

## XML 기반 분산 의료정보 통합 모델

김정민<sup>0</sup> 최진영

고려대학교 컴퓨터정보통신대학원

diqtqk@korea.ac.kr<sup>0</sup>, choi@formal.korea.ac.kr

### A XML-based Model for Integration of Distributed Clinical Information

JeongMin Kim<sup>0</sup> JinYoung Choi

Graduate School of Computer Information and Technology, Korea University

#### 요 약

질병을 예방하고 건강을 보호하기 위해서 개인의료정보를 지속적으로 관리하고, 관련 기관이 필요시 손쉽게 이용할 수 있어야 한다. 특히 서로 다른 기관으로부터 발생된 의료정보들을 통합, 관리할 필요가 있다. 이에 본 논문에서는 XML을 기반 기술로 이용하면서, XML만을 이용할 경우 XML 문서의 종류와 개수가 많아지면서 발생하는 문서관리상의 비효율성 문제를 해결하기 위하여 관계형 데이터베이스를 함께 이용한 의료정보 통합모델을 제안한다. 이를 통해 XML과 관계형 데이터베이스의 장점을 함께 이용하여 복잡하고 시간에 따라 변할 수 있는 의료정보를 구조적으로 표현하고, 유연하게 관리하며, 효율적으로 검색할 수 있으며, 향후 시스템 개발에 있어서도 확장성이 용이하다.

#### 1. 서 론

의료분야에서 정보화가 도입되면서 국내외적으로 전자의료기록이 의무화되었고, 의료기관간의 협력관계로 인해 환자의뢰, 감사의뢰, 및 경영정보의 교류가 확산되고 있고, 분산된 의료정보가 의료사고의 가장 큰 원인으로 자리잡으면서 의료기관간의 정보 통합은 중요한 사항이 되었다.[1]

질병을 예방하여 건강한 생활을 영위하고 나아가서는 건강을 증진시키기 위해서는 무엇보다도 개인의 의료정보가 지속적으로 관리되어 정보가 축적되어야 한다. 그러나 병원, 보건소 등의 의료기관에서 생성되는 각종 의료정보는 각 시스템에서 개별적으로 관리되고 있어 개인이 정보를 필요로 할 때 제대로 이용하지 못하고 있다.[2] 또한 의료정보는 환자에게 중요한 개인 정보이고, 병원에게는 항부로 유출할 수 없는 병원의 자산이다. 이런 이유로 의료정보통합은 현실적으로 많은 어려움이 있다. 그러나 이런 비 통합적인 체계는 정보의 재사용성과 환자에게 질 좋은 의료 서비스 제공을 저해하는 요소이다. 따라서 이질적인 의료정보를 통합적으로 관리해야 할 뿐 아니라 서로 다른 시스템들 사이의 데이터 교환을 원활하게 하기 위해 데이터 통합 시스템이 필요하다.[3]

이에 본 논문에서는 XML 기술을 기반으로 한 의료정보 통합 모델을 제안하고자 한다. 기존 연구들과 달리 본 연구에서는 XML을 기반 기술로 이용하면서, XML만을 이용할 경우 XML 문서의 종류와 개수가 많아지면서 발생하는 문서관리상의 비효율성 문제를 해결하기 위하여 관계형 데이터베이스를 함께 이용하고 있다는 점에서 차별화된다.

본 제안 모델은 의료정보를 체계적, 지속적으로 관리하며 관련기관이 필요 시 손쉽게 이용할 수 있도록 하는 데 목표를 두고 있으며, XML과 관계형 데이터베이스의 장점을 함께 이용하여 복잡하고 시간에 따라 변할 수 있는 의료정보를 구조적으로

표현하고, 유연하게 관리하며, 효율적으로 검색할 수 있도록 하고자 한다. 2장에서는 관련연구에 대해서 소개하고, 3장에서는 XML 기술을 기반으로 한 의료정보 통합모델에 대한 연구내용을 소개하며 마지막 4장에서는 결론 및 향후과제에 대해 제안하였다.

#### 2. 관련 연구

인터넷의 급속한 발전과 함께 표준 문서 형식으로 제안된 XML(eXtensible Markup Language)은 상이한 포맷으로 존재하는 문서를 통합 관리하는 데 용이하다. XML은 다양한 형태의 정보를 구조적으로 표현하고 태그에 의미를 부여할 수 있어 보다 상세하게 정보를 검색할 수 있는 등 유연한 정보 이용이 가능하며, 구조적인 정보를 데이터의 손실 없이 데이터베이스에 저장, 검색, 관리할 수 있게 할 뿐만 아니라 기존 정보의 변경이나 새로운 유형의 정보 추가도 쉬운 등 많은 장점을 가지고 있다. 따라서 XML은 데이터의 저장과 전송에 있어 표준언어가 되고 있다.[4]

최근 이러한 XML의 장점을 이용하여 복잡한 정보를 효율적으로 관리하고 이용하고자 하는 연구가 활발히 진행되고 있다. XML 문서를 합성적인(synthetic) 관점에서 재구성하여 정보를 효율적으로 도출하는 방법[4], 확정적이지 않은 확률적인 데이터를 다루기 위해 XML을 이용하는 방법[5], XML을 기반으로 하여 웹 사이트를 효율적으로 관리하는 방법[6], XML을 이용하여 반구조적인(semi-structured) 자료를 구조적으로 추론하는 방법[7] 등 효율적인 정보관리를 위한 여러 가지 연구들이 진행되고 있다.

