able to access the services, we must ensure a strong kind of security. Several IHE profiles support security aspects on different levels of the communication path. The IHE-ATNA (Audit Trail and Node Authentication) profile is one important profile for the validation of the identity of communication nodes. Each communication endpoint which wants to communicate with services of the eHealth platform, must be able to act as a secure node actor and the communication peers must perform a mutual certificate authentication. This means that the communication nodes must know and trust each other. The check of the validity of the certificate from the communication peer must be done by each node and therefore the node must also be able to act as a Time Client, specified by the IHE-CT (Consistent Time) profile. So using ATNA, the responsibility to provide access to the system (e.g., authentication / authorization) relies on each secure node. Depending on the security system and policies of the institution/enterprise, the user shall login to the node using different login mechanisms. Identities of users must be unique across the secure domain and a user may have more than one identity.

Logging of transactions and activities on the eHealth platform is an essential and important part, to be able to track changes, reconstruct possible error scenarios and access security breaches.

IHE-ATNA (Audit-Trail and Node Authentication) is the profile which covers this logging point also. It is important to ensure that logged data doesn't lead to security holes and enable possible attackers or administrators to get informations about the linkage between patient identifying and medical data. For example each access to a patients record should lead to entries in the audit-trail of the patient. The audit-trail tracks informations about "who" accessed "when", "which" documents and maybe as an option "why". The audit-trail must not disclose medical informations from reports and must ensure not to link to patients real identity. When the patient himself will be able to access his medical data, appropriate audit-trail informations will be stored too. Therefore not to disclose the relationship between the patient and his documents, one solution could be to store audit-trail data with pseudonyms.

Thinking about the eHealth platform we described so far, all nodes of the eHealth platform including the connector should implement the Time-Client and Secure-Node actor. Both

profiles, ATNA and CT are needed by the XDS and XCA profile for the management of different concerns like authentication, encryption (on transport layer) or audit trail as well as consistency among documents and so called submission sets by using consistent times for signature. The known recommendations concerning the user authentication defined in ATNA are not sufficient for our eHealth platform. They suggest that user authentication can be handled by the secure nodes, which means that an successful user authentication at one secure node (or in the local environment using EUA Enterprise User Authentication) leads to an access to the system, because the secure nodes in the Affinity Domain are in a trusted relationship. We would prefer, at least for the retrieval of documents, the strong authentication of users against the centralized authentication service of the platform by using XUA, because this prevents the system from corruption by bad local administrators which could otherwise try to misuse the system by creating wrong identities. Concerning the audit trail and logging informations which are gathered during the execution of transactions, it is not sufficient to store them at each actors side, we would prefer an enhancement to transmit those data to a centralized audit trail area. After a first trial implementation we will provide a full description as proposal for an extension of the profile.

13.4 IHE User Identification & Authentication XUA, XUA++

With ATNA and CT we can assure that the communication nodes trust each other and only known nodes are able to connect. Despite implementing ATNA, we still have to identify and authenticate the users which want to use the system. So for the eHealth platform a central storage for users and credentials is needed, where the users are registered and every user must authenticate before she or he gains access to services of the system.

Since the eHealth platform consists of many services the authenticated user can access, every service must know the user and accept and adopt the access right policies for this user. That's where IHE-XUA (Cross-Enterprise User Assertion) profile comes in place. It enables services which are called from an authenticated user, to determine the security access rights that the user has. XUA does not define actors for the User Authentication Provider and the Cross – Enterprise Assertion Provider; instead it defines two Actors X-Service User and X-Service Provider which are the participants in a web-service based transaction (e.g., XDS.b Registry Stored Query). When grouping XUA with XDS, which means that an X-Service User Actor will be grouped with the XDS Document Consumer actor, this actor will obtain a properly scoped XUA Assertion (e.g., SAML 2.0 – Assertion) which then will be used to access and authorize (on application level) calls on other services like for example the Repository or Registry.

This fits well to the concept of our eHealth platform where we defined a Security Token Service (STS) for centralized user authentication and authorization. Before a request to a service e.g., Document Registry could be established, the Document Consumer must be authenticated against the Security Token Service, which fulfills the Actors User Authentication Provider and X-Assertion Provider in one instance. After a successful authentication the X-Service User gets an X-User Assertion (e.g., a token) as part of the

response message from the STS and must use this Token as a part of further requests against X-Service provider actors (e.g., Document Registry). The STS can enrich the Token with user attributes and structural role informations.

The IHE-XUA++ (Cross-Enterprise User Assertion -Attribute Extension) profile deals with informational attributes e.g., for Role-Based access control, which could be part of the assertion (token) emitted from the Security Token Service. The profile so far is in an trial implementation state.

