

e-CRL: A Rule-based Language for Expressing Patient Electronic Consent

Cédric Pruski

CR SANTEC

Centre de Recherche Public - Henri Tudor

2A, rue Kalchesbrück L-1852, Luxembourg

Email: cedric.pruski@tudor.lu

Abstract—Since the advent of the Web, the health domain is progressively adopting emerging technologies what forces new paradigms, concepts and tools to be defined. Electronic consent is one of them. This recently defined notion aims at formalising electronically the agreement of the patient on sharing personal health information. However, existing approaches dealing with electronic consent do not provide the adequate concepts to express, in an unambiguous manner, patients' wishes with respect to the access and management of their personal health data. To correct this lack, we propose the e-CRL language. This language has been designed in order to facilitate the capture and to formalise the expression of patients' consent regarding the access and management of their health information. In this paper, we first discuss the objective such language must fulfil, we then introduce the syntax and the semantics of the e-CRL language and we eventually give some examples of e-CRL rules.

Keywords-electronic consent; data privacy; rule language; eHealth; legal aspects;

I. INTRODUCTION

In the early nineties, the explosion of the World Wide Web and its ever increasing popularity has deeply modified our contemporary society and our way of life. The health domain is part of this revolution and is progressively adopting emerging technologies. However, it has been pointed out in [1] that one of the main obstacles to this migration is doctors', or more generally health professionals, and patients' mentality and behaviour. Since health is a critical subject for most of the people, they will have to trust new technologies in order to change their mind and finally adopt it what will speed up the technological migration.

This is all the more so true when patients must give their consent (or agreement) for accessing their personal health data in an electronic way. In addition to security issues that are extremely important, concepts and tools have to be defined in order to facilitate patient's data entry. Actually, an intuitive interface has to be proposed to the patient in order to capture its consent. But what is more important is the way the entered data, that represent patient's consent regarding the access and management of its health data, is formalised. This will enable the system to manage in a rigorous and controlled way the access to and interaction with sensitive data. Actually, Coiera and Clarke have identified four distinct levels of consent [2]:

- 1) *General consent* - The patient consents to give a full access to his health data.
- 2) *General consent with specific exclusions* - In this case the patient gives a general agreement but some restrictions are defined. The restrictions are defined based upon specific elements like the persons, the data, and/or a particular purpose (i.e., location, date ...).
- 3) *General denial with specific consents* - This situation is complementary of the previous one but the priority is given to the restrictions. However, some exceptions permitting identified persons to interact with protected data can be defined.
- 4) *General denial* - This case is the opposite of that described at level 1.

If level 1 and 4 are easy to manage, from a technical point of view, the two others (i.e., levels 2 and 3) are more problematic. Actually, patients have to express, in a non ambiguous way, what persons will be allowed to access their data (or a subset of it) in a rigorously controlled manner. This aspect is often neglected in existing approaches dealing with e-consent [2][3][4] where natural language like English or French is, in most of the cases, used to express patient's consent. However, since natural language is ambiguous by nature, semantic problems stand out when the consent is interpreted which, in turn, leads to security issues regarding confidential data access and management.

In order to overcome this lack, we propose the e-CRL language. It has been proposed for e-consent system designers and developers in order to facilitate the capture as well as to formalise the expression of patient's consent regarding the access and the management of his personal health information. The language has a well defined BNF-based syntax and semantics defined based on first-order logic and set theory which allow eHealth systems to fully control the access to critical health data.

The remainder of the paper is structured as follows: Section II surveys state-of-the-art in the field e-consent. Section III introduces the motivations and the various requirements for the definition of the e-CRL language. Section IV presents the e-CRL language. We first describe the syntax, then the semantics of the language and we

eventually propose examples of e-CRL rules. Finally, Section V wraps up with concluding remarks and outlines future work.

II. RELATED WORK

Although the formalisation of the expression of patient's consent is a really important problem in eHealth, it has not been so deeply studied by the scientific community. Actually, existing work dealing with e-consent are rather focusing on security aspects, as shown in [5][6][7], than on semantic issues. In this study, we focus on how patient consent is formalised in existing approaches.

In the work presented in [4], Song et al. introduce the concept of e-consent object. This object should contain all the elements concerning patient consent. Hence, it includes the so-called *rules of consent* which express the wishes of a patient regarding data access. Directly inspired from Clarke's [8] model, the authors proposed eight features for constructing rules (see section III). However, even if the proposition is interesting, it lacks semantics. Actually, the rules are not expressed in a formal language; therefore remain ambiguous at interpretation time. Observe that the idea of using an object for modelling patient content has also been exploited in [9] and [10].

