EHR4CR

Electronic Health Records for Clinical Research

Executive Summary for deliverable D5.1: Requirements and specifications of the security and privacy services

Project acronym: EHR4CR
Project full title: Electronic Health Records for Clinical Research
Grant agreement no.: 115189
Budget: 16 million EURO
Start: 01.03.2011 - End: 28.02.2015
Website: www.ehr4cr.eu

Coordinator: AstraZeneca

Managing Entity: EuroRec

imi

The EHR4CR project is partially funded by the IMI JU programme

efpia

Document description

Deliverable no:	D5.1 (executive summary)		
Deliverable title:	Requirements and specifications of the security and privacy services		
Description:	First deliverable of Work Package 5		
Status:	Final		
Version:	0.3	Date:	14/08/2012
Deadline:			
Editors:	Roland Krause		

Document history

Date	Revision	Author(s)	Changes
29/06/2012	0.1	Roland Krause	Initial draft
11/08/2012	0.2	Roland Krause	Version after revision
14/08/2012	0.3	Roland Krause	Final version

Table of Content

1. Executive Summary

Work Package 5 focuses on *Requirements and Specifications of the Security and Privacy Services* from a legal and technical point of view. This document provides an extended executive summary of the deliverable D5.1 (Report on Requirements and Specifications of the Security and Privacy Services).

Privacy is one of the fundamental issues in health care today and a trade-off between the patient's demands for data privacy protection as well as the society's need for improving efficiency and reducing costs of the health care systems is necessary. The systematic usage of electronic health records is expected to improve communication between all health care stakeholders and access to data and documentation, leading to better clinical and service quality.[1]

The following descriptions in the deliverable have been **drafted jointly by Work package 1 (Task 1.4) and Work package 5 (Task 5.1)** to complement functional requirements that have already been developed to specify how protocol feasibility services should be implemented. These requirements are called non-*functional requirements*, and primarily focus on the ethical, legal and regulatory requirements that the platform should meet.

There are several **legal and governance challenges** that are being tackled within this project:

1. Ensuring that the identity of individuals whose data are analysed, communicated and stored externally outside the hospital environment or with the help of the platform architecture is protected, and that the processing of all personal data conforms to EU data protection legislation and to national and regional policies;

2. Ensuring that the platform minimises the risk of identification of individuals and maximises organisational, jurisdictional and societal trust by keeping to a minimum the direct access to identifiable data, minimising the communication of patient-level data across borders, by using aggregated and non-identifiable, pseudomized data.

3. Ensuring that the information handled by the platform and its services are of sufficient quality, and have suitable provenance metadata, to be used for regulated clinical research;

4. Understanding the ethical and privacy protection concerns, and expectations, of stakeholders and decision makers in public bodies and industry;

5. Ensuring that the pilot sites have the necessary permissions to utilise their in-house EHR data for this project's research, and are applying good practices in information governance and information security; and

6. Identifying the key governance success factors for the business model and platform for long term acceptability.

[1] AHIMA. "HIM and Health IT: Discovering Common Ground in an Electronic Healthcare Environment" Journal of AHIMA 79, no.11 (November–December 2008): 69-74.

2. Scope and Overview

The goal of WP5 is to provide the tools and services to ensure that process and data flows on the EHR4CR platform, dealing with particularly sensitive information (patient-identifiable data), are compliant with legal and ethical requirements defined in WP1 Tasks 1.3 and 1.4. There is also a potential for interaction with Task 1.4, as it is necessary to identify the regulations and cross-border legislative issues that EHR4CR must address and comply with. This task is jointly undertaken with WP1. Therefore, this document will have cross-references to the corresponding document, D1.1, regarding requirements.