Since the Security Token Service is not only a X-Assertion Provider, it also can act as a User Authentication Provider, it enables Single – Sign On functionalities and needs access to databases or registers where user credentials and role informations are stored. Those could be databases or LDAP directories containing informations for example about health professionals.

A first trial implementation may also show how cross-enterprise user authentication between hospitals and the eHealth platform can be established. The important question is whether the platform's security accepts the user log-ons of an internal hospital IT-system. With this topic as well the topic of declared patient consent concerning data access is influenced.

13.5 IHE Patient Identification PIX, PAM, PDQ, XCPD

In a national eHealth platform the most important security relevant informations are the medical data and their relation to the corresponding patients. The planned eHealth platform uses a so called Trusted Third Party (TTP) to strictly separate the patient identifying data from the corresponding medical results.

Main job of the TTP service is to create, store, match and deliver pseudonyms for patients, based on their demographic data. Based on a set of demographic data of the patient, the right identity will be provided as a pseudonym. As the TTP is planned as a central service for patient identification and pseudonymization of the eHealth platform, it is in fact a kind of Master Patient Index System with pseudonymization.

So thinking about the infrastructure of the eHealth platform environment with many different participants e.g., hospitals, laboratories and so on, which use their own representation of patient data and patient identifiers, the TTP must be able to match these information belonging to the same patient but provided by different institutions.

The Patient Administration Management profile IHE-PAM defines actors and methods for the exchange of patient identity data. This profile, used f.e. in institutions, covers also transactions to exchange patient encounter/leave information or movements within a health enterprise. There are four Actors defined in PAM, the Patient Demographics Supplier (PDS), Patient Demographics Consumer (PDC), Patient Encounter Supplier (PES) and Patient Encounter Consumer (PEC). Whereas the first two Actors PDS and PDC are useful for the eHealth platform, the two encounter related Actors are not yet in focus because encounter event based transactions are not yet addressed by the eHealth platform. Therefore the Patient-Identity-Management is the one transaction which is the right one. As part of the content of

the transaction a HL7 Version 2.5 Patient identification (PID) segment containing patient demographic data must be used. As an additional and important functionality the two actors Patient Demographic Supplier and Patient Demographic Consumer supports the Merge, Link and Unlink options which also must be communicated as HL7 messages.

Referring to the needs of the eHealth platform, using IHE-PAM on the institution's side we would need implementations of a Patient Demographics Supplier Actor which is able to communicate the Patient Identity Management transaction with the Merge, Link and Unlink options.

On platform's side the Patient Demographics Consumer actor could be implemented by the TTP which must handle also the optional transactions as well as updates of patient demographic data. Since the TTP as a central point, will be the Patient Demographics Consumer for more than one Supplier, it is important for identity matching purpose, that the PID – segment which is transferred, contains the institution key of the supplier. Referring to the IHE-XCA profile this could be the homeCommunityId. Merge, Link and Unlink operations done in the local systems could be transmitted and should lead to appropriate changes on TTP side for records based on this patient identity in context of the referring institution. Matching, Linking and Unlinking of identities from different institutions inside the TTP is not yet covered by these transactions and has to be done in an separate process.

Another profile which has to be mentioned in context of patient identification is the IHE-PIX (Patient Identifier Cross-referencing) profile. This profile is used in environments with multiple patient identifier domains, where cross-referencing of patient identifiers is needed. Therefore a Patient Identity Cross-reference Manager actor is introduced, which is responsible for the mapping between the identifiers of the different domains. Local Patient Identity Source actors can use the Patient Identity Feed transaction to deliver their patient demographic data to the Cross-reference Manager. Options like Update, Merge, Link and Unlink are also possible. Main differences to the PAM profiles are the two additional transactions, PIX Query and PIX Update Notification which are supported when implementing the Patient Identifier Cross-reference Consumer actor.

If patient data are changed, the Consumer actor could be notified by the Manager actor and could be able to force queries (PIX Query transaction) to ask for a list of corresponding patient identifiers which are related to the patient identifier known by this institution. While IHE-PIX is designed to work with HL7 v2 messages underneath, the IHE-PIXv3 specification is designed to support Web-Service interfaces and the usage of HL7 v3.

Concerning our eHealth platform, the PIX and PAM profile do not fit exactly. As defined in both profiles we need functionalities of a Patient Identity Source actor (PIX) or Patient Demographics Supplier (PAM) to feed PID segment informations to the TTP. The options merge, link, unlink, update are important too. But instead of a Patient Demographic Consumer (in PAM) we need functionalities to process queries for patient identities against our TTP (Patient-Identifier Cross-reference Manager) but not exactly the way as defined in the PIX-Query transaction, because our TTP does not deliver any patient matching identifier to the systems outside.

