

Image-Centric Integrated Data Model of Medical Information by Diseases: Two Case Studies for AMI and Ischemic Stroke

Meeyeon Lee*, Ye-Seul Park*, and Jung-Won Lee*

Abstract

In the medical fields, many efforts have been made to develop and improve Hospital Information System (HIS) including Electronic Medical Record (EMR), Order Communication System (OCS), and Picture Archiving and Communication System (PACS). However, materials generated and used in medical fields have various types and forms. The current HISs separately store and manage them by different systems, even though they relate to each other and contain redundant data. These systems are not helpful particularly in emergency where medical experts cannot check all of clinical materials in the golden time. Therefore, in this paper, we propose a process to build an integrated data model for medical information currently stored in various HISs. The proposed data model integrates vast information by focusing on medical images since they are most important materials for the diagnosis and treatment. Moreover, the model is disease-specific to consider that medical information and clinical materials including images are different by diseases. Two case studies show the feasibility and the usefulness of our proposed data model by building models about two diseases, acute myocardial infarction (AMI) and ischemic stroke.

Keywords

Acute Myocardial Infarction, Data Model, Hospital Information System, Ischemic Stroke, Medical Image, Medical Information, Ontology

1. Introduction

Medical systems including Hospital Information System (HIS) are constantly evolving for digitalizing and efficiently managing medical information. Nowadays most of medical institutions like hospitals naturally use HIS. Recently, many studies [1-10] have tried to efficiently store, transfer, and manage clinical materials. However, HIS consists of several systems such as Order Communication System (OCS), Electronic Medical Record (EMR), and Picture Archiving and Communication System (PACS) and they generate clinical materials of different types. Those systems tend to separately and redundantly include one patient's information. In other words, since information about a patient's condition is contained in various materials, medical experts should read all of charts or reports and interpret all of images or videos before diagnosing. Therefore, systems should be able to provide and manage essential data required in emergency so that medical experts can make a quick decision.

* This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received September 8, 2016; first revision November 18; accepted November 24, 2016.

Corresponding Author: Jung-Won Lee (jungwony@ajou.ac.kr)

* Dept. of Electrical and Computer Engineering, Ajou University, Suwon, Korea (mylee@ajou.ac.kr, yeseuly777@gmail.com, jungwony@ajou.ac.kr)

Therefore, one of the primary issues in HIS is how to effectively manage and provide a patient's information for medical experts, particularly in urgent situations. Some institutions have presented schemes for describing medical and health data [12-14]. However, even though these standards are able to represent and integrate general medical data, it is hard to describe the detailed data and semantics dependent on a specific disease by using them.

In this paper, we propose a new data model which integrates medical information by diseases. The model can represent common medical data as well as disease-specific data. Besides, the data model will support the HISs in order to provide essential data for medical experts who should immediately make a decision about the diagnosis and treatment.

This paper is organized as follows: Section 2 discusses the existing works for medical systems and medical information management. In Section 3, we propose a process to build an integrated data model for a specific disease. Section 4 presents the details of the proposed data model through case studies. Section 5 concludes this paper.

2. Related Works

After HISs have become common, many efforts are ongoing to efficiently represent medical information in EMR [1,2] and to efficiently manage them [3,4]. Some previous works presented analysis or processing methods [5-8] and transfer or exchange systems [9,10] for medical images. However, since they have focused on a specific type of HISs or clinical materials, they did not consider the diversity and semantic relations of medical data in multi-systems. Recently, a few studies have tried to utilize various types of medical data for medical image retrieval using semantic relations with radiology reports [11] and for diagnosis using textural and morphological data [15].

Moreover, many efforts have been taken for the efficient modeling and management of personal health information including medical data. Some of them have focused on a specific type among various clinical materials. [16-18] presented ontology models to analyze and represent semantic data implicitly contained in medical images or videos. However, the scope of those models is limited to just one type of clinical materials and medical systems (PACS).

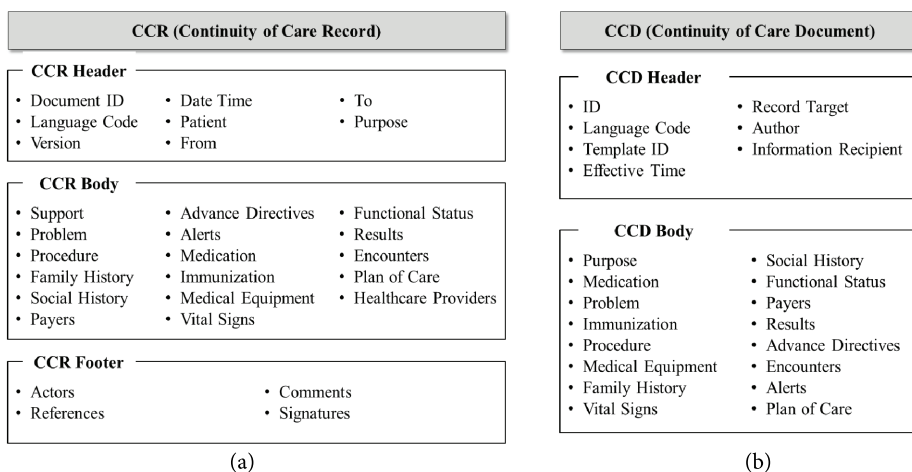


Fig. 1. Representative standard specifications for health records. (a) Schema of CCR, (b) schema of CCD.

As the specifications for personal health data, CCR, CDA, and CCD are representative standards presented by international institutions, American Society for Testing and Materials (ASTM) and Health Level 7 (HL7). Continuity of Care Record (CCR) [12] is a health record standard specification to contain all of a patient's relevant medical history. Fig. 1(a) shows the overview of CCR schema. Clinical Document Architecture (CDA) [13] is a standard to specify the structure and semantics of clinical documents. Continuity of Care Document (CCD) [14] shown in Fig. 1(b) is an XML-based markup standard for specifying a summary of a patient's clinical document. These standards can describe general medical and health information. However, they cannot represent particular data dependent on a specific disease.