In [11], Win et al. describe the "e-Medical Book" e-consent mechanism. The system provides an interface through which patient can express their consent. However, since this solution is hard coded, it provides very little flexibility for the patient. As eHealth systems are evolving very fast, with respect to legal and technological evolutions, the interface may change frequently which can disturb users somehow. Furthermore, the entered data are not formalised which can prevent its future usage. Such system has also been proposed by Win and Fulcher [12] and has the same drawbacks.

Coiera and Clarke [2] argue that the form of the consent depends on the objectives to be met by the e-consent system. Three broad functions have been identified: The first one refer to a legal aspect. In this case the consent expression is captured in a very basic form. In the two other scenarios, the system plays a more active role: gatekeeper. Hence, in this case, the consent record must be more detailed in order to be used by the system since it will grant permissions to users to access confidential data.

O'Keefe et al. have presented another model of e-consent object [9]. As it is already the case in [4], the e-consent object proposed by O'Keefe et al. also contains statements for expressing patient wishes as regard data access. The statements are expressed in natural language or natural language inspired one. As evoked in the introduction of this document, such statements are problematic when they are interpreted due to natural language ambiguity.

As illustrated by this non exhaustive survey, the

major approaches dealing with e-consent use an object that contains data of interest and rules expressing the consent of patient. Nevertheless, to our knowledge, there is no formal language that can be used to construct such rules. In consequence, they cannot be interpreted in a non ambiguous way by systems which, in turn, can cause serious security breaches.

III. MOTIVATIONS AND REQUIREMENTS

As illustrated in the previous section, existing approaches do not provide the appropriate concepts for patients to express their agreement regarding the access and management of their personal health information. Nevertheless, some propositions have been made and deserve to be considered in order to overcome these lacks. In [4], Song et al. have identified 8 features directly linked to the expression of patient wishes. Some of them are relevant, the others can be considered as optional. Actually, combining technological, legal and cultural approaches of e-consent several aspects need to be taken into account.

First, systems dealing with patient consent need to identify the various persons who want to access to health data. Therefore, the patient will have to specify, in a non ambiguous way, **who** can access to its data. In [13], the author pretends that considering a hierarchy of role that can be assigned to users is sufficient for managing data access. The so-defined hierarchy can be seen as nested sets where the set that included all the others represents the more general role. Although this notion is important, because it can reduce the complexity of the expressed rules, considering roles only is not enough. We believe that patients must be able to define exception to the roles for example: If a patient does not trust in a particular doctor anymore, he should be able to specify this particular doctor and modify his access rights.

Second, the system needs to decide **what kind of consent** is given to the user requesting an access to the data. In our approach, only two types of consent are considered: *denial* or *agreement*. If the *denial* case is easy to manage from a technical point of view insofar as the data access is prohibited, the second case needs to be discussed because when a patient allows someone to access his data, additional aspects come into play. In fact, the system will have to control what actions can be done on the data which must be implicitly or explicitly defined in the access right. To this end, we have identified a set of 9 actions. Data can either be read only or modified. The latter case can be derived into: adding data, erasing data, deleting the whole record, moving data (i.e., copy data), deleting the whole data, and all privileges. As in most of cases patients are owner of their personal health information, they must have at their disposal a way to express such privileges (i.e., who can access data in which way).

Third, the **kind** and the **sensitivity** of the data that can be accessed and handled are of utmost importance. In order

$WHAT(Type);$
 $PERM(Rights);$
 $DATA(A);$
 $START(Date);$
 $DURATION(Period);$
 $WHY(String).$

$A := All$
 $|expr$

$expr := String$
 $|expr \& expr$
 $|expr | expr$
 $|expr - expr$

$Type := agreement$
 $|denial$

$Rights := Read | Modify | none$
 $Modify := 'Write'$
 $| 'Copy'$
 $| 'Write, Copy'$
 $| 'Write, Delete'$
 $| 'Write, Copy, Delete'$
 $| 'Delete'$
 $| 'DeleteAll'$
 $| 'All'$