다양한 이질적인 형식의 데이터를 통합하는 방법에 대해서는 오래 전부터 많은 연구가 이루어져 왔다. 대규모의 이질적인 데이터베이스들의 상호운영과 관련하여 근본적인 어려움인 일관성 유지 문제를 해결하기 위해 메타 데이터를 이용하여 이질적인 데이터베이스들의 페더레이션(federation)에 대한 진화 모델을 제시

하기도 했다[8]. 최근에는 이질적인 형식의 데이터를 통합 관리하기 위해 XML과 관계형 모델을 함께 이용하는 정보관리 방법에 대하여 많은 연구를 진행하고 있다.

보건의료계는 다양하고 복잡한 형태의 정보교환이 이루어지고 있는 가장 대표적인 정보산업분야이다[2]. 상이한 포맷으로 존재하는 문서를 통합 관리하는 데 용이하여 실질적인 정보교환 표준이 되고 있는 XML은 다양한 형태의 복잡한 정보를 다루며 관련 기관간의 정보교환이 활발히 이루어져야 하는 보건의료 분야에 유용하게 활용될 수 있다. 따라서 최근 XML을 이용하여 병원 및 관련기관 사이에 정보를 교환하고 공유하고자 하는 연구들이 진행되고 있다[9]. 보건의료 분야에서 XML을 직접 이용하지는 않았으나 관련된 연구로는, 진료정보 전송 표준인 HL7을 기반으로 한 병원정보시스템 설계에 관한 연구가 있다.

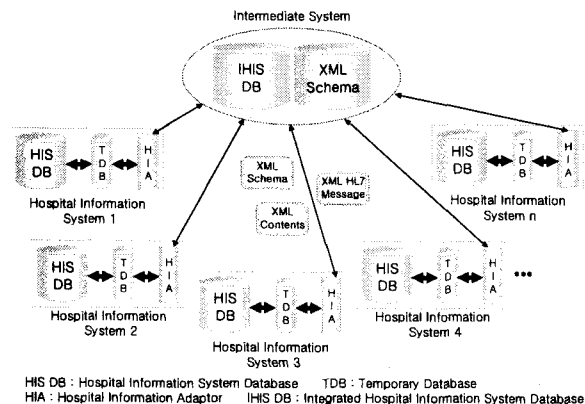
### 3. 연구 내용

#### 3.1 의료정보 통합모델의 전체구조

의료정보 통합관리를 위한 모델의 전체적인 구조는 [그림1]와 같이 각 의료기관의 의료정보시스템(Hospital Information System; HIS)이 중앙의 중개시스템(Intermediate System)과 연결되어 있으며 메시지를 주고받는 형태를 취한다. 여기서 의료정보시스템이란 개인의 건강정보가 생성되고 관리되는 의료기관 개별 시스템 각각을 가리키며, 중개시스템이란 개별 의료정보시스템에서 생성된 정보를 통합적으로 관리하면서 개별 의료정보시스템에서 필요로 하는 의료정보를 적절한 형태로 제공하는 의료정보 관리에 있어 중개 역할을 하는 시스템을 가리킨다.

각 의료정보시스템에서 생성된 의료정보는 의료정보시스템의 데이터베이스(HIS DB)에 저장되고, 지속적인 건강관리를 위해 필요한 의료정보는 별도의 임시 데이터베이스(TDB)에 저장된다. TDB에 저장된 의료정보는 의료정보 어댑터(HIA)에 의해 변환되어 중개시스템으로 전송된다 [그림1].

중개시스템으로 전송된 의료정보는 중개시스템의 통합의료정보 데이터베이스(IHIS DB)에 저장되어 관리된다. 전에 이용했던 의료기관이 아닌 다른 곳에서 필요한 정보를 확인하고자 할 때에는 수시로 중개시스템에 저장되어 있는 개인의의료정보를 검색할 수 있다[그림1].



HIS DB : Hospital Information System Database TDB : Temporary Database  
HIA : Hospital Information Adaptor IHIS DB : Integrated Hospital Information System Database

그림1. Configuration of integrated hospital information system

#### 3.2 의료정보 표현과 관리방법

본 연구 모델에서는 XML과 관계형 데이터베이스 모델을 함께 이용하여 복잡하며 시간에 따라 변할 수 있는 의료정보를 계층적으로 표현하여 관리하고 있다. [그림2]에서 볼 수 있듯이 정보 표현의 형태는 크게 두개의 계층 즉, 관계형 모델을 기반으로 한 상위 계층(Upper level hierarchy)과 XML을 기반으로 한 하위 계층(Lower level hierarchy)으로 구분할 수 있다.