3. Integrated Data Model of Medical Information

We suggest a new data model which has the following characteristics or strengths in order to manage vast medical data as a single record. With the data model, a HIS can efficiently manage a patient's information and quickly provide core data which is necessary for a decision about the diagnosis and treatment to medical experts.

- **Data integration:** the current HISs manage a patient's information with diverse materials such as reports (documents), charts, and images/videos. However, it is not easy for medical experts (physicians) to check those materials one by one, particularly in emergency cases. Thus, the proposed model integrates medical data stored in various materials into a single record.
- **Image-centric:** among various clinical materials, the medical images or videos which are generated by medical examinations like magnetic resonance imaging (MRI) and computed tomography (CT) play key roles in diagnosing and treating diseases. Besides, they explicitly and implicitly contain core information about a patient's conditions. In order to analyze correlations between clinical materials and integrate their data, the proposed data model focuses on medical images/videos.
- **Disease-specificity:** medical information is not identical for all of diseases, since types of medical data, clinical materials, and examinations for the diagnosis and treatment of each disease are different. Even in the current hospitals, medical departments have used charts or reports of their own forms. Therefore, a data model should reflect the disease-specificity.

Fig. 2 shows a process to build an integrated data model for a specific disease. The process can define a specification of data model by analyzing clinical materials and the anatomy, centered at medical images. The following section will explain the details about each step through two case studies.

We define a data model as shown in Fig. 3 by applying this process to two diseases. The data model is described as an ontology in order to richly represent relations of integrated data. It consists of two parts; data element and property. The 'Data Element' part contains medical data extracted from clinical materials and is organized with hierarchical classes in an ontology. This part is composed of two sub-parts. 'General Data' is common to all diseases, and 'Disease-specific Data' depend on diseases. The 'Property' part includes attributes and relations of data elements. The details of the data model in Fig. 3 will be explained in the next section.

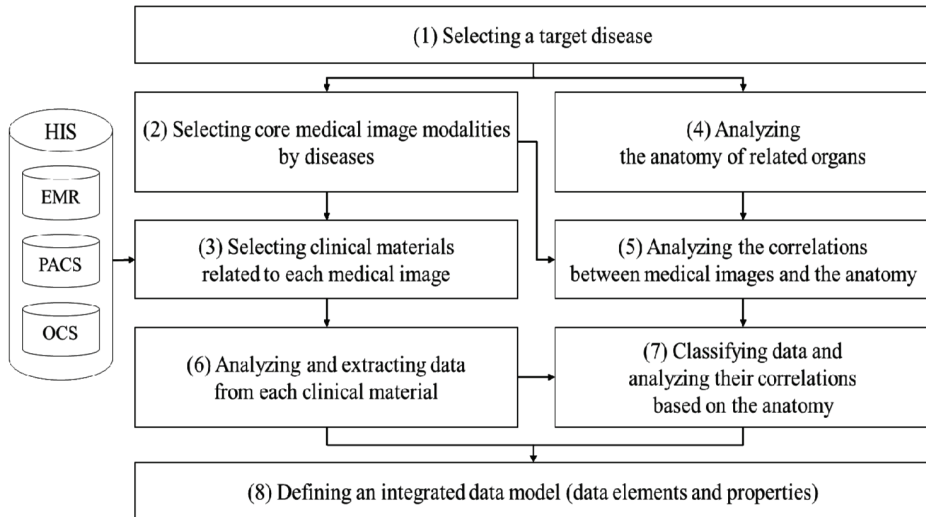


Fig. 2. A process to build an integrated data model by diseases.

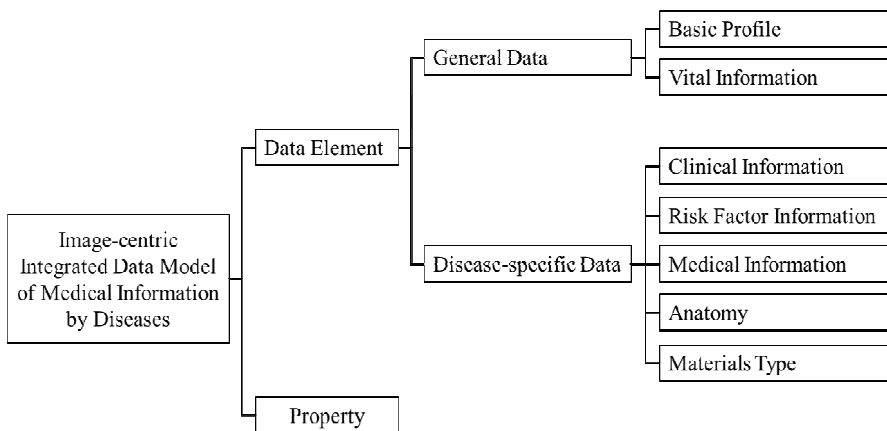


Fig. 3. A brief structure of the proposed data model.

4. Case Study

4.1 Acute Myocardial Infarction

We built a data model for two urgent diseases through the process in Fig. 2. An integrated data model for the first target disease, acute myocardial infarction (AMI), is defined by the following procedures.

- (1) Selecting a target disease: AMI occurs as coronary arteries become blocked or narrowed due to a buildup of plaque. When a patient who suspects AMI comes to an emergency room, medical experts should diagnose and treat him/her within 90 minutes at most by checking his/her medical information. Therefore, AMI is a proper case which needs our model.