The deliverable is organized into the three sections:

- Part 1: State of the Art: Inventory of the privacy and security requirements for the EHR4CR Platform
  - Design of requirements
- Part 2: Analysis of Data Flow in the first 2 Scenarios and Analysis of Auditing functions
  - This work-in-progress document will be revised after each reporting stage, thus complementing the document throughout the project lifecycle.
- Part 3: Applicable laws and bylaws on a European, National and local level compilation *(not contained in this executive summary*).

In the initial stage of the project, the technical requirements for these different scenarios will be drawn up from the identified relevant laws and regulations (WP1, Tasks 1.3 and 1.4). They will serve as input for defining (and implementing) a standard definition of based trust, security and privacy-protecting layer of the EHR4CR architecture.

The data and process flows of the detailed scenarios defined in Task 1.2 will be evaluated to ensure that the EHR4CR data protection framework covers all needs encountered in practice. At this stage, the first two scenarios, protocol feasibility and recruitment, are suitable for analysis and will be examined in depth. For example, in the “protocol feasibility scenario” highlighted above, we will utilize solutions to concentrate on processing of patient-identifiable data at the data source side and that only export aggregated result sets (which constitutes an effective means of data protection). Architectural design is based on the description in Document “D3.1 Initial EHR4CR architecture and interoperability framework specifications” for the Protocol feasibility scenario.

Further, the query engine process is analysed and indications given as to which legal safeguards apply to each of the steps involved. This is equally applicable to the audit model. Auditing is an important aspect for achieving compliance. Basic auditing mechanisms will typically allow individual events within a closed system to be logged and interconnected. Few standards and solutions are available for providing manageable audit trails in distributed systems, and may need to be developed according to the requirements identified. Standards efforts to be taken into account include:

- ISO standards 13606 Part 4 (Health informatics -- Electronic health record communication -- Part 4: Security);

- ISO 27789 (Audit trails for electronic health records, draft); and
- Implementation standards such as IHE ATNA (Audit Trail and Node Authentication).

Anonymisation makes long-term storage possible; therefore, researchers should use anonymous data and material whenever possible. However, in many cases of research, the correct association between a single patient's data from distinct sources or distinct points of time is crucial. Some scenarios, such as the Serious *Adverse Event Reporting* (EHR4CR Scenario 4), even require a way back to the identity. Pseudonyms are a viable solution for long term storage.

Several legal conditions increase restrictions for the transfer or storage of data for research purposes, for example data protection regulations and professional confidentiality. Further details about the applicable laws are explained in the full deliverable. The use of data and material for research also presupposes the informed consent by the patients. But even with consent, data must be transferred and stored only for a defined purpose, with restricted time frames and explicitly authorised users, required to be mentioned in the protocols' patients' information.

Extending these restrictions is possible in certain circumstances, but only by applying additional safeguards and conditions in a rigid organizational framework, where the risk of re-identification is closely monitored. It is of utmost importance that EHR4CR complies with these requirements, as outlined in an overview below.

2.1. Content Description and Objectives

The objectives of this document are to formally specify the data protection framework for the envisaged EHR4CR System. The result of the inventory will be used:

- to define a first version of the EHR4CR data protection framework (task 5.2)
- to examine current development stages of the architecture with regard to legal compliance
- To outline the legal framework to be taken into consideration by other EHR4CR work packages in designing the platform.

There are five objectives that have been tackled during year one. These are:

Protecting the identity of individuals

The main year one activity for this objective has been to produce an inventory of relevant legislation, standards, regional and national policies and regulatory obligations. This work has been undertaken in collaboration with WP 1 and the Ethics Board of EHR4CR. This inventory and the analysis of the environment are presented in this document. The main contribution to the design of the platform is the collection of legal requirements.

Ensuring the data used by the platform meets regulatory requirements

The work ensuring the platform data meets regulatory requirements is at an early stage, with the higher-level requirements having been included in Deliverable 5.1, and a more formal specification to be developed during year two in collaboration with the architecture work package 3.