상위 계층은 관계형 모델의 구조에 기반을 두고 다시 세 개의 서브 계층 즉, 메타 테이블 계층, 데이터 테이블 계층, 데이터 필드 계층으로 구분된다[그림2]. 가장 상위의 서브계층인 메타 테이블 계층은 하위 서브계층에 있는 데이터 테이블들 사이의 연관 관계 또는 데이터 테이블의 활용을 위해 필요한 정보를 가지고 있는 계층으로, 데이터 테이블에 저장된 정보의 분산 내역에 관한 데이터, 데이터 테이블이 필드로 가질 수 있는 XML 정보에 대한 데이터 등이 정의되어 있다. 정보의 분산 내역에 관한 데이터란 검색 효율을 높이기 위해 서버의 규모에 맞추어 데이터를 분할 또는 분산시킬 때의 상세 내역을 의미하며, XML 정보에 대한 데이터란 저장하고자 하는 정보의 세부 주제별로 필요에 따라 만들어 지는 XML 문자열(string)에 대한 ID, 스키마 등의 데이터를 의미한다.

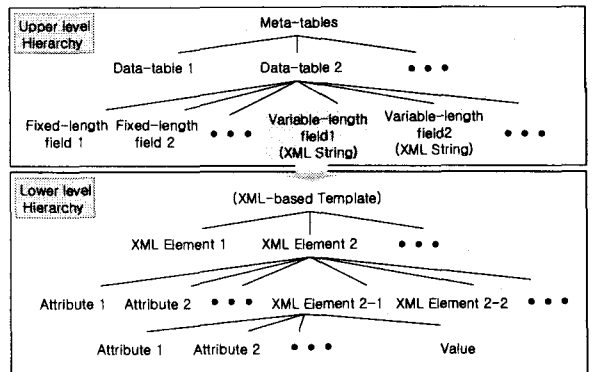


그림2. Hierarchical structure of information representation

그 다음 서브계층인 데이터 테이블 계층은 실제 데이터가 저장되어 있는 데이터 테이블들로 구성되어 있으며, 각 데이터 테이블은 여러 개의 필드들을 포함하고 있다. 데이터 테이블 계층의 하위에 있는 서브계층은 데이터 필드 계층이다. 데이터 필드 계층에 있는 데이터 필드는 크게 두 가지 종류로 구분된다. 즉, 데이터의 구조와 형식이 간단하고 시간의 흐름에 따라 변할 가능성이 거의 없는 고정길이필드(Fixed-length field)와 데이터의 구조와 형식이 복잡하고 시간의 흐름에 따라 변할 가능성이 많은 가변길이필드(Variable-length field)로 분류된다. 고정길이필드는 일반적인 관계형 모델의 데이터 필드와 같은 방식으로 표현되고, 가변길이 필드는 XML로 표현된 문자열로 표현된다.

하위 계층은 상위의 데이터 필드 계층의 필드가 XML로 표현된 가변길이필드인 경우에 적용되는 계층이다. 이 하위 계층은 XML로 표현된 정보가 일반적으로 여러 개의 계층으로 표현되는 것과 동일하게 다수의 계층으로 구성된다[그림2]. 계층의 수는 일반적인 XML로 표현된 정보처럼 데이터의 내용에 따라 다르다. 가변길이

이필드가 가질 수 있는 XML 정보에 대한 스키마는 상위 계층의 메타 테이블에 정의되어 있다. 따라서 본 연구에서는 이렇게 데이터의 구조와 형식이 복잡하고 시간에 따라 변할 수 있는 정보는 XML로 표현하여 가변길이필드 형태로 저장한다.