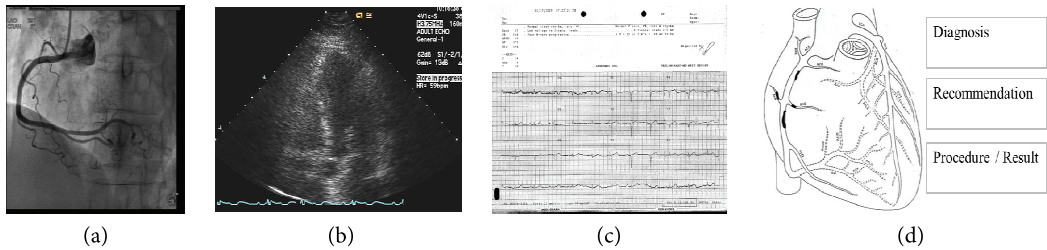


Fig. 4. Samples of four image/video modalities related to AMI. (a) A captured image of a CAG video sample, (b) a captured image of an echocardiogram video sample, (c) a sample of EKG image, and (d) a sample of coronary arteriogram image.

- (2) Selecting core medical image modalities: in a case of AMI, four image/video modalities are essentially used for the diagnosis and treatment and Fig. 4 shows samples. Two modalities are coronary angiography (CAG) and echocardiogram which generate materials of a video type. Electrocardiogram (EKG) produces an image as the examination result. The final modality is a coronary arteriogram which is an image describing a summary of a patient's condition.
- (3) Selecting clinical materials: we chose eight materials including four types of images/videos and four types of reports (documents); CAG, echocardiogram, EKG, coronary arteriogram, CAG report, PTCA and stent deployment report, cardiology lab sheet, and echocardiography laboratory.
- (4) Analyzing the anatomy: Since clinical materials notate names of the organ related to a disease, a data model should include anatomical knowledge. AMI is a disease related to the coronary arteries. Then, we analyze the coronary anatomy including right coronary artery (RCA) and left coronary artery (LCA) and their relations.
- (5) Analyzing the correlations: the medical images/videos selected in step 2 have relations to the anatomy analyzed in step 4. For example, a CAG video of a specific angiographic view shows certain parts of the whole coronary arteries. These relations will be the basis of an image-based information provision service.
- (6) Extracting data: this step analyzes and derives meaningful data explicitly and implicitly contained in eight clinical materials. For instance, a CAG video has explicit data like metadata including date, equipment model and measurement axis. In addition, it has semantic data such as the site of lesion and the degree of severity, which should be interpreted by medical experts.
- (7) Classifying data and analyzing their correlations: some of extracted data may be duplicated in multiple materials. Even though some of them are contained in different materials, they may be related to each other. Therefore, after the previous step, we classify extracted data and analyze their relations, regardless of their source materials.
- (8) Defining data elements and properties: the final step defines a single record as a data model. The data model is described with classes and properties of ontology schema.

The data model has the structure as shown in Fig. 3. 'Basic Profile' of 'General Data' shown in Fig. 5 represents personal information of a patient. Fig. 6 shows 'Vital Information' which includes fundamental information about medical states.

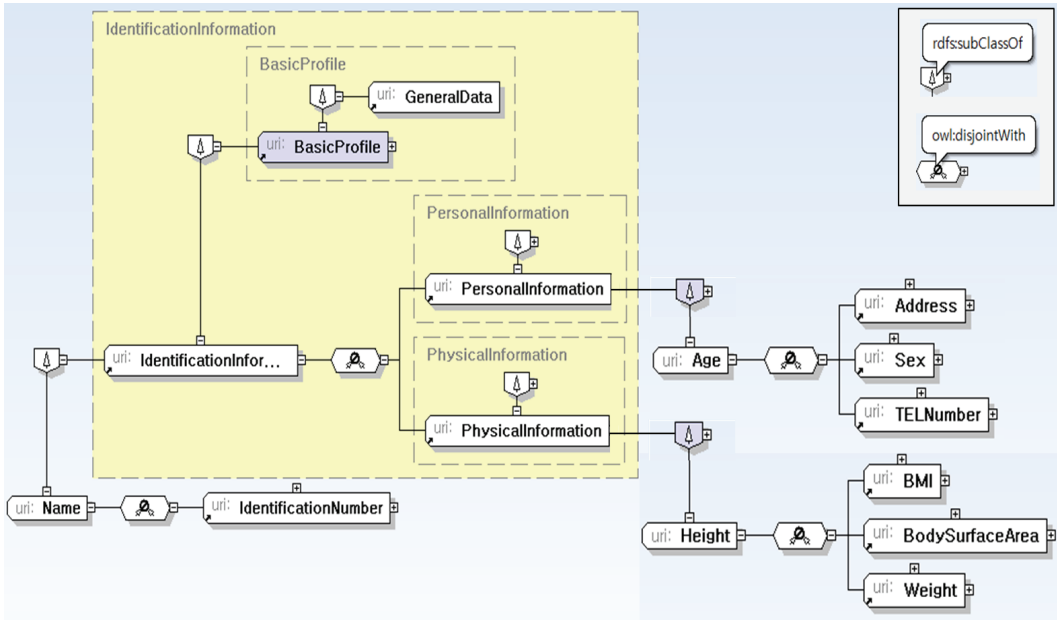


Fig. 5. Structure of BasicProfile.