Wider stakeholder viewpoints

A detailed interview questionnaire, and briefing pack for interviewers, has been developed through Task 1.1, reported in that Deliverable. That task report also reports on an ethics workshop specifically focusing on the EHR4CR platform. Key findings from that workshop served as input to the WP5 Workshop that was held in January 2012 on Legal and ethical issues.

Governance of the pilots

Building on work undertaken in past EC projects, a comprehensive template has been developed by WP 1 to capture and collate the necessary information from the pilot sites regarding their legal status. In parallel, the pilot work package, WP 7, has been refining their EHR data requirements from each site (cf WP7 Round 1 Data Export). The sites are now in a position to report on the approvals and protection measures needed to utilize this EHR data, and so are beginning to populate the forms "*Register of ethical-legal and security policies"*. The project will centrally and securely hold a register of these approvals and measures for project governance purposes and, in the longer term, produce a guidance pack for new hospitals connecting to the final EHR4CR Platform. Evaluation of this template with regard to legal issues was given.

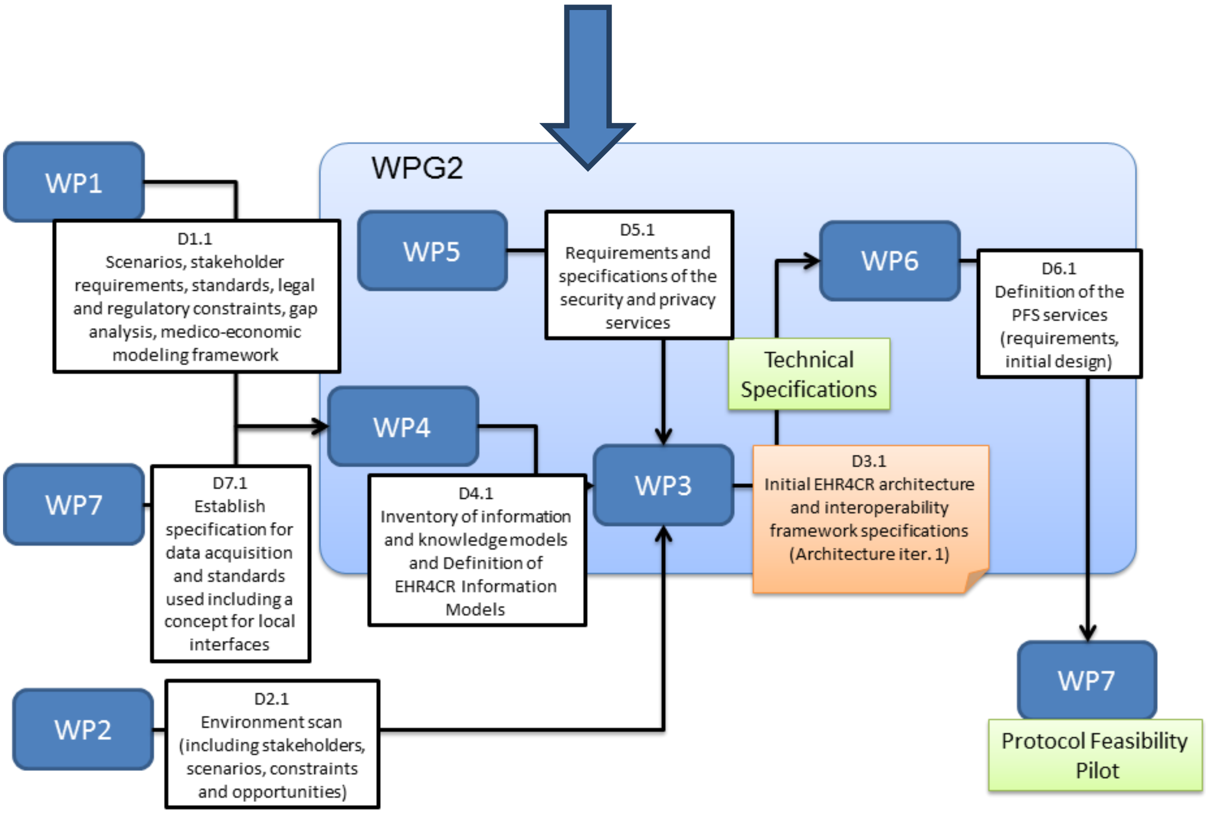
Ethical inputs to the Business Model

Specific questions exploring the importance of ethical issues were included within the Environment Scan survey undertaken in the summer of 2011. The results of this have been published in December 2011, and are reported in Deliverable 2.1. Ethical issues impact to a certain extent the legal framework, as biomedical research in many countries is subject to approval from an Ethical Review Board (ERB) at an organizational, regional or national level.

2.2. Relationship to other parts of the project

- The output of work package 5 has a crucial relationship with the majority of the requirements of the project. The platform must comply with the requirements set out in this work package, thus complying with legal and ethical requirements. A platform that does not meet the legal requirements discussed herein, cannot be operated and therefore would not constitute a basis for neither a viable nor sustainable business model. In the further design process, there must be a close interaction between the architectural design group [WP3] and this work package team.
- As this work package builds on the requirements established in task 1.4, the leaders of work package 1 and work package 5 have decided to join forces and combine their efforts towards a detailed work package 5.1 output. As such, the present document contains the output of work package 1.4 and 5.1 and is serving as a reference documents for WP 1.3 and 1.4.

The interaction of Task 5.1 within WPG2 as a subgroup in EHR4CR can be summarized graphically as follows:


WPG2
WP1
D1.1 Scenarios, stakeholder requirements, standards, legal and regulatory constraints, gap analysis, medico-economic modeling framework
WP5
D5.1 Requirements and specifications of the security and privacy services
WP6
D6.1 Definition of the PFS services (requirements, initial design)
Technical Specifications
WP4
WP7
D7.1 Establish specification for data acquisition and standards used including a concept for local interfaces
D4.1 Inventory of information and knowledge models and Definition of EHR4CR Information Models
WP3
D3.1 Initial EHR4CR architecture and interoperability framework specifications (Architecture iter. 1)
WP2
D2.1 Environment scan (including stakeholders, scenarios, constraints and opportunities)
WP7
Protocol Feasibility Pilot

- Graphic © WP 3 Electronic Health Records for Clinical Research News EHR4CR_D3.1_ Architecture Description.docx

3. Part 1: State of the Art: Inventory of the privacy and security requirements

3.1. General data protection within EHR4CR

At this stage of the project the general applicable principles are laid out. The exact implementation and orchestration within EHR4CR can only be made once all use cases / scenarios are defined. Therefore, the following will include general guidelines and rules:

3.1.1 Data controller within EHR4CR

The research institutions cooperating within EHR4CR at the current stages would have to look into establishing a legal entity ([WP2] which would be the Data Controller for the EHR4CR project), defining the legal responsibilities for the consortium with a clear focus and data protection and privacy and appoint a Data Privacy Officer.

This appointed Data Privacy Officer shall oversee all operations of the platform within an independent legal EHR4CR entity to be formed. He/she shall advice the partners on implementing the data protection framework and ensures that data provider sites understand and implement the rules established for the use of the platform. A chain of trust and compliance shall be established (further detail will be explored in Task 5.2).

3.1.2 Trusted third party service

At the current time, there is no Trusted Third Party (TTP) service used in the Protocol Feasibility scenario.[2] It is likely to be used in Scenario 4.

3.1.3 Data linking

The platform architecture shall ensure that data related to each specific subject of care are only identified through a project-specific pseudo-identifier (a code) issued and managed in accordance with good practices as defined in ISO TS 25237. The matching of specific individuals to enable pseudo-identifier generation and subsequent linkage at any healthcare site shall only occur within systems and services located and governed in accordance with ethical approvals pertaining to that site. These systems and services may be designed and implemented by EHR4CR or by accredited third parties.