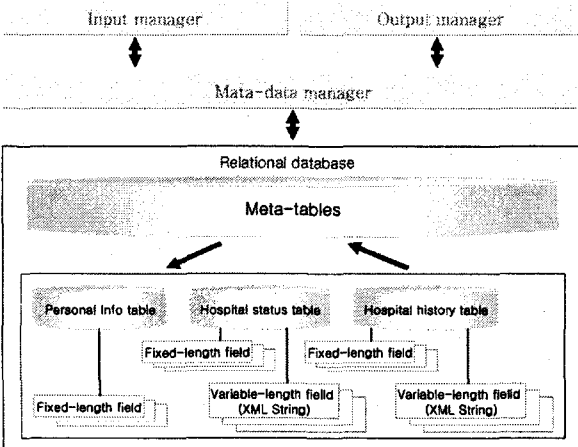


그림3. Architecture of intermediate system

3.3 의료정보 중개시스템

3.3.1 중개시스템 구조

의료정보 중개시스템의 전반적인 구조는 [그림3]와 같이 일반적인 정보시스템의 구조와 비슷하다. 개별 의료정보시스템에서 전송되어 온 수신 메시지는 입력관리자(Input Manager)를 통하여 데이터베이스에 저장되고, 개별 의료정보시스템으로 전송될 메시지는 출력관리자(Output Manager)를 통하여 작성되어 전송된다. 데이터베이스의 물리적인 구조는 일반적인 관계형 데이터베이스와 같다. 그러나 정보가 저장된 데이터베이스 내부의 논리적인 구조는 일반적인 관계형 데이터베이스만을 이용하는 경우와 다르다. 일반적인 데이터베이스 테이블의 구조를 정의하고 있는 메타 테이블들이 상위에 있고, 그 하위에 실제 데이터가 저장되어 있는 데이터 테이블들이 있다. 각 데이터 테이블 내부의 논리적인 구조는 고정길이필드인 경우와 가변길이필드인 경우로 구분하여 다른 구조를 취한다. 가변길이필드의 구조는 메타 테이블에 정의되어 있으며, 필요에 따라 가변길이필드에 대한 구조를 추가하고 데이터를 추가할 수 있어 시변 환경에 쉽게 적용할 수 있다. 가변길이필드 추가 및 변경에 대한 정보는 전송 메시지 입력력 시 메타 데이터 관리자(Meta-data Manager)를 통하여 메타 테이블에 저장한다.

3.3.2 중개시스템의 데이터베이스 구성

개인의료정보의 통합적 관리를 위한 데이터베이스 테이블은 크게 메타 테이블과 데이터 테이블로 구분된다. 메타 테이블에는 분산색인 테이블(tblDistIndex)과 XML스키마 테이블(tblXMLSchema)이 있으며, 데이터 테이블에는 개인신상정보 테이블(tblPersonalInfo), 최신의료정보 테이블(tblHospitalStatus), 과거의료정보 테이블(tblHospitalHistory)이 있다( 표1).

메타 테이블과 개인신상정보 테이블은 고정길이필드로만 구성

되어 있는 일반 관계형 데이터베이스 테이블과 동일하다. 개인신상정보 테이블은 회원별 1개의 레코드가 존재한다.

<표 1> Database tables of intermediate system

Table	Field
Distributon Index Table (tblDistIndex)	*DL_DataTableID(Data Table ID)
	*DI_FirstMemberID(First Member ID)
	DI_LastMemberID(Last member ID)
Meta Table	...
	*XS_DataTableID(Data Table ID)
	*XS_XMLSchemaID(XML schema ID) XS_XMLContents(XML schema details)
Personal Information Table (tblPersonalInfo)	...
	*PI_MemberID(Member ID) PI_MemberName(Member ID) PI_civiNum(Civilization number) PI_Gender(Gender) PI_BloodTypeRH(Rh blood type)
	...
Hospital Status Table (tblHospitalStatus)	*PI_MemberID(Member ID) HS_Date(Check date) HS_Place(Check place) HS_Inspector(Inspector) *HS_Classification(Classification code of hospital information) HS_Result(Check Result):XML String
	...
	*PI_MemberID(Member ID) HH_Date(Check date) HH_Place(Check place) HH_Inspector(Inspector) *HH_Classification(Classification code of hospital information) HH_Result(Check Result):XML String
Hospital History Table (tblHospitalHistory)	...
	*PI_MemberID(Member ID) HS_Date(Check date) HS_Place(Check place) HS_Inspector(Inspector) *HS_Classification(Classification code of hospital information) HS_Result(Check Result):XML String
	...

\* field for primary key

최신의료정보 테이블과 과거의료정보 테이블은 같은 테이블의 구조를 취하고 있으며, 둘 다 5개의 고정길이필드와 1개의 가변 길이필드로 구성되어 있다. 고정길이필드는 회원번호, 측정일, 측정장소, 측정인, 건강정보분류의 5개 필드이고, 가변길이필드는 건강정보 분류별로 각각 다른 형태의 XML 문자열로 표현된다.

최신의료정보 테이블은 회원별 의료정보 분류별로 1개의 레코드가 존재하며, 과거의료정보 테이블은 회원별 의료정보 분류별 측정일별로 1개의 레코드가 존재한다. 최신의료정보 테이블과 과거의료정보 테이블의 스키마가 동일하지만 최신의료정보 테이블을 별도의 테이블로 구성한 이유는 가장 자주검색되는 정보이므로 정보검색시간을 줄이기 위해서이다.

3.3.3 중개시스템의 의료정보 관리방법

관리하고자 하는 정보 중 데이터의 구조나 형식 등 시간의 흐름에 따라 변하지 않는 경우에는 일반적인 관계형 데이터베이스 설계 시처럼 고정길이필드로 만든다. 정보의 입력, 수정 및 삭제도 일반적인 관계형 데이터베이스를 이용할 때처럼 관리한다.