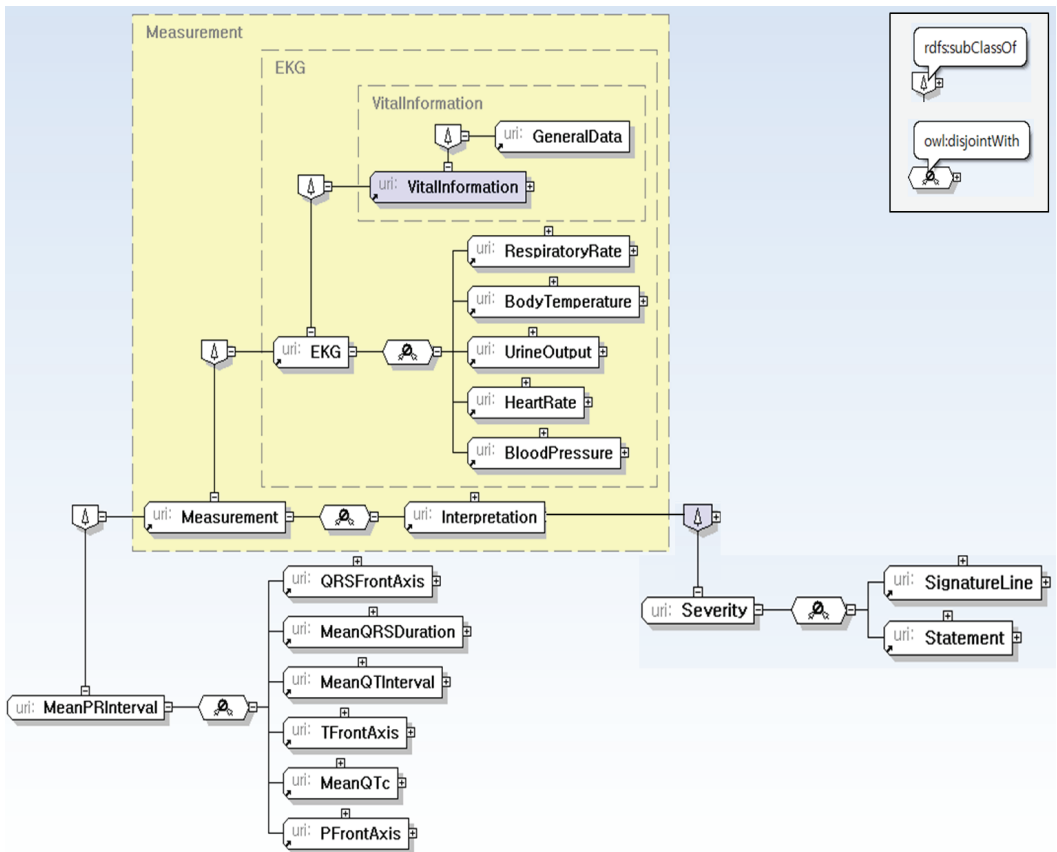


Fig. 6. Structure of VitalInformation.

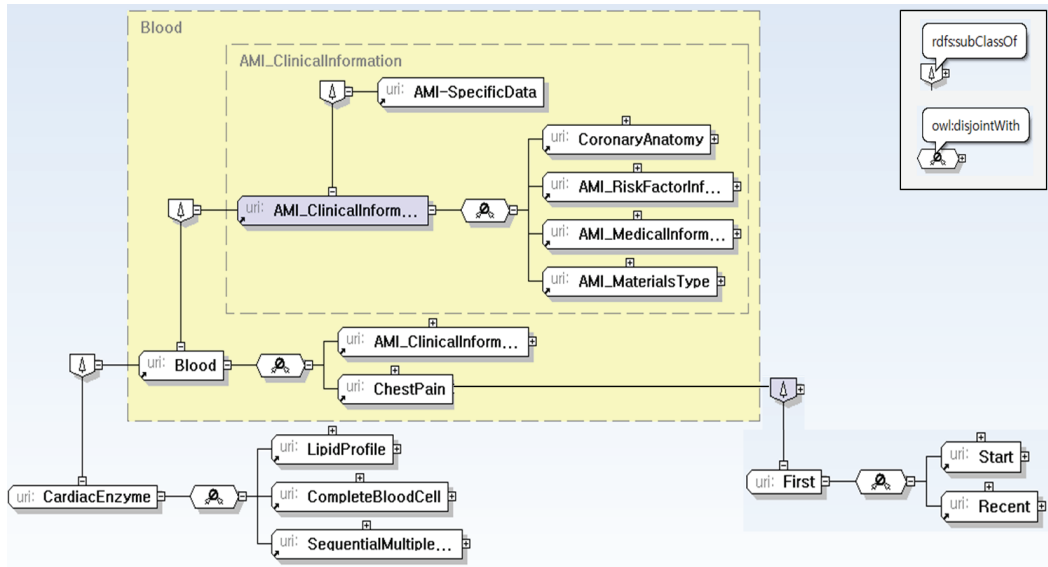


Fig. 7. Structure of AMI_ClinicalInformation.

‘Disease-specific Data’ includes unique data for each disease and consists of the following five categories.

- (1) Clinical Information: a patient’s results of various medical examinations except image/video examinations (as shown in Fig. 7).
- (2) Materials Type: types of reports, documents, images, and videos (Fig. 8(a)).
- (3) Risk Factor Information: factors, such as habits or environmental conditions, which cause a particular disease (Fig. 8(b)).
- (4) Anatomy: anatomical knowledge for the disease (Fig. 9).
- (5) Medical Information: data related to the diagnoses and treatment (Fig. 10).

