Original Article

Healthc Inform Res. 2014 January;20(1):39-44.
http://dx.doi.org/10.4258/hir.2014.20.1.39
pISSN 2093-3681 • eISSN 2093-369X

HIR
Healthcare Informatics Research

CCR+: Metadata Based Extended Personal Health Record Data Model Interoperable with the ASTM CCR Standard

Yu Rang Park, PhD[1], Young Jo Yoon, BS[2], Tae Hun Jang, BS[2], Hwa Jeong Seo, PhD[3], Ju Han Kim, MD, PhD[2,4]

[1]Department of Biomedical Informatics, Asan Medical Center, Seoul; [2]Seoul National University Biomedical Informatics (SNUBI), Seoul National University College of Medicine, Seoul; [3]Medical Informatics, Graduate School of Public Health, Gachon University of Medicine and Science, Incheon; [4]Division of Biomedical Informatics, Seoul National University College of Medicine, Seoul, Korea

Objectives: Extension of the standard model while retaining compliance with it is a challenging issue because there is currently no method for semantically or syntactically verifying an extended data model. A metadata-based extended model, named CCR+, was designed and implemented to achieve interoperability between standard and extended models. **Methods:** Furthermore, a multilayered validation method was devised to validate the standard and extended models. The American Society for Testing and Materials (ASTM) Community Care Record (CCR) standard was selected to evaluate the CCR+ model; two CCR and one CCR+ XML files were evaluated. **Results:** In total, 188 metadata were extracted from the ASTM CCR standard; these metadata are semantically interconnected and registered in the metadata registry. An extended-data-model-specific validation file was generated from these metadata. This file can be used in a smartphone application (Health Avatar CCR+) as a part of a multilayered validation. The new CCR+ model was successfully evaluated via a patient-centric exchange scenario involving multiple hospitals, with the results supporting both syntactic and semantic interoperability between the standard CCR and extended, CCR+, model. **Conclusions:** A feasible method for delivering an extended model that complies with the standard model is presented herein. There is a great need to extend static standard models such as the ASTM CCR in various domains: the methods presented here represent an important reference for achieving interoperability between standard and extended models.

Keywords: Standards, Metadata, Interoperability, Personal Health Record

Submitted: November 15, 2013
Revised: 1st, January 21, 2014; 2nd, January 27, 2014
Accepted: January 27, 2014

Corresponding Author
Ju Han Kim, MD, PhD
Seoul National University Biomedical Informatics (SNUBI), Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 110-799, Korea. Tel: +82-2-740-8320, Fax: +82-2-747-8928, E-mail: juhan@snu.ac.kr

I. Introduction

Standardization is a consensus-driven activity that is carried out by and for the interested parties themselves. It is based on openness and transparency within independent organizations, and it aims to establish the voluntary adoption of and compliance with standards [1]. A standard model is useful in the early stage of scientific domains to represent and share data [2,3]. However, when the domain is immature, the standard model may not be applicable to new needs [4-10]. For instance, the American Society for Testing and Materials (ASTM) established the Continuity of Care Record (CCR) for providing standard personal health record (PHR) contents. The CCR includes a summary of the patient's health status (e.g., problems, medications, and allergies) and basic information about insurance, advance directives, care documentation, and the patient's care plan. Subsequently, the Continuity of Care Document standard provided a specification for CCR implementation that incorporated aspects of the Health Level Seven Clinical Document Architecture. Many healthcare stakeholders have adopted and implemented these standards for the management of PHRs, dental history, genetic test data, and sensor data, and for testing Electronic Health Record (EHR) interoperability [5,6,11-14]. However, these standards provide a restricted set of PHR contents [15], which has prompted many organizations and researchers to extend the standard model; however, this has largely been done without consideration of the interoperability between standard models [5,11,14]. Although some of the research has provided validation methods for extended models [11,14], if the other PHR systems do not support those validation methods, these extended data models cannot achieve interoperability or compliance with the standard model. Another problem with extended data models is that additional attributes or contents cannot be validated semantically or syntactically. While the Health Level Seven (HL7) standards provide generic structure based on HL7 standard artifacts, such as RIM, data types and vocabularies or assigned external coding systems, the ASTM CCR provide a static model.