Patient Demographic Consumer and Patient Demographic Supplier are two actors which also belong to the IHE-PDQ (Patient Demographics Query) profile, which offers transactions to query for demographic data of patients, by submitting parts of demographic data. The Patient Demographic Supplier could in fact be implemented like our TTP and can determine the right patient data by matching with its internal structures. The response of such a query transaction could be a list of patient demographic data, a data consumer can choose from. This differs from our TTP specification, because the TTP should respond a state like “matched” and a “messageid” if parameters match a given demographic dataset instead of delivering patient demographics data back to the caller. Nevertheless, it could be possible that such a functionality will be needed later on. While IHE-PDQ is designed to work with HL7 v2 messages underneath, the IHE-PDQv3 specification is designed to support Web-Service interfaces and the usage of HL7 v3.

Another profile concerning patient identification, in context of using communication gateways as defined in the IHE-XCA profile, is the IHE-XCPD (Cross-Community Patient Discovery) profile. One transaction, the Cross Gateway Patient Discovery transaction, is used to support a query to find patients by submitting demographic data. Based on the submitted set of patient data this transaction queries for matching patients from the community of the responding gateway.

The Patient Location Query transaction, the second transaction introduced with XCPD, could be used to retrieve informations about communities which may have health data of particular patients.

The first transaction seems somehow with slightly changes to fit in that way that our TTP must be able to work with identity feed and requesting for a patient by matching demographics data. The main difference is about the type of response values which could be provided by our Responding Gateway.

After a first trial implementation of the eHealth platform we may summarize our extension needs on the profile and share those ideas with the IHE community.

13.6 IHE Consent Management BPPC

The consent management and the ability to provide patients the opportunity to configure for whom they want to grant access to their medical data and how and under which circumstances, is one big goal defined for the eHealth platform.

When using the IHE-XDS profile, documents can be marked with a so called confidentialityCode to classify the basic security classification. These codes must be defined for an Affinity-Domain. But they do not define how the patient consent is realized for each single document or group of documents. For the latter the IHE-BPPC (Basic Patient Privacy Consents) profile comes in place.

With BPPC an Affinity-Domain is able to define a set of policies which can unambiguously be identified with different Object-Identifiers (OIDs). The policies should be defined in a way

that they can be understood by patients, providers and systems. The patient could then decide which policies should be used for his documents. Patients acknowledgments or signature of policies should be captured using CDA (Clinical Document Architecture) documents. Those consent declarations could be scanned copies of ink-signed paper based consent forms or electronically signed by the patient.

When a document is used, BPPC demands that the retrieving Document Consumer actor is responsible to enforce the security policies and enforce the acceptable use, so that only documents could be retrieved where the Document Consumer is allowed to. The Document Consumer actor is required to block the access to documents when the access is not authorized. If the Document Consumer actor is not able to understand at least one OID of the consent management, the access to the document must be blocked.

With BPPC it is not yet possible to define policies where patients can grant or revoke individuals (persons) the rights to access their data. Instead, more general policies are defined like opt-in/opt-out for some kind of institutions or departments, or limited access for functional (e.g., direct care) or structural (e.g., radiologist, cardiologist) roles.

The assignment of consent to the medical documents must be done by the Document Source actor as part of the XDS Metadata confidentialityCode. Therefore, based on the Affinity Domain Policy, the confidentialityCode must be set to the right OIDs of the Privacy Consent Policy Identifiers. The XDS Registry has to check that all OIDs which are used for the confidentialityCode are valid for this Affinity Domain.

Regarding to our planned eHealth platform, we face some security problems when using the IHE-BPPC profile: For our needs it is not appropriate that the Document Consumer is responsible to enforce his own access rights. In our concept the access rights as well as the proof of the consent declarations must be checked on document Registry and Repository side, so that unauthorized access can not happen “accidentally” by a Document Consumer who has not restricted himself enough. Wrong or corrupted Document Consumer implementations could lead to a big security hole. Also it is not sufficient to restrict the consent assignment to the Document Source, the document owner (patient) should also be able to make changes to the consent of the document. With a first trial implementation of these aspects inside the eHealth platform we will summarize our experience and offer them to the IHE community.

13.7 IHE Content Profiles for Medical Documents XDS-MS

Medical reports and their metadata, which are created on document source side, must conform to standards defined for the Affinity Domain. Using international accepted standards will enforce being able to share medical documents with other authorized users in purpose of the medical treatment of a patient at least across national borders.

The XDS-MS (Cross-Enterprise Sharing of Medical Summary) profile is based on the HL7 CDA (Clinical Document Architecture) implementation. It describes which medical record entries should be used and how they should be transferred. XDS-MS defines two actors, the Content Creator and Content Consumer Actor which could be used in conjunction with other profiles e.g., XDS, XDR or XDM for the transmission. XDS-MS itself is only a content

profile used to describe how the content of a document could/should be structured. The content profile is based on CDA-CRS (Care Record Summary) definition and extends it with constraints (mandatory/optional) and sections dependent on the use cases.