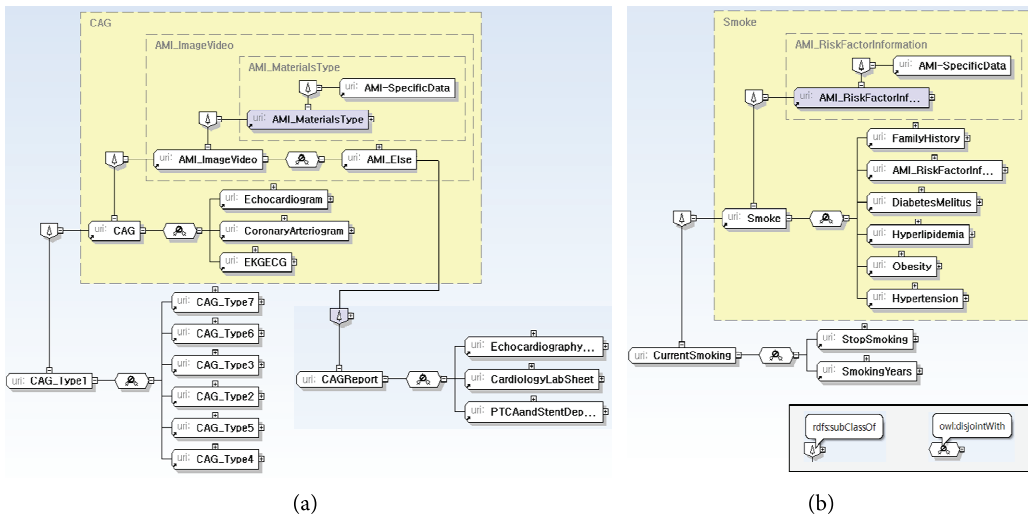


Fig. 8. Structure of AMI_MaterialsType (a) and AMI_RiskFactorInformation (b).

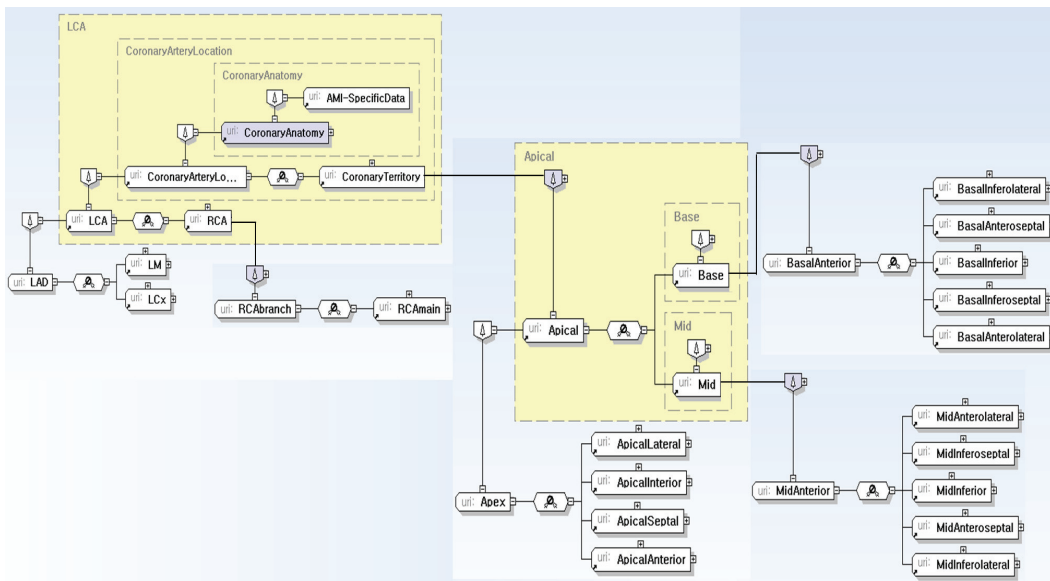


Fig. 9. Structure of CoronaryAnatomy.

In addition to data elements described as classes in ontology, the data model represents their attributes and relations as properties. In our case studies, the ‘Property’ is a common part for two diseases, and it has two groups as explained in Table 1. Object property in ontology specifies relations between data elements (classes), and data property means their attributes.

Table 1. Properties in the data model

Property	Description	
Object property	canShow	The correlations between images/videos and the anatomy (a specific type of images/videos can show a specific part of the organs)
Data property	hasCreator	Institutions or experts who created or modified the data element
	hasDate	The date on which the data element was created or changed
	hasImportanceLevel	A level meaning how important the data element is
	isMandatoryOptional	Whether the data element is mandatory or optional for decision about diagnoses or treatment
	hasReferenceMetadata	Metadata of reference materials related to a data element

4.2 Ischemic Stroke

The second target disease is ischemic stroke. We briefly explain the procedures to build a data model, since they are similar to that of AMI in the previous sub-section.

- (1) Selecting a target disease: ischemic stroke which is one of urgent diseases and occurs when an artery to the brain is blocked.

- (2) Selecting core medical image modalities: CT, CT angiography (CTA), MR, transfemoral cerebral angiography (TFCA), and EKG.
- (3) Selecting clinical materials: in addition to five images/videos, CT report, CTA report, MR report, TFCA report, and EKG report, clinical laboratory sheet, and neurology exam report.
- (4) Analyzing the anatomy: brain anatomy such as brainstem and cerebrovascular supply.
- (5) Analyzing the correlations: analyzing relations between the brain anatomy and five images/videos.
- (6) Extracting data: analyzing and extracting data from twelve materials.
- (7) Classifying data and analyzing their correlations: analyzing redundancy and associations of extracted data.
- (8) Defining data elements and properties.