In this article we explore how extended data can be integrated and made to comply with the static standard model and how it can be validated both semantically and syntactically. An additional system is needed to manage the metadata from the standard and extended models; this study applied a metadata registry (MDR) and a method defined in the ISO/IEC 11179 metadata standard model [16]. The ISO/IEC 11179 standard addresses management of the semantics of data elements: it provides a standard metadata model for the representation of data elements and provides a methodology for the registration of the descriptions of data elements through this standard model to an MDR [17]. In a metadata standard, if the user wants to add additional data, he/she registers the metadata of data in the MDR in metadata format. After metadata registration, the user obtains a metadata ID from the MDR. This metadata ID may then be used as a semantic identifier in another system. In this research, we employed ISO/IEC 11179 parts 3, 4, and 6. This paper focuses on an implementation of processes needed to construct extended data model in compliance with static standard model and a demonstration of achieving interoperability using extended PHR data in patient-centric multicenter data exchange.

II. Methods

The feasibility of interoperability between metadata-based extended data and a standard model was demonstrated by designing and implementing four steps as described below.

First, metadata were extracted from the ASTM CCR, which is an XML-based data model. This metadata extraction process employed ISO/IEC 11179 part 5: naming and identification principles. Each XML tag was defined as metadata. In this process, semantic relationships between metadata were added, such as dependency or composite relationships. These relationships were defined in our previous research on representing semantic and syntactic relationships among metadata [8]. The extracted metadata were registered into an ISO/IEC-11179-based MDR, the Clinico-Histopathological Metadata Registry (CHMR). This MDR was build based on ISO/IEC 11179 parts 3 (registry metamodel and basic attributes) and 6 (registration).

Second, indiscriminate extension for incorporating of patient- or organization-specific data into the standard model was avoided by applying the metadata registration and management method defined in ISO/IEC 11179. The CCR data model was examined to identify where metadata would be included and the various attributes that would need to be added to the CCR data model in order to represent these metadata appropriately. This process resulted in the following four attributes being extended in all sections of the CCR containing metadata ID, question labels, and response values: *Question, MetadataID, QuestionLabel,* and *Value*. These attributes were used as semantic identifiers of the metadata. Detailed definitions of the extended attributes are provided in the Results. Based on this extended data model, an XSD (XML schema definition) file was defined for validating extended data. These attributes were defined as metadata and registered in the CHMR, and the metadata were interconnected with other metadata from the CCR data model using dependency or composite relationships.

Third, a validation method was devised to validate not only the metadata-based extended XML file, but also the standard-based XML file. The following factors were also considered throughout the development of the validation process: 1) the extended CCR+ attributes should be ignored when validating the standard XML file, 2) the extended CCR+ attributes should be validated, 3) semantic annotation of data such as concepts from controlled vocabulary should be validated, 4) the validation process should be implemented

http://dx.doi.org/10.4258/hir.2014.20.1.39

in PHR systems without requiring user interaction, and 5) the successfully validated XML file should be loaded into the PHR systems.

Fourth, the CCR+ data model was evaluated to determine whether it can exchange data and be integrated with the standard CCR data model. The processes of this evaluation consisted of loading CCR and CCR+ XML files into a Health Avatar CCR+ and observing whether the data from each XML file were seamlessly selected and visualized without additional modification. Two CCRs from Ajou University Medical Center and Pusan National University Hospital, and one CCR+ from Gachon University Gil Medical Center were used in this evaluation.

III. Results

1. CCR Metadata Registration

Before metadata extraction, the CCR comprised three core components: *Header, Body*, and *Footer*. The CCR *Header* defines all information relevant to the document, such as its unique identifier, language, version, and information about the patient of the CCR. The CCR *Body* contains patient-specific health data in 17 sections (*Payer, AdvanceDirective, Support, FunctionalStatus, Problems, FamilyHistory, SocialHistory, Alerts, Medications, MedicalEquipment, Immunization, VitalSigns, Results, Procedures, Encounters, PlanOfCare*, and *HealthCareProviders*). The CCR *Footer* describes the actors in the care context, and comments and references to external document and information [18].

In total, 188 metadata were extracted from the CCR data model: 14, 164, and 10 from the *Header, Body*, and *Footer*, respectively. These metadata were hierarchically and semantically interconnected using composite or dependency relationships. Among the 188 metadata, 273 composite or dependency relationships were defined. The composite relationship means that the metadata are included at a semantically high level; for example, the CCR *Body* contains 17 sections, and so the *Body* metadata have 17 composite relationships between 17 sections. Some clinical data elements are supposed to be activated or deactivated by the response (value) to a different data element; the association between those data elements is defined as a dependency relationship. The *DateTime* type is determined by a *DateTimeType* value; thus, a dependent relationship is defined between *DateTime* and *DateTimeType*.