의료정보 중 데이터의 구조와 형식이 시간의 흐름에 따라 변할 수 있는 가변길이필드를 갖는 세부 의료정보는 몇 단계의 과정을 거쳐 정보를 관리한다. 먼저 정보의 내용과 용도에 따라 세부의료정보를 분류하고, 세부 의료정보별로 XML 문서 형식으로 스키마를 설계한다. XML 스키마에 대한 설계가 완료되면, XML 스키마

를 구분할 수 있는 XML 스키마 ID를 부여한 후 새로 작성한 XML 스키마를 메타데이터에 저장한다.

사용하고자 하는 가변길이필드에 대한 XML 스키마가 메타 데이터에 저장되면, 데이터 테이블에 정보를 입력할 수 있다. 정보 입력 시 가변길이필드가 있는 경우에는 해당 XML 스키마를 기준으로 정보의 유효성(Validity)을 검사한다. 이러한 입력 작업은 입력 관리자(Input Manager)를 통하여 이루어진다.

정보의 수정과 삭제 역시 입력 관리자를 통하여 이루어진다. 기존 XML 스키마의 변경 없이 정보를 수정할 경우에는 새로운 정보를 추가하는 경우와 마찬가지로 가변길이필드가 있으면 메타 데이터에 저장된 해당 XML 스키마를 기준으로 유효성 검사작업을 거치게 된다.

기존의 XML 스키마를 변경하여 정보를 수정할 경우에는 기존 XML 스키마는 그대로 두고 변경되는 XML 스키마를 추가하는 방식으로 처리한다. 이때 추가되는 XML 스키마에게는 신규 추가시 처럼 새로운 ID를 부여한다. 기존 XML 스키마를 변경하면 기존에 입력된 정보를 관리할 수 없으므로 그대로 두고 계속 사용할 수 있게 한다. XML 스키마가 변하더라도 정보를 검색할 때 해당 가변길이필드에 저장되어 있는 XML 트리 전체를 하나의 자료로 처리하므로 쉽게 처리할 수 있으며, XML 트리의 세부구조 및 내용은 메타 데이터에 저장되어 있는 해당 XML 스키마를 통하여 확인할 수 있다.

데이터가 고정길이필드이건 가변길이필드이건 관계없이 모든 데이터가 관계형 데이터베이스의 일반 데이터처럼 필드별로 저장되어 있다. 단지 가변길이 필드인 경우에는 하나의 필드 내에 XML 형태의 정보가 포함되어 있다는 점에 차이가 있다. 따라서 정보의 검색은 관계형 데이터베이스에 저장되어 있는 정보를 검색하는 일반적인 방법으로 처리하면 된다. XML 형태로 저장되어 있는 정보중의 특정 데이터를 검색하고자 할 경우에는 지금까지의 많은 연구개발 결과를 토대로 하여 일반적으로 많이 이용하고 있는 XML 정보 검색 방법을 이용하면 된다.

3.4 건강정보 전송 메시지의 종류와 구조

각각의 의료정보시스템과 중개시스템 간에 전송되는 메시지의 종류는 (표 2)와 같이 크게 4가지 의료정보 스키마, 의료정보 내용, XML형태의 HL7 메시지, 그리고 확인 메시지로 구분된다. 의료정보 스키마는 중개시스템에서 작성되어 관리되며 변화될 때마다 개별 의료정보시스템으로 전송된다. 의료정보 내용과 HL7 메시지는 개별 의료정보시스템에서 생성되어 변화과정을 거쳐 중개시스템으로 전송되고, 중개시스템에서 통합적으로 관리된다. 의료정보 내용과 HL7 메시지는 개별 시스템에서 필요 시 중개시스템에 축적되어 있는 데이터베이스로부터 개별 시스템으로 전송된다. 모든 의료정보 전송 메시지는 XML 문서의 구조로 표현된다. 의료정보 스키마의 세부구조는 XML 스키마의 구조를 취하고,

의료정보 내용의 세부구조는 XML 문서의 구조를 취한다. HL7 메시지의 세부구조는 XML 형태의 HL7 메시지와 같다. [그림 4]는 XML 문서의 구조로 작성된 의료정보 내용 메시지의 한 예이다. 모든 메시지는 메시지의 종류와 ID 등의 정보를 포함하고 있는 메시지 헤더를 XML 요소로 가지고 있다. [그림 5]은 수신확인 메시지의 한 예이다.

```
<?xml version="1.0"?>
<MessageHeader>
  <MessageType>HI_Contents</MessageType>
  <MessageID>HI_20060101001</MessageID>
</MessageHeader>
<BodyComposition>
  <Stature>172cm</Stature>
  <Weight>65kg</Weight>
  <BMI>20.67</BMI>
  <PercentBodyFat>20</PercentBodyFat>
  ...
</BodyComposition>
```

그림4. Example of transmission messages

```
<?xml version="1.0"?>
<MessageHeader>
  <MessageType>HI_ACK</MessageType>
  <MessageID>HI_20060101001</MessageID>
</MessageHeader>
<Result>OK</Result>
```

그림5. Example of transmission acknowledgement messages

3.5 의료정보 어댑터

개별 의료정보시스템에 필요한 의료정보 어댑터(Hospital Information Adaptor: HIA)는 각 의료정보시스템에서 생성되는 개인의 의료정보를 변환하여 중개시스템으로 전송하는 기능을 수행하며, [그림 6]과 같이 크게 메시지 변환처리 모듈(Message translation module)과 메시지 전송 모듈(Message transmission module)로 구성되어 있다.