$Date := Number / Number / Number$

$Period := Number$

$Char := 'a'..'z'$
 $| 'A'..'Z'$

$String := Char | Number | \epsilon$

$Number := \mathbb{N}$

We try to keep the syntax of the e-CRL language as intuitive as possible in order to facilitate patients work. In consequence, an e-CRL rule is made up of either a **DELEGATE** statement that takes as argument the name of the person concerned with the delegation task or a more complex expression. In such expression, the *WHO* primitive refers to the users that have requested patient's consent. *WHAT* refers to the kind of consent that is given (agreement or denial). *PERM* represents the types of permissions that are assigned to the person(s) mentioned as arguments in the *WHO* statement as discussed in the previous section. *DATA* denotes the targeted data (i.e., the set of data that are accessible). *START* contains the date when the consent starts. Observe that *DATE* can contain a date that is not well formed (e.g., a date like 31/02/2004). The validation of the

information provided by the submitted rules will be done at interpretation time (see section IV-B for more details). *DURATION* contains a natural number representing the length (in days) of the period of validity. Last, the *WHY* constructor refers to the purpose of the consent thus, a string of characters. Examples of e-CRL rules are given section IV-C. Observe that the *WHO*, *DATA* and *WHY* constructors have alphanumerical strings as parameters.

Since we provide a Backus Naur Form (BNF)-based syntax of the e-CRL language, it is now much easier for software developers to build tools, like parsers or syntactic analysers, for treating e-CRL rules. However, the semantics of the language need to be defined in order to fully implement e-CRL in existing applications dealing with patient consent.

B. Semantics of e-CRL

In this section we introduce the semantics of the e-CRL language. Since the identified requirements are easily formalised using sets, especially the ones concerning persons who need to access the data and data itself, and, moreover, the rules build using the e-CRL language can be compared to axioms, we decided to use first-order logic and set theory in order to give the semantics of the language. Before proceeding, we need to introduce some notions. In the following axioms that define the semantics of the e-CRL language, Σ denotes all strings of characters build upon any alphabet (including words of the natural language), Ω_{hp} denotes the set of identified healthcare professionals, Ω_{pa} is the set of existing individuals and, Ω_d the set of identified data. In consequence,

$\forall i \in \Omega_{pa} \quad DELEGATE_TO(i)$
 $\forall i \in \Omega_{pa} \quad DELEGATE_WITH(i)$
 $\forall i \in \Omega_{pa} \quad DELEGATE_REMOVE(i)$

$WHO(p) \Rightarrow \exists p \in \Omega_{hp} \oplus \exists p \subseteq \Omega_{hp}$
 $WHO(p \& q) \Rightarrow (\exists p \in \Omega_{hp} \vee \exists p \subseteq \Omega_{hp}) \wedge (\exists q \subseteq \Omega_{hp} \vee \exists q \in \Omega_{hp})$
 $WHO(p | q) \Rightarrow (\exists p \in \Omega_{hp} \vee \exists p \subseteq \Omega_{hp}) \vee (\exists q \subseteq \Omega_{hp} \vee \exists q \in \Omega_{hp})$
 $WHO(p - q) \Rightarrow (\exists p \in \Omega_{hp} \setminus q) \vee (\exists p \subseteq \Omega_{hp} \setminus q)$
 $\forall p \in \Omega_{hp} \quad WHO(all)$

$WHAT(x) \Rightarrow x \in \{ 'agreement', 'denial' \} \subset \Sigma$

$PERM(x) \Rightarrow x = \{ 'Read' \} \in \Sigma \vee x = \{ 'Write' \} \in \Sigma \vee x = \{ 'Copy' \} \in \Sigma \vee x = \{ 'Write, Copy' \} \in \Sigma \vee x = \{ 'Write, Delete' \} \in \Sigma \vee x = \{ 'Write, Copy, Delete' \} \in \Sigma \vee x = \{ 'Delete' \} \in \Sigma \vee x = \{ 'DeleteAll' \} \in \Sigma \vee x = \{ 'All' \} \in \Sigma \vee$

$DATA(d) \Rightarrow \exists d \subseteq \Omega_d$

$DATA(d|e) \Rightarrow (\exists d \subseteq \Omega_d) \wedge (\exists e \subseteq \Omega_d)$
 $DATA(d\&e) \Rightarrow (\exists d \subseteq \Omega_d) \vee (\exists e \subseteq \Omega_d)$
 $DATA(d - e) \Rightarrow \exists d \subseteq \Omega_d \setminus e$
 $\forall d \in \Omega_d \quad DATA(all)$