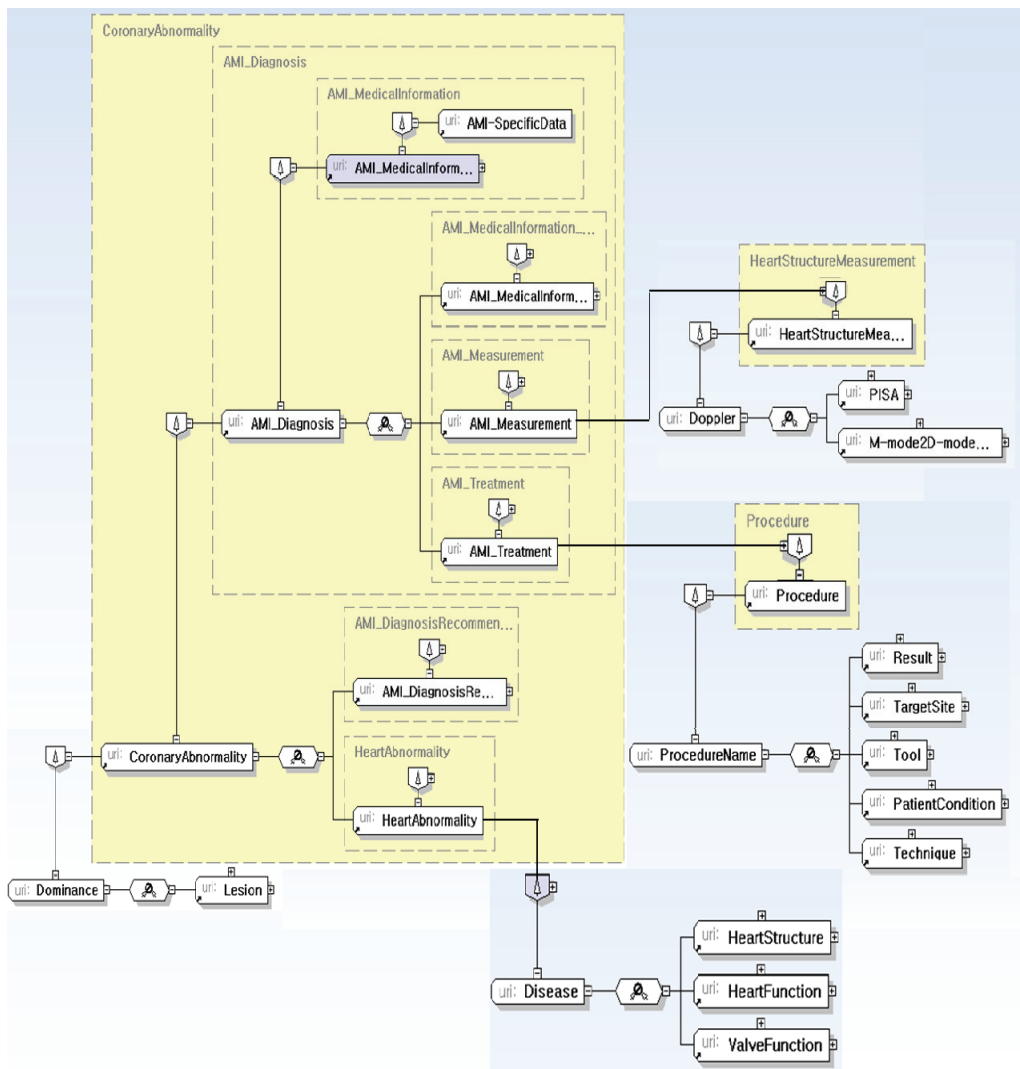


Fig. 10. Structure of AMI_MedicalInformation.

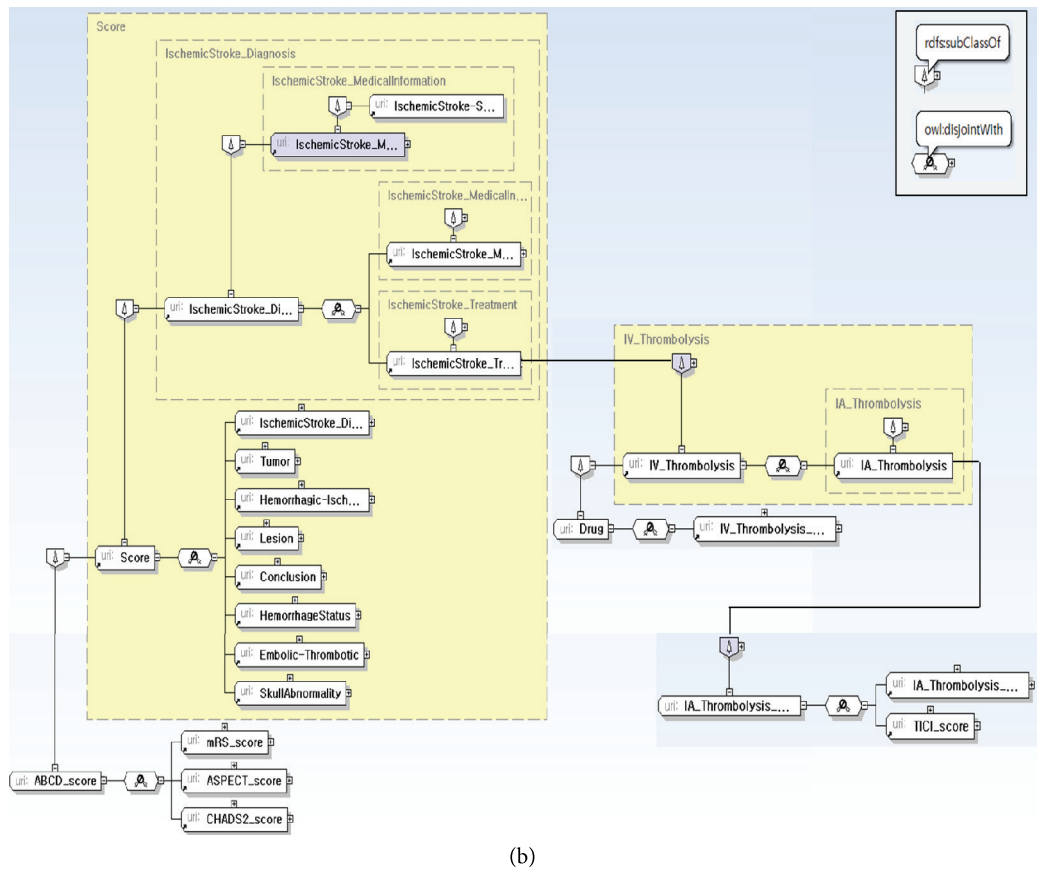
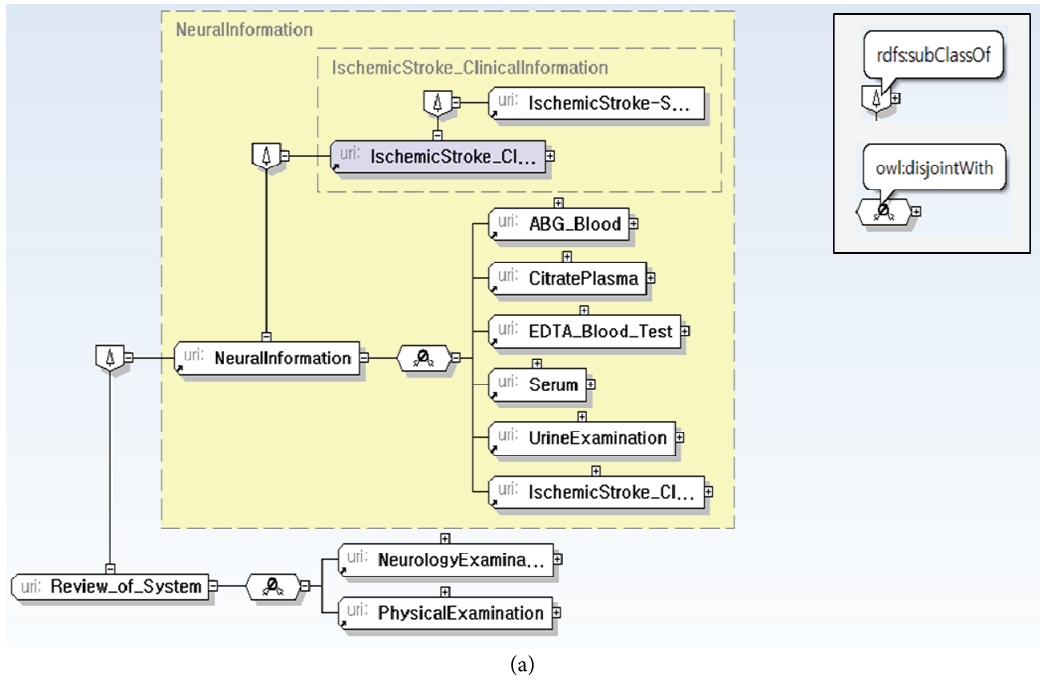