2. Extension of the CCR in a Metadata-Standard Manner

The CCR data model was extended in a metadata-standard manner to avoid indiscriminate extension for incorporating of patient- or organization-specific data into the standard model. The following four attributes were added to all sections of the CCR: *Question, MetadataID, QuestionLabel*, and *Value*. *Question* is a parent attribute of *MetadataID, QuestionLabel*, and *Value*, and is a container for representing the metadata identifier, question label, and response value of the patient; it is defined for each metadata item. The *MetadataID* attribute is required if the *Question* attribute is included. This attribute is a semantic identifier that connects to the MDR.

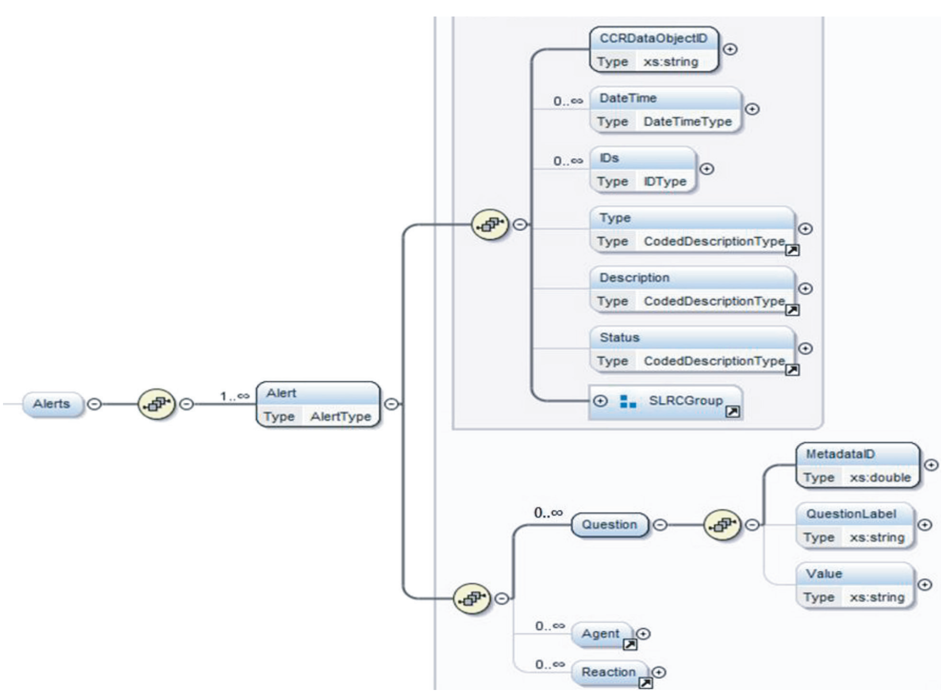

Figure 1. A part of screenshot of *Alert* section in CCR+ model. CCR: Community Care Record.

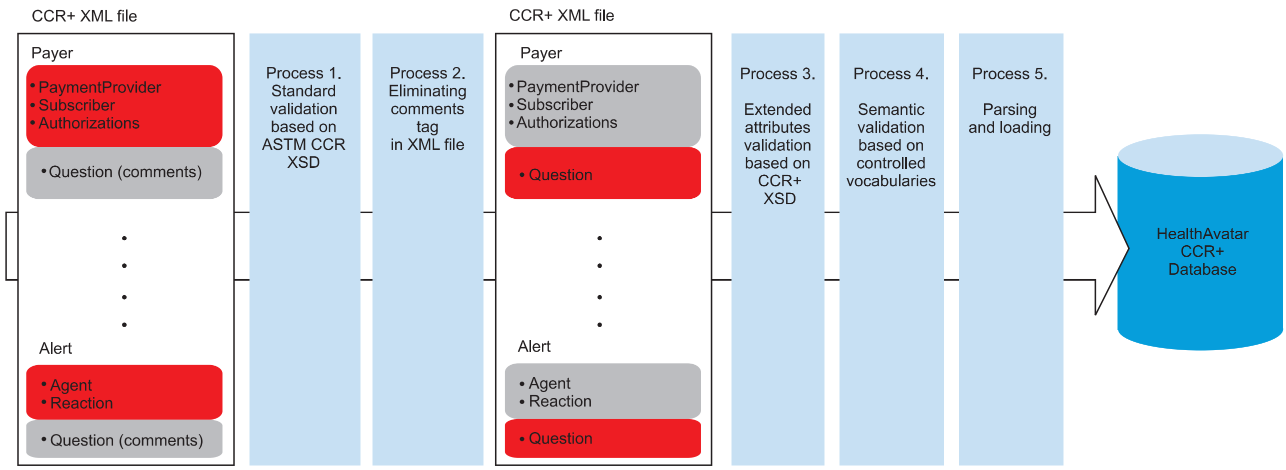

Figure 2. Architecture of multi-layered validation. CCR: Community Care Record, XML: extensible markup language, ASTM: American Society for Testing and Materials.