$START(d/m/y) \Rightarrow (d \in \mathbb{N} \wedge 1 \leq d \leq 31) \wedge (m \in \mathbb{N} \wedge 1 \leq m \leq 12) \wedge (y \in \mathbb{N})$

$\forall d \in \mathbb{N}^* \quad DURATION(d)$

$\forall b \subset \Sigma \quad WHY(b)$

From a more pragmatic point of view, the so defined axioms can be understood in the following way. DELEGATE_TO, DELEGATE_WITH, DELEGATE_REMOVE, WHO, WHAT, PERMS, DATA, START, DURATION and WHY which are the primitives of the languages are considered as relationships. DELEGATE_TO takes an alphanumeric string of characters as inputs therefore its associated relationship is verified if the argument is an alphanumeric string.

Regarding the WHO constructor, arguments can be either an element (an individual) or a set (a group of individuals) or a mathematic expression combining sets. Hence, the WHO relationship is verified if the individuals or group given in argument are declared in the system.

The WHAT relationship is verified if the argument is either 'agreement' or 'denial'. The PERM constructor is defined on the same idea.

The DATA constructor takes as argument a label denoting the set of concerned data. This set can be defined using dedicated operators which respect the semantics of the basic operations on sets (union, intersection, difference).

The START primitive takes as argument a date on the form dd/mm/yyyy where dd and mm have to be smaller than 31 and 12 respectively while yyyy must be a natural number. This allows a better control of the entered date but some inconsistencies remain and will be corrected at interpretation time.

The DURATION constructor is verified if its argument is a natural number. This integer denotes the length (in days) of the validity of the consent.

Lastly, the WHY primitive, indicating the purpose of the consent, is verified if the parameter is a string of alphanumeric characters.

Actually, to be rigorous, we needed to define the appropriate logic structure (including domains, relations, functions and distinguished elements) as well as a function that maps every element of the language to the corresponding elements of the domain to define properly the previous constructors and elements of the language. This logic structure also allows e-CRL rules to be verified. In that sense, since e-CRL has an axiomatic semantics,

e-CRL rules can be compared to first-order logic formulae that can easily be verified on the evoked logic structure. Once established, the consistency of the set of e-CRL rules can be checked. This can be done for instance through the use of any solvers. The rules can then be interpreted by the system in charge of managing data access and management when someone wants to access medical data. This point will be explained more in details in the following section using concrete examples.

C. Examples of e-CRL Rules

As e-CRL allows users to construct rules, in this section, examples of e-CRL rules are provided. The proposed examples want to be realistic (i.e., the proposed rules can be submitted by patients).

First, assume that the patient is not able to take decision regarding the access of his health information. In that case, he will have to delegate his rights from the 5th of October 2009. The appropriate e-CRL rule will consist in using the DELEGATE_TO primitive with the identifier of the person that will be in charge of managing the patient consent as parameter with the corresponding date as argument of the START primitive. If for instance the identifier of the person is 124424345 the corresponding e-CRL rule is:

```

DELEGATE_TO('124424345');
START(05/10/2009).

```

At interpretation time, the system will replace the existing identifier by the new one given in argument at the specified date. In consequence, the patient who has delegated his rights will not be able to manage his data and add new e-CRL rules anymore until having removed his delegation.

In this second example, the patient wants to give to all radiologists and GPs an access to his radiological data during a 30 days long period. Moreover, this access also includes the permission for the targeted person to modify and copy his data. To do so, the patient will have to write the following e-CRL rule:

```

WHO('radiologist' & 'GP');
WHAT('agreement');
PERM('Read,Write,Copy');
DATA('radiological');
START(05/10/2009);
DURATION(30);
WHY('medical control').

```

After having identified the persons who want to access radiological data, the system, according to this rule, will check if the person is either a GP or a radiologist. If so the system will allow him to read, write and copy any radiological data from October the 5th 2009 and for a period of 30 days. So if for instance a radiologist wants to access the data on the 12th of February 2010, the system

will refuse this action.

In this last example a patient wants to deny access to its data to every GP except doctor Smith. The corresponding rules will be:

```
WHO('GP'-'Smith');
WHAT('denial');
PERM('Read,Write,Copy,Delete');
DATA('all');
START(05/10/2009);
DURATION(3000);
WHY('trust reason').
```

The interpretation of this rule will be processed as follows. First, the system will identify the user wanting to access the data. If the user is a GP, the system will refuse any actions specified in the PERM primitive to all data contained in the system from October the 5th 2009 and for a period of 3000 days. Now if the user is Smith, the system will allow him to read, write, copy and delete all the data for the specified period starting from the specified date. Furthermore, the identified user may belong to any of the specified groups or users. In that case, for security reasons, the system will refuse all access to all the stored data.