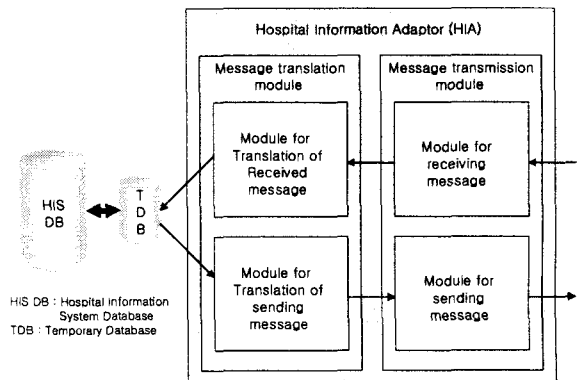


그림6. Architecture of hospital information adaptor(HIA)

<표 2> Classification of transmission messages

Message type	Contents of message
HLSchema	Schema of hospital information
HL_Contents	Contents of hospital information
HL_HL7	HL7 message
HL_ACK	Transmission acknowledgement message

메시지 변환처리 모듈은 메시지 송신변환 모듈(Module for translation of sending message)과 메시지 수신변환 모듈(Module for translation of received message)로 세분화되어 있다. 메시지 송신변환 모듈은 임시 데이터베이스(Temporary Database: TDB)에 저장되어 있는 전송 대상 건강정보를 중개시스템으로부터 전송받은 표준 메시지 포맷을 기준으로 하여 XML

문서로 변환하고 그 결과를 메시지 전송모듈로 보낸다.

메시지 전송 모듈은 메시지 송신 모듈(Module for sending message)과 메시지 수신 모듈(Module for receiving message)로 세분화되어 있다. 메시지 송신모듈은 메시지 변환처리 모듈로부터 받은 XML 문서 형식의 건강정보 전송 메시지를 중개시스템으로 전송한다. 메시지 수신 모듈은 중개시스템으로부터 전송받은 건강정보 메시지를 메시지 변환처리 모듈로 보낸다.

메시지 변환처리 모듈 중 메시지 수신변환 모듈은 메시지 전송 모듈로부터 전송받은 의료정보 메시지를 표준 메시지 포맷을 기준으로 하여 XML 문서 형식의 메시지를 다시 일반 정보의 형태로 변환하고 그 결과를 임시 데이터베이스(Temporary Database: TDB)에 저장한다.

### 3.6 실험 및 분석

#### 3.6.1 정보 표현과 관리 측면의 효과 분석

정보 표현과 관리 측면에서의 효과 분석은, 중개시스템의 설계 및 구현의 과정과 결과를 기초로 하여, 기존의 정보 표현 및 관리 방법들과 정성적으로 비교 분석하였고, 본 논문에서 제안하는 XML과 관계형 데이터베이스(RDB)를 함께 이용한 방법을 기존의 RDB만을 이용하거나 파일형태를 기반으로 한 방법과 정보 표현 및 관리 측면에서 비교하였다. 정보표현의 효율성 측면에서 XML과 RDB를 함께 이용한 방법은 다른 방법들에 비해, 복잡하고 다양한 정보를 XML의 장점을 이용하여 구조적으로 이해하기 쉽게 잘 표현할 수 있으며, 기존의 정보를 수정하거나 새로운 유형의 정보를 추가해야 하는 경우 쉽게 대처할 수 있었다. 본 연구에서 다루는 의료정보는 개인의 다양한 정보로, 그 종류가 많은 편이다. 따라서 측정환경이나 개인의 상태 등에 따라 수정 또는 추가해야 하는 항목들이 많다. 그러나 본 프레임워크 상에서는 복잡한 가변 길이 정보의 데이터 각 항목을 RDB나 파일의 각각의 필드로 다루지 않고, 전체의 내용을 XML 스트림으로 작성하여 RDB의 하나의 필드로 저장하여 관리함으로써 변경을 쉽게 처리할 수 있었다.

#### 3.6.2 중개시스템의 성능 분석

중개시스템의 정보 검색에 대한 성능을 측정하고 그 효율성을 분석하는 것은 본 연구에서 제안한 방법인 관계형 데이터베이스(RDB)와 XML을 혼용한 방법(방법1)과 다른 2가지 방법 즉, 모든 정보를 RDB의 일반적인 필드로 구성하는 방법(방법2), 모든 정보를 XML로 표현하고 관계형 데이터베이스 관리 시스템에서 제공하는 XML 저장 및 검색 기능을 이용하여 XML 형태로 RDB에 저장하고 검색하는 방법(방법3)의 성능을 비교하는 방식으로 실시하였다.