CRS (Care Record Summary) was the first introduced implementation guide for medical records using CDA r2. CRS covers the CDA Levels 1 and 2 and should specify Level 3 templates (CDA entries).

Another standard for a patient health summary medical record which is often referenced is CCR (Continuity of Care Record). CCR is a document structure definition based on using XML, introduced by the ASTM (American Society for Testing Materials) organization. The aim is to define documents which contain the relevant and current health informations about a patient. These documents could be useful at the time of the beginning of a new medical treatment or at clinical encounter. For example Google Health uses a subset of the CCR specification for vendors of healthcare systems to provide their informations to the eHealth platform. A main critic with CCR is, that it is not able to deal with more complex information e.g., treatment and workflow inside hospitals.

Another CDA content structure which should cover this limitations is called CCD (Continuity of Care Document). ASTM (American Society for Testing Materials) and the HL7 group joined their effort to create a harmonized data format based on the CCR and HL7's CDA specification.

Although the CCD specification was applicable in the United States only, the new version distinguishes between a universally applicable version and a derived/constrained version for use in the US. The CCD standard is the preferred way to communicate electronic health records with the Microsoft HealthVault platform, although HealthVault is also able to deal with CCR documents.

When thinking about our planned eHealth platform, content definitions for medical summaries should base on the given standards already defined in the content – module sections of the IHE-PCC (Patient Care Coordination) framework and the specific content module sections for the proposed healthcare domains e.g., laboratories, which cover the specific requirements. The IHE XD-LAB (sharing laboratory reports) content module is one of these specifications which covers the vocabularies and clinical document structures for laboratories.

The efforts for detailed definitions and the check for usability of certain document structure definitions according to the needs of the Luxembourgish healthcare system and the eHealth platform, have to be done in detail regarding to the planned development steps of the eHealth platform. The specific requirements of the eHealth platform regarding the security concepts of documents has also to be considered.

Apart from the specifications for the contents of documents, there are other important issues which have to be recognized when sharing medical documents. One important issue is the way document states are represented. For example, the XDS metadata for a XDSDocumentEntry contains a field which is called availabilityStatus. This field could represent the following document states like:

Approved = completed; Deprecated = obsolete; Deleted = nullified. Accessibility status could for example be used for a document replacement. In IHE XDS such a replacement of a

document must be done using a transaction which deprecates (`availabilityStatus = Deprecated`) the old document, submits the new document (`availabilityStatus = Approved`) and creates a reference between both documents with the replacement type (RPLC) .

Although these states are necessary and useful when accessing a document, but this field does not cover different workflow states of document usage. Common states during a document usage could be e.g., for a referral document: referral placed, document read, admission in progress and so on. The IHE-XDW (Cross-Enterprise Document Workflow) profile which is currently still in development phase, should deal with such workflow state transitions of documents. Therefore an additional document should be stored and registered which contains as body sections of a CDA document the states which the original document has passed.

13.8 Closing the Gaps with Connectors

We can not expect that primary systems are prepared to provide their information to the eHealth platform. Therefore a connector between each primary system and the eHealth platform is foreseen. A connector is a part of software running on the side of a primary system. It prepares the compiled information syntactically and with respect to catalog mappings, semi-semantically. The connector sends the prepared information to the eHealth platform. Main jobs of the connector are:

- Copy-Extraction of the patient identifying data
- Extraction and generation of relevant meta data out of the documents
- Institutional or transport signature of the documents
- Encrypting the documents with TTP's public key
- Send the encrypted document, the meta data, and the encrypted document-key to PMIP
- Send the patient identifying data to the TTP
- (Statistics) Strip off person identifying data from the original document
- (Statistics) Encrypt stripped document with public key of dedicated statistical user
- (Statistics) Provide the stripped document to a statistical component of the PMIP

In contrast to pushing connectors, the pulling connectors connect primary systems with the eHealth platform for the purpose of querying results out of the platform. They split up queries in patient identifying data for the TTP and send the remaining part of the medical query to PMIP. A pulling connector will be one part of the web-server that enables access to the eHealth platform for health professionals and for patients. Larger institutions like hospitals or laboratories may get a direct access to the eHealth platform without a web-server in between. As consequence they have pulling facilities included in their connectors. Details of the connectors will be defined and then described in a technical reference document.

14 Iterative Evolution

The authors suggest an iterative approach for (1) research, (2) concepts & requirement definition and (3) realization phases of a test-bed as well as (4) for the productive system. Concerning research and the resulting concept, *this document* is the first iteration. The

following tasks are iterative by themselves.