The previous examples and their interpretation by the system allow the reader to better understand the philosophy of the e-CRL language which provides the basic elements for expressing patient consent. However, due to legal, technological or cultural evolutions, the patient consent notion is supposed to evolve over time. In that case, new features can easily be added to e-CRL which will enrich the expressivity of the language.

V. CONCLUSION AND FUTURE WORK

Patient consent is the cornerstone of applications dealing with eHealth and therefore need to be addressed with care. In this paper we have introduced the e-CRL rule language for expressing patient consent with respect to the access and the management of critical personal health data. Our work has consisted first in identifying a set of requirements laying down the foundation for the construction of such a language and, in a second time, we have designed the e-CRL language including the definition of its syntax and semantics. To this end, we have discussed the set of requirements taking into account business needs as well as technological and legal issues. Moreover, we have provided a BNF-based syntax and a first-order logic and set theory based semantics for the e-CRL language. As a result, patient consent can now be expressed in a formal way and can be interpreted unambiguously. Nevertheless, we need now to provide additional elements to fully benefit from the e-CRL language. In consequence, in our future work, we will propose an intuitive interface for capturing patient consent and we will also provide an inference engine able first to

test the consistency of the set of rules and second to manage the access and interaction for users.

REFERENCES

- [1] R. Whiddett, I. Hunter, J. Engelbrecht, and J. Handy, "Patients attitudes towards sharing their health information," *International Journal of Medical Informatics*, vol. 75, no. 7, pp. 530–541, 2006.
- [2] E. Coiera and R. Clarke, "e-consent: the design and implementation of consumer consent mechanisms in an electronic environment," *Journal of the American Medical Informatics Association*, vol. 11, no. 2, pp. 129–140, 2004.
- [3] C. O'Keefe, A. Goodchild, P. Greenfield, A. Waugh, E. Cheung, and D. Austin, "Implementation of electronic consent mechanisms," Final Analysis Paper, August 2002.
- [4] H. Song, T. K. Win, and P. Croll, "Patient e-consent mechanism: Models and technologies," in *Proceedings of COLLECTeR*, Melbourne, Australia, 2002.
- [5] J. F. Reid, I. Cheong, M. P. Henricksen, and J. Smith, "A novel use of RBAC to protect privacy in distributed health care information systems," in *8th Australasian Conference on Information Security and Privacy (ACISP 2003)*, July 9–11 2003.
- [6] X. Chen, D. Berry, and W. Grimson, "Identity management to support access control in e-health systems," in *4th European Conference of the International Federation for Medical and Biological Engineering ECIFMBE 2008*. Springer Berlin Heidelberg, November 2008, pp. 880–886.
- [7] B. Blobel, "Authorization and access control for electronic health record systems," *International Journal of Medical Informatics*, vol. 73, pp. 251–257, 2004.
- [8] R. Clarke, "e-consent: A critical element of trust in e-business," in *Proc. 15th Bled Electronic Commerce Conference*, Bled, Slovenia, 17–19 June 2002. [Online]. Available: <http://www.rogerclarke.com/EC/eConsent.html> (last access November 2009)
- [9] C. O'Keefe, P. Greenfield, and A. Goodchild, "A decentralised approach to electronic consent and health information access control," *Journal of Research and Practice in Information Technology*, vol. 37, no. 2, pp. 161–178, 2005.
- [10] J. Bergmann, O. J. Botta, D. P. Pretschner, and R. Haux, "An e-consent-based shared EHR system architecture for integrated healthcare networks," *International Journal of Medical Informatics*, vol. 76, no. 2–3, pp. 130–136, 2007.
- [11] K. Win, H. Song, P. Croll, and J. Cooper, "Implementing patient's consent in electronic health record systems," in *Proceedings of COLLECTeR*, Melbourne, Australia, 2002.
- [12] K. T. Win and J. A. Fulcher, "Consent mechanisms for electronic health record systems: A simple yet unresolved issue," *J. Med. Syst.*, vol. 31, no. 2, pp. 91–96, 2007.
- [13] C. Ruan, "UML specification of e-consent requirements in a health care system," in *CSA '08: Proceedings of the International Symposium on Computer Science and its Applications*. IEEE Computer Society, 2008, pp. 275–280.