Fig. 11. Structure of IschemicStroke_ClinicalInformation (a) and IschemicStroke_MedicalInformation (b).

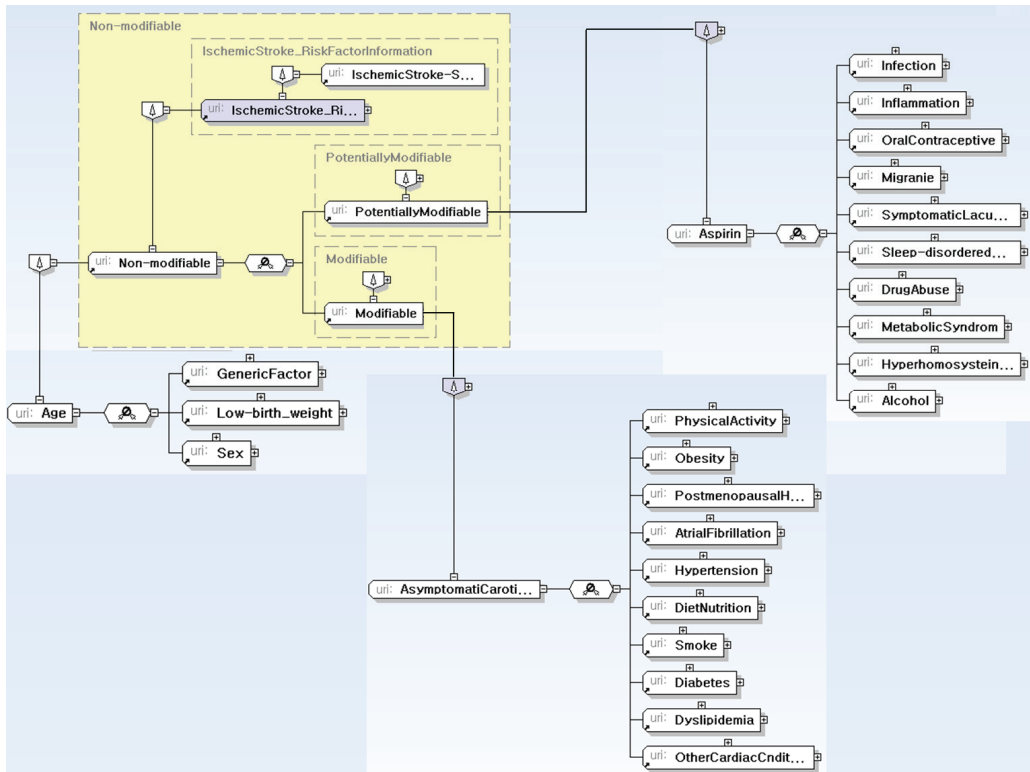


Fig. 12. Structure of IschemicStroke_RiskFactorInformation.

5. Conclusion

The current medical information systems have managed a patient's information by using different systems like EMR, PACS, and OCS. Therefore, since information even for one patient is stored in heterogeneous materials of various systems, medical experts cannot efficiently check important information, especially in urgent situations. In order to improve the way of managing and providing medical data, we have proposed a new scheme for medical data model. The model can integrate heterogeneous information which is distributed in different systems into a single record. A system can provide essential data, rather than materials or files themselves, for medical experts. Moreover, the model focuses on clinical materials of image or video types, since they include core data about a patient's condition. In addition, the data model can consider that essential data for the diagnosis and treatment is different by diseases. That is, the proposed data model integrates heterogeneous medical information for a specific disease by focusing on images and videos.