The *QuestionLabel* and *Value* tags are optional. *QuestionLabel* is a text that briefly describes the question or data, while *Value* contains the patient response value for the question. An XSD file was also designed based on the extended CCR data model. Figure 1 shows a partial screenshot of the *Alerts* section in the XML validation file. This XSD file is used in a multilayered validation process.

3. CCR+ Validation Processes

A multilayered validation process was devised and implemented for validating both the CCR+ XML and CCR XML files. Multilayered validation consists of following five subprocesses: 1) standard validation based on the ASTM CCR XSD, 2) eliminating the comments tag in the XML file, 3) extended tag validation based on the CCR+ XSD, 4) semantic validation based on controlled vocabularies, and 5) parsing and loading. Figure 2 shows the structure of this multilayered validation process.

The input file for the multilayered validation is the CCR XML or CCR+ XML file. First, the standard validation process checks the XML file based on the ASTM CCR XSD. The extended four attributes use the comments tag to be ignored during this process. The XML Comments start with “<!–” and end with “-->”. Supplementary Figure 1 shows an example of the input CCR+ XML file with the comments tag. Second, the comments process is eliminated by removing the comments tag in the XML file for the following process. The CCR XML file just passes through this process. Third, the extended attribute validation checks the XML file based on the CCR+ XSD. This process involves checking the extended attributes. Fourth, semantic validation involves checking the concept of a controlled vocabulary, which is performed using only the International Classification of Diseases (ICD)-9 or ICD-10 codes. This process is both time-consuming and optional. Finally, after successful validation, the XML file is parsed and loaded into the Health Avatar CCR+ database.

This multilayered validation processed was additionally developed for a smartphone application, named Health Avatar CCR+, which was itself developed in our previous work as a patient-centered health data management system [19].

4. Evaluation

The CCR+ data model was evaluated using two CCR XML files from Ajou University Medical Center and Pusan National University Hospital, and one CCR+ XML file from Gachon University Gil Medical Center. Thirty-one questions were added to the CCR+ XML file from Gachon University Gil Medical Center, with 3, 5, 2, and 21 questions being added to the *Payer*, *Problems*, *Results*, and *Procedures* sections, respectively. Supplementary Figure 2 shows the evaluation results. Before the evaluation test, two CCR XML files that had already been parsed were uploaded to the Health Avatar CCR+ database. Supplementary Figure 2A and B show the parsing and loading results for the CCR+ XML file uploaded to the Health Avatar CCR+ for each section. In total, 1, 1, 10, 3, 18, 6, 8, and 6 records were added to the *Payer, Alerts, FunctionalStatus, SocialHistory, Results, VitalSigns, Medications*, and *Encounters* sections, respectively. Supplementary Figure 2C shows the details of the uploaded data in the *Payer* section, to which the following two questions were added: *InsuranceName* and *InsuranceCompanyName* (Supplementary Figure 2). The ASTM CCR already provides architecture to record insurance and company names; however, these data were recorded in the *Source* XML tag without definite attributes. This situation means that natural language processing would need to be applied to extract the insurance

http://dx.doi.org/10.4258/hir.2014.20.1.39

and company name in the CCR XML file, and so the extraction efficiency was improved by adding the two aforementioned questions. To retain compliance with the standard model, the values of these two questions were also recorded in the *Source* XML tag.