The platform and its health data repositories shall be able to include or reference information that supports the unique identification of the subject of care without relying upon demographic descriptors, photographs, biometric properties, health service issued patient identifiers or other explicitly recognising traits. Further, EHR4CR shall specify the relationship of any other persons who are the subject of EHR information to the subject of care, if this is explicitly given in the source EHR data e.g. “relative/spouse/father” but shall not in any other way identify such persons (no third-party identifiable personal data shall be recorded).

3.1.4 Access control

General access control to any component of the EHR4CR system shall ensure that no unauthorized person has access to the system. It must be decided whether there are different levels of access limitations for different categories of authorized users. Currently there is no need for providing federated identity management, but this might change with future scenarios. Upon first access to the EHR4CR platform, the end-user must be authenticated. The end-user interface itself is not responsible for performing the actual authentication, but delegates this to the central EHR4CR authentication provider (cf. Architecture Description (D3.1) 7.4.2.2). There shall be a declaration of fair usage, which must include the following non-exclusive list of items, not limited to the scope of the project, but also taking into consideration concerns voiced by pilot sites regarding data use in general:

- No reverse engineering with regard to patients data;

[2] Architectural design is based on the description in Document “D3.1 Initial EHR4CR architecture and interoperability framework specifications”.

- No intent to retrieve quality data about individual hospitals;
- No singling out of individual patients through various repeated queries and
- Obligations not to divulge access codes to the system to any 3rd party
- Monitoring of any action taken in the system.

It is further recommended that:

- The Platform shall enable the implementation of policies that control access for use and access for onward disclosure of EHR data items (= patient information) and aggregate information to authorized individuals and computer systems;
- The Platform shall enable access to EHR data items in the hospital warehouses and aggregate data to be filtered to match the access permissions of the individual or organization in accordance with their role, context and purpose of use;
- The Platform should enable the implementation of policies that dictate specific periods for data retention, if explicit or implied consent or health system policies or jurisdictional regulations do not allow EHR data to be held indefinitely;
- The Platform should desirably enable the implementation of policies that specify the circumstances under which record linkage to other databases is authorized as well as those under which it is expressly prohibited;
- The communication of EHR extracts with the Clinical Data Warehouses (as original data items or in aggregate form) shall comply with the access policies pertaining to the information being communicated;
- EHR extracts communicated to the Platform or by the Platform shall be able to reference any policies that specifically apply to the information being communicated;
- Platform access policies shall permit authorized (governance) users to grant or to restrict access to nominated EHR information to identified individuals, to specific structural or functional role groups, to specified organizations, for specific data uses as required;
- Platform access policies shall specify a time interval over which their stipulations apply;
- Platform access policies shall identify their author and may be attested;
- The Platform shall support conformance to legislation, national and international mandates and directives on the protection of personal health data; in cases where these differ between countries the platform shall support country-specific policies to be implemented;
- The Platform shall comply with legislation pertaining to subject access requests, if personally identifiable data are held or processed by Platform components and services;
- The Platform shall enable authorized persons to retrieve all of the personally identifiable health information relating to any particular data subject that is included within any Platform-governed repository, including who created, , used, changed, disclosed or destroyed it, as permitted by governance policies that apply to the information; and
- The Platform shall be able to represent changes in the underlying Clinical Data Warehouses made by subjects of care or their legal representatives to correct errors or append notes (as determined by relevant legislation), including amendments to the consents applicable to the data.

3.1.5 Patient informed consent

The Platform shall support obtaining, recording and tracking the status of consent to access the whole or specified sections of every EHR, for defined EHR4CR purposes.

The Platform shall be able to indicate the EHR4CR-relevant purposes for which consent is granted. A special attention shall be given to the possible future use of data stored, as a narrow, project specific declaration of consent may hinder future use while on the other hand; a very broad declaration of informed consent by the patient may not fulfill the legal requirements for usage at all. This issue is to be studied in year 2 of the project.