We have applied the proposed building process to two urgent diseases, AMI and ischemic stroke, and then built a data model. The model consists of two parts; the general part represents data common to all of diseases, and the disease-specific part describes data dependent on a specific disease.

Our future works will focus on extending the data model to cover more diseases and implementing an image-based information provision service based on the proposed data model in order to prove its efficiency.

Acknowledgement

This research was supported by the Ministry of Science, ICT and Future Planning (MSIP), Korea, under the Information Technology Research Center (ITRC) support program (IITP-2016-R2718-16-0015) supervised by the IITP (National IT Industry Promotion Agency).

References

- [1] W. MacKinnon and M. Wasserman, "Implementing electronic medical record systems," *IT Professional*, vol. 11, no. 6, pp. 50-53, 2009.
- [2] C. Zhao and L. Zhang, "Research of information presentation for electronic medical record based on ontology," in *Proceeding of the 6th International Conference on Information Management, Innovation Management and Industrial Engineering*, Xian, China, 2013, pp. 489-492.
- [3] H. S. Na, S. Y. Yun, and S. C. Park, "Design and implementation of mapping system for effective health information data exchange in multi-platform environment," *Journal of Korean Institute of Information Technology*, vol. 10, no. 12, pp. 143-150, 2012.
- [4] S. Perera, C. Henson, K. Thirunarayan, A. Sheth, and S. Nair, "Semantics driven approach for knowledge acquisition from EMRs," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 2, pp. 515-524, 2014.
- [5] H. Greenspan H and A.T. Pinhas, "Medical image categorization and retrieval for PACS using the GMM-KL framework," *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 2, pp. 190-202, 2007.
- [6] Y. Tao, Z. Peng, A. Krishnan, and X. S. Zhou, "Robust learning-based parsing and annotation of medical radiographs," *IEEE Transactions on Medical Imaging*, vol. 30, no. 2, pp. 338-350, 2011.
- [7] J. Agarwal and S. S. Bedi, "Implementation of hybrid image fusion technique for feature enhancement in medical diagnosis," *Human-Centric Computing and Information Sciences*, vol. 5, no. 3, pp. 1-17, 2015.
- [8] S. U. Khan, W. Y. Chai, C. S. See, and A. Khan, "X-ray image enhancement using a boundary division wiener filter and wavelet-based image fusion approach," *Journal of Information Processing Systems*, vol. 12, no. 1, pp. 35-45, 2016.
- [9] F. Valente, C. Viana-Ferreira, C. Costa, and J. L. Oliveira, "A RESTful image gateway for multiple medical image repositories," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 356-364, 2012.
- [10] L. R. Alvarez and R. C. Vargas Solis, "DICOM RIS/PACS telemedicine network implementation using free open source software," *IEEE Latin America Transactions*, vol. 11, no. 1, pp. 168-171, 2013.
- [11] J. Ramos, T. T. Kockelkorn, I. Ramos, R. Ramos, J. Grutters, M. A. Viergever, B. V. Ginneken, and A. Campilho, "Content-based image retrieval by metric learning from radiology reports: application to interstitial lung diseases," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 1, pp. 281-292, 2016.
- [12] *Standard specification for Continuity of Care Record (CCR)*, ASTM E2369-12, 2012.
- [13] Health Level 7 International, "HL7 implementation guide for CDA release 2," [Online]. Available: http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7.
- [14] Health Level 7 International, "HL7 implementation guide: CDA release 2 – continuity of care document (CCD)," 2007 [Online]. Available: http://www.hl7.org/implement/standards/product_brief.cfm?product_id=6.
- [15] K. M. Prabusankarlal, P. Thirumoorthy, and R. Manavalan, "Assessment of combined textural and morphological features for diagnosis of breast masses in ultrasound," *Human-Centric Computing and Information Sciences*, vol. 5, no. 12, pp. 1-17, 2015.
- [16] S. Mhiri, S. Despres, and E. Zagrouba, "Ontologies for the semantic-based medical image indexing: an overview," in *Proceeding of the 2008 International Conference on Information & Knowledge Engineering (IKE)*, Las Vegas, NV, 2008, pp. 311-317.

- [17] D. K. Iakovidis, D. Schober, M. Boeker, and S. Schulz, "An ontology of image representations for medical image mining" in *Proceeding of the 9th International Conference on Information Technology and Applications in Biomedicine*, Larnaca, Cyprus, 2009, pp. 1-4.
- [18] D. L. Rubin, "Finding the meaning in images: annotation and image markup," *Philosophy, Psychiatry, & Psychology*, vol. 18, no. 4, pp. 311-318, 2011.



Meeyeon Lee

She is a lecture professor of the Department of Electrical and Computer Engineering at Ajou University, Korea. She received Ph.D. degree in Computer Science and Engineering from Ewha Womans University, Korea, in 2012. For 2012–2014, she was a post-doctoral researcher in Ubiquitous Convergence Research Institute (UCRi). She was a research assistant professor in Ajou University, Korea, for 2014–2015. Her research interests include ubiquitous computing, mobile computing, context-awareness, data modeling, bio-medical data modeling, and ontology.



Ye-Seul Park

She is in Master's course in Department of Electrical and Computer Engineering at Ajou University, Suwon, Korea. She received B.S. degree in Electrical and Computer Engineering from Ajou University in 2015. Her research interests include data modeling for bio-medical images, ontology and embedded software.



Jung-Won Lee

She is an associate professor of the Department of Electrical and Computer Engineering at Ajou University, Korea, since 2006. For 2003–2006, she was the BK professor in Ewha Womans University. She received her Ph.D. degree in Computer Science Engineering from Ewha Womans University, Korea in 2003. She was a researcher of LG Electronics and did an internship in the IBM Almaden Research Center, USA. Her research interests include mobile context awareness, ontology, bio-medical data modeling and embedded and software.