- Research and state-of-art.
- Concept description of product backlog together with the future users.
- Implementation for the test-bed and description of the technical requirements.
- Implementation for the productive EHR system.

Here the outcomes of an iteration of one task impacts the next iteration of its successor task. i.e., the outcome of the *research* impacts the *concept*, the *concept* drives the *test-bed* implementation as well as the technical requirements definitions. *Requirements definitions* are the input for the *productive realization*.

Example: While the research is in iteration n+3, the conceptual phase is in n+2, the test-bed is in n+1, and the productive implementation is in iteration n.

As of March 2011, the research task is in the 2nd iteration, the concept task will close its 1st iteration with *this document*. The implementation of the test-bed will just start its 1st iteration in short. Therefore the product backlog for the 1st test-bed iteration is described next. The productive system is not yet specified as of today.

15 Test-Bed Backlog for 1st and/or 2nd Iteration

After studying the IHE profiles in chapter 13 including their recommendations and references to real standards of HL7, W3C, ANSI, etc., the next steps to go are different proof of concepts within a test-bed. The setup of an appropriate infrastructure of components and its services will give an impression of the complexity. The aim is the realization of the described concept (iteration 1) with the build-in security, regarding pseudonymization and cryptography.

TTP and PMIP Separation

The main parts of the architecture are the Trusted Third Party Provider TTP with the services of pseudonymization and re-encryption of symmetric document keys, and the Pseudonymized Medical Information Provider PMIP as information storage for encrypted medical documents and encrypted symmetric keys.

PKI and STS

Access protection is given by a surrounding PKI infrastructure which checks the identities of potential users, a Security Token Service STS assigning access rights to those users according to their declared user access rights. The STS itself uses the PKI certification authority together with the health professional registers and general citizen registers. Those register may get their initial information out of existing registers of the communes or the health insurances. Update information may be delivered by the source registers as well.

Pass-through Webserver

Associated to this depicted infrastructure is a pass-through Webserver which routes through the encrypted patient identifying information from Web-Users to the TTP and routes back

again the results belonging to requests from the PMIP to the Web-User. Because the routed material is encrypted the Web-Server is not a target for successful hacker attacks or curious insiders.

IT-consent Management

In the internal part, associated to the PMIP, an IT-consent Management Component ICMC crosscheck every query of a requester, whether answering is allowed or not. – This with respect to the declared IT-consent of the patient. In a first version the IT-consent is a “yes-all or no-nothing” implementation.

CMReg, CMRepo(s), DSA(s)

The Central Medical Registry CMReg as component operated by the PMIP is the central access point for every medical query. CMReg holds the meta data, it crosschecks the patient's given IT-consent before addressing the Central Medical Repositories CMRepo or the Decentralized Staging Areas DSA for retrieving the encrypted documents.

DSAs are provided by health care delivering organizations, i.e., hospitals, laboratories, or operated by a group of general practitioners or medical specialists. From the viewpoint of the eHealth platform, CMRepo is an internal repository while DSAs are external repositories. The CMRepo as well as DSAs contain encrypted medical documents, which are addressed via the central CMReg entries. Only one central CMRepo is foreseen for the first iteration of the PMIP in the test-bed.

XDR Implementation based on XDS Infrastructure

With the solution mentioned in section 7.10 – PKI for Authentication, for Signature and for Cryptography – the extra maintenance of a PKI infrastructure for encryption keys (besides the PKI for authentication and signature) is avoided. All trusted actions are based on the authentication and signature features of the certificate. The drawback of this procedure is that the simple communication P2R – provider to known receiver – can not be implemented by simple secure eMail services. – There is no public key available for the known receiver! A pull-component where the intended receiver asks for the information is necessary. This requester delivers his ad-hoc generated public key, and the sender (here the eHealth platform) has to re-encrypt the result (the symmetric document key) with this public key. The good news: This infrastructure is already in place with the implementation of P2MUR – provider to multiple unknown receivers, i.e., the core of the EHR implementation.

Primary System Connectors

At least for the beginning, the information providers and consumers may not be able to connect to the platform's components with their operational software systems. Therefore Primary System Connectors PSC are in place. Logically they belong to the platform, but they are operated by the health care providers, i.e., on the locations of the primary systems. The PSCs connect operative inhouse systems with the platform. Over time the functionality of the PSCs will be re-implemented more and more inside of the primary software systems. If once the communication of a primary system with the platform components is fully standardized,

the PSCs are (nearly) empty with respect to functionality. Nevertheless they stay alive for future non-standard extensions of the platform or for older operative systems of data providers.