The Platform should be able to generate a list of all patient consents and/or authorizations held for particular subjects, EHR systems or countries. Throughout the project, a general informed consent framework document shall be developed (year 2).

In the context of the Protocol Feasibility Scenario, no personal information is leaving the source systems but aggregated data only. As such, although explicit consent is advised if feasible, obtaining and tracking consent is not required. This statement however only holds under the provision that the aggregated information is truly non-identifying, i.e. sufficient safeguards are necessary to thwart statistical analysis attacks leading to re-identification.[3]

On July 13, 2011, the Article 29 Data Protection Working Party (hereafter Article 29 WP) adopted Opinion 15/2011 on the Definition of Consent.[4] Article 29 WP: "This means that there must be no risk of deception, intimidation or significant negative consequences for the data subject if he/she does not consent. Data processing operations in the employment environment where there is an element of subordination, as well as in the context of government services such as health may require careful assessment of whether individuals are free to consent." The working group examines various examples of informed consent and its implications.

The first example is about the creation of a summary record which is absolutely voluntary, and the patient will still receive treatment whether or not he or she has consented to the creation of a summary record. In this case consent for the creation of the summary record is freely given because the patient will suffer no disadvantage if consent is not given or is withheld.

In the second example, a moderate financial incentive to choose the e-health record instead of the paper file is proposed. Patients refusing the e-health record would not suffer disadvantage in the sense that the costs do not change for them. Thus, it could be considered here as well that they are free to consent or not to the new system.

In the third example, patients refusing the e-health system have to pay a substantial extra cost compared to the previous tariff system and the processing of their file is considerably delayed. This signifies a clear disadvantage for those not consenting, with the purpose to bring all citizens within the e-health system in a scheduled deadline. Consent is therefore not sufficiently free.

[3] D3.1 Initial EHR4CR architecture and interoperability framework specifications, Section 9.

[4] http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp187_en.pdf.

In regard to EHR4CR, it can be seen as falling within example one, as there is no cost involved for the patient, he is free to consent and does not suffer any disadvantage if not doing so.

The definition of sensitive personal data is very broad, being defined as personal data that reveal "racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life" and processing of such data is in principle prohibited, with a very limited list of exceptions (Article 8.2(a) of Directive 95/46/EC). Article 8.2(a) Directive 95/46/EC requires explicit consent to process sensitive data.

EXPLICIT consent:

In legal terms "explicit consent" is understood as having the same meaning as express consent. The difference here is that, whereas with regular personal data, for consent to be valid it must be unambiguous (and explicit/express consent is one among many ways to show unambiguous consent), in case of sensitive personal data, explicit/express consent is the ONLY valid way to show valid consent. The Article 29 WP defines this as follows: "meaning an active response, oral or in writing, whereby the individual expresses his/her wish to have his/her data processed for certain purposes. Therefore, express consent cannot be obtained by the presence of a pre-ticked box. The data subject must take some positive action to signify consent and must be free not to consent." An example is given: A patient who is informed by a clinic that his medical file will be transferred to a researcher unless he objects (by calling a number), will not meet the requirement of explicit consent. It clearly constitutes a lack of full information about the extent of data to be transferred. Also: "Consent does not have to be recordable to be valid. However, it is in the interest of the data controller to retain evidence."

4. Conclusion

WP 5 aims to outline all the constraints defined regarding security. Services must provide a useful and efficient tool for the first step of a clinical research, the feasibility study and be compliant with all legal and ethical requirements. At month M12, some of the constraints are still not finalised, this led to the WP 5 team making some assumptions about the design of the platform, maintaining a bird's eye view towards the architecture in Scenario one and two. However, the core legal framework is well identified and defined. It will be described and applied to EHR4CR in deliverable 5.2 at length.