IV. Discussion

This study implemented a metadata-based extended data model that was demonstrated to comply with the standard model. This metadata-based interoperability method has already been used for implementing a EHR system, named DiaNet, and a personal big-data management platform, named Health Avatar platform. The present study was subject to several limitations. The focus was to determine the feasibility of the CCR+ data model, and we evaluated a simple scenario involving patient-centric data exchange. Future work should involve designing and evaluating various scenarios to enable the exchange of clinical data or documents in diverse environments. Further research is also needed to extend our approach to other static models in order to explore its generalizability. In addition, local implementation of the extended data model should be examined. The extension of static standard models, such as the ASTM CCR, is required in various types of research [5,6,10,11,14]; the methods described herein represent an important reference toward achieving interoperability between standard and extended models.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

This work was supported by the National Research Foundation of Korea grant funded by the Korea government (2010-0028631). YRP's educational training was supported in part by a grant from the Korea Health 21 R&D Project, Ministry of Health, Welfare, and Family Affairs, Korea (A112020).

Supplementary Materials

Supplementary materials can be found via http://pdf.medrang.co.kr/Hir/2014/020/Hir020-01-06-s001.pdf. Figure 1. A part of example of CCR+ XML file. Figure 2. Screenshots of Health Avatar CCR+.

References

1. Brusse B, Wenning R. Benefits of standards and standardization [Internet]. Brussels, Belgium: COPRAS; c2010 [cited at 2013 Nov 13]. Available from: http://www.w3.org/2004/copras/docu/D15.html#benefits.
2. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet 2001;29(4):365-71.
3. HL7: an application protocol for electronic data exchange in healthcare environments (HL standard ver. 2.3.1.). Ann Arbor (MI): Health Level Seven; 1999.
4. Smith B, Ceusters W. HL7 RIM: an incoherent standard. Stud Health Technol Inform 2006;124:133-8.
5. Jing X, Kay S, Marley T, Hardiker NR, Cimino JJ. Incorporating personalized gene sequence variants, molecular genetics knowledge, and health knowledge into an EHR prototype based on the Continuity of Care Record standard. J Biomed Inform 2012;45(1):82-92.
6. Spyropoulos B, Botsivaly M, Tzavaras A, Koutsourakis K. Extending the use of DRGs to estimate mean Home-Care cost by employing an adapted ASTM E2369-05 Continuity of Care Record. Conf Proc IEEE Eng Med Biol Soc 2006;1:5210-3.
7. Park YR, Lee HW, Cho SB, Kim JH. Cancer genomics object model: an object model for multiple functional genomics data for cancer research. Stud Health Technol Inform 2007;129(Pt 2):1235-9.
8. Park YR, Yoon YJ, Kim HH, Kim JH. Establishing semantic interoperability of biomedical metadata registries using extended semantic relationships. Stud Health Technol Inform 2013;192:618-21.
9. Park YR, Kim JH. Achieving interoperability for metadata registries using comparative object modeling. Stud Health Technol Inform 2010;160(Pt 2):1136-9.
10. Jones A, Hunt E, Wastling JM, Pizarro A, Stoeckert CJ Jr. An object model and database for functional genomics. Bioinformatics 2004;20(10):1583-90.
11. Wolf KH, Schirmer S, Marschollek M, Haux R. Representing sensor data using the HL7 CDA personal healthcare monitoring report draft. Stud Health Technol Inform 2009;150:480-4.
12. Veseli H, Kopanitsa G, Demski H. Standardized EHR interoperability: preliminary results of a German pilot project using the archetype methodology. Stud Health Technol Inform 2012;180:646-50.
13. Do NV, Barnhill R, Heermann-Do KA, Salzman KL, Gimbel RW. The military health system's personal health

record pilot with Microsoft HealthVault and Google Health. J Am Med Inform Assoc 2011;18(2):118-24.
14. American Dental Association. Dental extension to the ASTM Continuity of Care Record. Chicago (IL): American Dental Association; 2008. (ANSI/ADA Standard No. 1040).
15. ASTM International. Standard specification for Continuity Of Care Record (CCR). West Conshohocken (PA): ASTM International; 2005. (ASTM E2369).
16. International Organization for Standardization. Information technology: metadata registries (MDR). Part 3. Registry metamodel and basic attributes. Geneva: International Organization for Standardization; 2003. (ISO/IEC 11179-3:2003).
17. Sinaci AA, Laleci Erturkmen GB. A federated semantic metadata registry framework for enabling interoperability across clinical research and care domains. J Biomed Inform 2013;46(5):784-94.
18. Blobel B. EHR architectures - comparison and trends. Stud Health Technol Inform 2008;134:59-73.
19. Park YR, Kim H, An EY, Kim HH, Kim JH, Park RW, et al. Establishing semantic interoperability in the course of clinical document exchange using international standard for metadata registry. J Korean Med Assoc 2012;55(8):729-40.