Putting it all Together

The implementation starts with a setup of a rudimentary infrastructure of components as mentioned above. They must communicate with each other according to the concept. The test-bed implementation will start with the CMReg and one CMRepo. The commercial available TTP of the Luxembourgeois Biobank (IBBL) project can be reused. Alternatively the available SANTEC TTP can be used. This decision depends on the main needs and the usability of the commercial TTP versions. A rudimentary STS is next, including test access to a certification authority service for the PKI. The first version of the IT-consent management component will be an all-or-nothing solution, i.e., the access to each information-item is flagged in the CMReg's meta data with YES or NO.

A health professional register with fictive names will be used for test cases, allowing or disallowing test users the access to the TTP and PMIP. The implementation of the user rights database will be rudimentary via SQL inserts to the underlying database. A maintenance application for the user access rights is not in a first iteration because the knowledge how to build this is neither new nor critical.

The web-server for passing through a patient's name that is encrypted has to be part of the proof of concept. The requesting web-application has to hide the patient's identity from the corresponding web-session. The realization may be an applet or any other appropriate solution.

A further part for the proof of concept is the encryption and re-encryption process with its related components. This part of the eHealth platform has to be proofed to work efficient and reliable.

This rough description is an overview of “proof of concepts” that should be demonstrated within an initial test-bed. With the experience that SANTEC will build-up during the implementation of that test-bed *edition one*, the technical requirement descriptions will be evolved, noted, tested, confirmed, and lastly recommended for a tender for industrial production. This description will be the basis for negotiations with the industrial IT companies that should implement and operate the components of the productive eHealth platform in Luxembourg – and elsewhere.

16 Next Iterations for Research, Concept, and Test-Bed

The next proposed research topics, state-of-art analysis and planed conceptual descriptions are sketched here for short. They describe possible next evolution steps in the test-bed implementation. The decision which of them will be in the product backlog for the next test-bed iteration cycle, has to be taken by the stakeholders in short future. A more detailed planing of the selected backlog for a next iteration will take these opportunities as an input. According to the agile methodology SCRUM the decision will be taken during the planing meeting for the product backlogs.

Maintenance Application for Role Based Access Rules

The health professional register and/or user register is maintained via direct database modification of the STS database tables per SQL. A maintenance application for insert, modify, and delete operations on the users registry may be useful for getting further experience with the needs of eHealth platform administrators. This can be a standard application for maintaining user access rights and registering users with their card identity parameters.

Data Aging and Archiving

Data aging and over-encryption in archives is not in the initial test-bed implementation. Nevertheless, the concept is already described in section 7.6. This concept can be part of the product backlog for a next iteration of the test-bed implementation.

Patients' Access to their own Records

The access of patients as users to their own health records is an important topic for the test-bed product backlog. It is mandatory for the first version of the productive system. Here the anonymity of the patient has to be guaranteed – i.e., the patient has to stay anonymous for the web-server where he is logged in – as a user. The Security Token Service STS gets some extra work with this requirement. In the product backlog for the next concept iteration this topic can be included. The final decision again belongs to the stakeholders.

Alert Functions and Access Logs

Every access to patient data inside the eHealth platform has to be logged. The patient will be able to monitor who has stored, updated or retrieved information of his health record. In section 7.11 is already mentioned that the patient can declare how he wants to be informed about curious accesses to his records. This may be per SMS, eMail, or the yearly paper print-out as default. Besides this automated notifications, the patient should be able to check the access history of his record interactively. This feature should be in one of the first product backlogs, as it is mandatory already concerning the Luxembourgian law.

Meta Structuring EHR in Pseudonymized Space

Chapter 11.2 describes the work of organizing the health record in an appropriate structure, i.e., composing all the different information pieces provided by the primary care delivery organization. For the beginning a chronological order of all received information items (labo reports, discharge letters, consultation letters, etc.) forms the log-book of the corresponding EHR. The contents of those documents are encrypted. Only the meta data inform about the document type. The patient's identity is replaced by his pseudonym. Without resolving the pseudonym or violating the encryption, the knowledge that can be used for structuring the EHR, is limited to the meta data of the documents. Additionally – in case a human performs the structuring work with the EHR – this one may know the content of some documents and use this knowledge for better structuring the EHR. This on the other hand results in the risk of so called profiling. A reference doctor for example may start to reorganize the EHR of a

patient. As he knows the content of a special laboratory report, he may decide to classify this report inside the EHR structure. At the end, the EHR of a patient will get organized based on the meta data of the encrypted documents, while the patient's identifying are unknown at the PMIP.

After gathering the first experiences on this basis, a good structure for an EHR will evolve as best practice out of these first rounds.

IT-Consent: Rule based, Role-Based and Item based

The next generation of the IT-consent concept goes beyond the “yes-all or no-nothing” solution of the first version. IT-consent can be declared with rules and based on health professional roles. The so called habilitation matrix helps for a proper design with roles. An implementation concept for rule and role based IT-consent is a topic of current research. Afterward the concept leads to an implementation in a later stage of the test-bed.