실험에 적용한 개인의료정보의 내용은 6가지 분류 즉, 신체구성(BC), 병력(MH), 의학검사(ME), 체력측정(FM), 운동처방(EP),

식이처방(DP)에 국한하였으며, 검색은 점차적으로 방대해지는 데이터용량으로 인하여 시스템의 성능을 가장 크게 좌우하는 과거 의료정보를 검색하는 부분으로 한정하였다. 3가지 방법 각각에 대하여 과거의료정보를 저장 및 검색할 수 있도록 데이터베이스를 구성하였다. 본 연구에서 제안하는 방법1의 데이터베이스 구성은 앞장에서 설계된 대로 과거의료정보가 1개의 테이블에 저장되도록 만들었다.

모든 정보를 RDB의 일반적인 필드로 구성하는 방법2의 경우, 의료정보의 분류에 따라 다른 스키마를 갖게 되므로 과거의료정보 테이블을 하나의 테이블로 구성하지 않고 의료정보 분류에 따라 분류하여 여러 개의 테이블로 구성하였다. 본 실험에서는 6가지 의료정보 분류에 대하여 적용하였으므로, 방법2에서의 테이블의 개수는 6개이다.

모든 정보를 XML로 표현하여 RDB에 XML 형태로 저장하는 방법3의 경우, 방법2에서와 마찬가지로 의료정보의 분류에 따라 다른 스키마를 갖게 되므로 과거의료정보 테이블을 하나의 테이블로 구성하지 않고 의료정보 분류에 따라 분류하여 여러 개의 테이블로 구성하였다. 따라서 방법2에서와 마찬가지로 방법3에서의 테이블의 개수도 6개이다.

본 연구에서는 웹 환경에서 정보를 관리하는 것을 전제로 하고 있으므로, 정보검색의 결과는 XML 형태로 보여주는 것을 기본으로 하였다. 따라서 과거의료정보가 XML 형태로 저장되지 않는 방법2의 경우에는 정보검색시간에 XML 형태로 변환해 주는 시간이 포함되어 있다. 3가지 방법에 대하여 데이터베이스에 저장되어 있는 레코드의 건수를 1,000건부터 160,000건까지의 범위에서 변화의 추이를 감안하여 레코드 수를 증가시키면서 정보 검색 시간을 측정하였으며, 그 결과는 (표 3)에 요약되어 있다.

[그림 7]에서 볼 수 있듯이 데이터베이스에 저장되어 있는 레코드의 건수가 5만 건 이하일때는 제안한 방법(방법1)이 다른 2가지 방법에 비해 정보검색시간이 월등히 짧은 결과를 보이면서 큰 변화를 보이지 않았다. 그러나 데이터베이스에 저장되어 있는 레코드의 건수가 6만 건 정도를 넘으면서부터는 제안 방법의 정보 검색 시간이 다른 2가지 방법에 가까워지기 시작했고 7만 건 정도부터는 거의 비슷한 시간이 소요되었다. 13만 건 정보가 되면서부터는 검색시간이 그 전에 비해 급격히 길어지기 시작했으며, 15만 건 정도가 되면서부터는 그 전에 비해 꽤 많은 시간이 소요되었다.

방법1에서는 모든 과거의료정보가 1개의 테이블에 저장되어 있는 반면, 방법2와 방법3의 경우는 의료정보 분류별로 6개의 테이블에 나누어 저장되므로 방법1에 비해 테이블 당 레코드 건수가 6분의 1에 불과하다. 따라서 상대적으로 레코드 건수의 증가에 영향을 덜 받은 것으로 분석되었다.

따라서 본 연구에서 제안하는 방법에서는 데이터량(레코드 수)의 증가에 따라 서버의 성능과 용량을 감안하여 적절하게 데이터

<표 3> Information search time by number of records

Number of records*	1	5	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160
Method 1	1.4	0.7	1.7	1.1	1.5	3.3	2.0	3.8	4.3	6.4	4.2	4.9	5.3	5.5	9.9	9.3	131.7	873.7
Method 2	4.7	4.9	4.9	4.0	4.5	5.0	5.9	4.6	5.3	7.2	6.0	4.9	8.3	6.8	11.6	8.6	96.6	469.0
Method 3	4.2	3.9	4.2	4.4	3.9	4.4	5.3	4.1	4.3	5.3	4.0	5.3	5.2	5.0	10.6	5.3	8.5	19.4

\* Unit of number of records: thousand, + Unit of search time: ms

테이블을 분할 또는 분산시키는 것이 필요하다. 데이터 테이블 분할 및 분산에 관한 세부정보는 앞에서 기술했듯이 메타 테이블 계층에서 정의하게 된다.