Embedding external Service for Affiliation Check

As first example the embedding of the affiliation checks can be implemented and tested in the test-bed. The web-service will be designed like described above and offered in an internal test system. The TTP variant will be implemented so that its security conformance implementation can be proofed. As result a recommendation for the productive implementation will be proposed.

Embedding external Service for Demographic Data Improvement

The TTP deals with patient identifying data. An access of the TTP to a national person register helps to improve the quality of demographic data inside the primary systems and so inside the TTP. If the TTP detects two sets of patient demographic data as potential duplicates, the TTP can inform the primary systems about. The user on side of the primary system has face-to-face contact with the patient and can verify or falsify the hint of the TTP. In case of verification the primary system initiates a patient-data-merge and publish this to the TTP. TTP trust on such merge notifications and merge both patient-data-sets as well.

Outsourced Global Services

Besides the eHealth platform components and the blind Webserver, additional services will be established working on the PMIP and TTP information. For example, a service for overview reports can be addressed by a general practitioner with the job to create an overview report for a patient. The access rights of the doctor are inherited to the service and the service can act on the TTP and PMIP, respecting the IT-consent of that patient – given for the requesting doctor.

Decision Support Systems

Based on the accumulated medical information medical decision support systems can act on the data. SANTEC runs already a project on clinical decision support. As soon as the project produces some results, they may be useful as well for the eHealth platform usage. Here the

critical point is the encryption of any document information inside the eHealth platform repositories. Again a solution can base on the concept of outsourced global services, inheriting the doctor's access rights and respecting the patient's IT-consent accordingly.

Versioning of Documents and Environments

Documents and structure giving folders can be replaced by newer versions of the same document. For a first step each document is considered as a new one and it may replace its older pendant. Research work has to be invested in the versioning of documents. In the IHE context the so called XDS Metadata Update is available for trial implementation as of August 10, 2010. Here the notation for different versions of a document are noted. Not every new version may replace an older one in a way that the older one can be erased.

As example a document B refers to a document A. At the same time when a new version of A should replace the current A, the first version of document B is put in place. So both versions of A are needed in parallel. Three days later a document C may refer to the first version of B. If now four days later B gets a version update, then C still refers to the first B, which itself still refers to the the first A. Unfortunately the referenced versions can not be replaced automatically by their successors because the reference reflects to the content and meaning of the documents. For each document a hint can be placed that there is a new version available. Handling this situation in an intuitive way can be a task in a next backlog of an next iteration. Concerning the environment, a future PDF/JPG/X-Ray/... reader may show more or less than the version of today. A user in the future does not see the same even if he accesses the same PDF/JPG/X-Ray/... document. This happens for any software that is involved.

Scheduled Pseudonym Exchange and Multilevel Pseudonymization

In case of serious attacks and the insight that the security level of information protection is not sufficient, these two extensions can be implemented additionally. Pseudonyms will be exchanged periodically between TTP and PMIP which stays transparent for the users. Only for a data theft the barrier is higher, because his two stolen copies of the TTP database and the PMIP database have to be done at the same time. Multilevel pseudonymization enlarges the number of needed criminal administrators over the initial two (TTP and PMIP) up to N. The costs may explode because of the (N-1) separated organizations which operate the (N-1) pseudonymization services.

Integration of HomeCare data collections

An EHR is a perfect place for collecting measurement data of home care monitoring devices. A concept for automated integration of such data, that is send out of measuring devices at the patient's home may enrich the usefulness of the EHR information.

Open for Further Extensions

Any new idea is welcome as input towards the next iteration steps of research, concept drawing, test-bed implementation or productive environment.

17 References

- [BE-GOV2010] Frank Robben: *The eHealth platform as a support of high quality healthcare and administrative simplification*, June 2010, https://www.ehealth.fgov.be/fr/page_menu/website/home/platform/mission/presentation.html, last visit February 17 2011.
- [BeSi2011] S. Benzschawel, M. Da Silveira: *Protecting Patient Privacy when Sharing Medical Data*, eTELEMED 2011, *The Third International Conference on eHealth, Telemedicine, and Social Medicine*, Guadeloupe, 2011
- [D2D-Padok] *Die Telematik-Plattform der Kassenärztlichen Vereinigungen*, <http://www.d2d.de/index.php?id=57>, access to the internal Padok Homepage is protected. But access passwords are named there.
- [EFES-2010-WP10] eSanté EFES Report WP10-1, *concept paper on a national EHR*, eSanté team, Version 1.1, April 2011, [http://www.santec.lu/project/esante/efes/Deliverable 10](http://www.santec.lu/project/esante/efes/Deliverable%2010), last visit May 06 2011.
- [EU-Dir95_46_EC] *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data*
- [Fraunhofer2005] Fraunhofer Institut: *Spezifikation der Lösungsarchitektur zur Umsetzung der Anwendungen der elektronischen Gesundheitskarte (2005)* (referenced by Neubauer)
- [GaDa2006] Dave Garets and Mike Davis : *Electronic Medical Records vs. Electronic Health Records: Yes, There Is a Difference . A HIMSS Analytics™ White Paper . Updated January 26, 2006*
- [Gartner2003] Gartner Research, *Comparing the TCO of Centralized vs. Decentralized ERP*, published January 2003, last access Mai 16 2011, <http://www.gartner.com/resources/112700/112745/112745.pdf>.
- [Haas2005] Peter Haas: *Medizinische Informationssysteme und Elektronische Krankenakten*, Springer-Verlag, 2005
- [Hohmann2010] J. Hohmann: *The use of Medical Data in Research and eHealth applications - Can European Data Protection Law keep pace?*, 18th World Congress on Medical Law, Zagreb, Croatia, August 2010.
- [IHE-profiles] IHE International. *IHE Profiles*. URL: <http://www.ihe.net/profiles/>
- [IHE ITI TF2010] *IHE Technical Framework Volume 1 - 3, Revision 7.0 Final Text*, August 2010
- [IHE ITI XCA2010] *IHE Technical Framework Supplement Cross-Community Access (XCA) Trial Implementation*, August 2010
- [IHE ITI PCC2010] *IHE Patient Care Coordination (PCC) Volume 1 + 2, Revision 6.0 Final Text*, August 2010
- [IHE ITI DSUB2010] *IHE Document Metadata Subscription (DSUB), Trial Implementation Supplement*, August 2009

- [IHE ITI XDW2011] IHE Cross-Enterprise Document Workflow (XDW), Draft in preparation for Public Comment, January 2011
- [IHE ITI XUA++2010] IHE Cross-Enterprise User Assertion – Attribute Extension (XUA++), Trial Implementation, August 2010
- [IHE ITI XCPD2010] IHE Cross-Community Patient Discovery, Trial Implementation, August 2010
- [Kerckhoffs1883] Kerckhoffs's Principle. Explained in http://en.wikipedia.org/wiki/Kerckhoffs%27s_Principle, last access Nov 24 2010.
- [Mann2008] Ministerialrat Bernhard Mann, LDI NRW: Offene rechtliche Aspekte im Zusammenhang mit der Konzeption einer eEA, Landesbeauftragte für den Datenschutz und Informationsfreiheit Nordrhein-Westfalen, 2008
- [Neubauer2009] T. Neubauer, M. Kolb: An Evaluation of Technologies for the Pseudonymization of Medical Data, in: R. Lee et al. (Eds.): Computer and Information Science, SCI 208, pp.47-60, Springer-Verlag, 2009
- [OASIS ebXML-RIM2005] OASIS ebXML Registry Information Model, Version 3.0, May 2005
- [Peterson2003] R.Peterson: Encryption system for allowing immediate universal access to medical records while maintaining complete patient control over privacy. US Patent Application Publication, No.: US 2003/0074564 A1 (2003) (referenced by Neubauer)
- [Pommerening2004] K. Pommerening, M. Reng: Secondary use of the Electronic Health Record via pseudonymisation. In: Medical And Care Compunetics 1, pp. 441–446. IOS Press, Amsterdam (2004) (referenced by Neubauer)
- [PRU2010] C. Pruski, e-CRL: A Rule-based Language for Expressing Patient Electronic Consent, Proceedings of the Second International Conference on eHealth, Telemedicine, and Social Medicine (eTELEMED 2010), IEEE Computer Society, St. Maarten, (Netherlands Antilles), 2010
- [PwC-study-2010] Luxembourg Ministry of Health , eHealth Service Platform Study , pre-final version: eHSPS_Report_v0.53.docx , 2010
- [Quebec2011] Santé et Services sociaux Québec. Data Security. Explained in http://www.dossierdesante.gouv.qc.ca/en_citoyens_securite.phtml, last access May 2011.
- [Stingl2007] C. Stingl, D. Slamanig: Berechtigungskonzept für ein e-health-portal. In: Schreier, G., Hayn, D., Ammenwerth, E. (eds.) eHealth 2007 - Medical Informatics meets eHealth, vol. 227, pp. 135–140. Österreichische Computer Gesellschaft (2007) (referenced by Neubauer)
- [Thielscher2005] C. Thielscher et. al.: Patent: Data processing system for patient data. Int. Patent, WO 03/034294 A2 (2005) (referenced by Neubauer) February 2011.