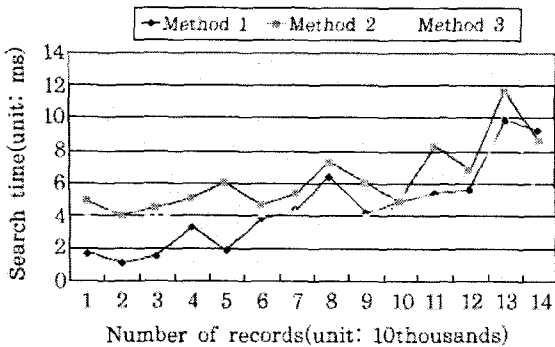


그림7. Analysis of information search time by number of records

4. 결론 및 향후 과제

본 논문에서는 국민의 건강관리 및 건강증진에 필요한 개인의 의료정보를 체계적, 지속적으로 관리하며 관련기관 및 개인이 필요시 손쉽게 이용할 수 있도록하는 의료정보시스템 통합모델을 제안하였다. 이를 위해 본 모델에서는 XML과 관계형 데이터베이스의 장점을 함께 이용하고 있다. 실험을 통해 본 논문에서 제안하는 XML과 관계형 데이터베이스를 함께 이용한 개인의료정보 통합모델은 XML을 이용하여 정보를 표현함으로써 복잡하고 다양한 정보를 구조적으로 표현하고, 필요에 따라 정보를 여러 가지의 주제로 분류하여 관리함으로써 복잡하고 방대한 정보를 관리하는데 있어서의 어려움을 극복할 수 있음이 확인되었다. 또한 시간에 따라 변하는 스키마에 관한 모든 내용을 상위 개념의 메타 테이블에 저장하여 지속적으로 관리함으로써 시스템은 스스로 진화하게 되고, 정보 검색 시 필요한 세부 데이터의 구조를 쉽게 확인하고 이용할 수 있게 한다.

XML과 관계형 데이터베이스를 함께 이용한 중개시스템의 정보 검색에 대한 성능을 측정하고 그 효율성을 분석한 결과, 모든 정보를 RDB의 일반적인 필드로 구성하는 방법(방법2)과 모든 정보를 XML로 표현하고 관계형 데이터베이스 관리 시스템에서 제공하는 XML 저장 및 검색 기능을 이용하여 XML 형태로 RDB에 저장하고 검색하는 방법(방법3)에 비해, 운영 환경에 맞추어 데이터 테이블을 적절히 분할 또는 분산시키는 것을 전제로 할 때, 정보 검색시간이 짧은 결과를 보여 성능이 우수함이 확인되었다.

본 논문에서는 아직까지는 가장 손쉽게 많이 이용하고 있는 관계형 데이터베이스를 기본적으로 사용하는 것을 전제로 하였다. 따라서 최근 개발되어 부분적으로 활용되고 있는 객체지향 데이터베이스를 기반으로 한 XML 전용 데이터베이스는 고려하지 않았다. 향후 XML 전용 데이터베이스를 이용하는 경우와 비교하여 실험해 보고 성능을 평가하는 작업이 필요하다. 개인의료정보는 개인의 프라이버시에 해당하는 정보를 많이 포함하고 있다. 따라서 정보의 보안관리가 아주 중요하다. 본 논문에서는 이 부분에 대해서는 다루지 않았다. 또한 사업적 측면, 사회적(법제도) 측면 등 기술 외적인 부분도 다루지 않았다. 본 논문에서 제안하는 의료정보 통합관리모델을 현장에 적용하기 위해서는 정보의 보안관

리와 기술 외적인 부분에 대한 세부적인 연구도 병행되어야 할 것이다.

5. 참고문헌

[1] 김재필, 최영선, 박희경, 최진욱, HL7 과 MFER 표준을 이용한 원격생체정보 공유 기술 개발, 대한의료정보학회지 제10권 제4호, 2004.  
 [2] Park SH, Bom HS. A Web-based medical information exchange system using HL7 protocols. The Korean Society of Medical Informatics 8(2):11-16, 2002.  
 [3] 장해정, 곽연식, Health Level Seven과 HL7 한국지부 소개, 대한의료정보학회지 제7권 제2호, 2001.  
 [4] Cannataro M, Guzzo A, Pugliese A. Knowledge management and XML: derivation of synthetic views over semi-structured data. ACM SIGAPP Applied Computing Review 10(1):32-36, 2002.  
 [5] Nierman A, Jagadish HV. ProTDB: probabilistic data in XML. Proc. of the 28th VLDB Conference 646-657, 2002.  
 [6] Van HL. A system based on XML for supporting the management of educational websites. ACM SIGAPP Applied Computing Review 10(1):37-42, 2002.  
 [7] Key JS, Wong RK. Structural inference for semi-structured data. Proc. of the 10th International Conference on Information and Knowledge Management 159-166, 2001.  
 [8] Pittas N, Jones AC, Gray WA. Evolution support in large-scale interoperable systems: a metadata driven approach. Proc. of the 12th Australasian Conference on Database Technologies 2001:161-168.  
 [9] Yoo S, Kim B, Park H, Choi J, Chun J. Design and implementation of HL7 based real-time clinical data integration system. Proc. of the METBMS'03 222-227, 2003.