

KEM과 OWL의 바인딩

장병철*, 차재혁**, 함달호***

요약

근래 e-러닝이 활성화됨에 사용자 맞춤형 교육에 필요한 효과적인 e-러닝 콘텐츠 검색의 필요성이 대두되었다. 이에 시맨틱 웹 기술을 이용하여 e-러닝 콘텐츠를 검색하기 위한 다양한 연구가 수행되고 있다. 또한 시맨틱 웹의 핵심 기술인 온톨로지를 e-러닝에 활용하려는 각종 연구가 수행되고 있다. 이에 본 연구에서는 기존의 메타데이터 표현 방식이 가지는 문제점을 해결하여 온톨로지를 이용한 효과적인 검색에 메타데이터를 활용하기 위하여 2004년 12월 국내 최초 e-러닝 관련 기술표준규격(KS X 7001)으로 제정된 KEM(Korea Education Metadata)를 owl로 바인딩하였다. 또한 그 과정에서 발생하는 각종 문제와 해결방안에 대하여 자세하게 논하였다.

The KEM OWL Binding

Byung-Chul Chang*, Jae-Hyuk Cha**, Dall-ho Ham***

Abstract

In the recent year, the need for effective searching for e-learning contents has emerged as a revitalization of e-learning. Therefore, there are many studies for the contents searching method that make use of semantic web technology. Also, there are many studies to apply the ontology as the core technology of semantic web to e-learning. In this study, we bind the KEM(Korea Education Metadata) that enacted as the first technical standard of e-learning, KS X7001, in Dec. 2004 with OWL to use the metadata for a effective searching with ontology through the solving of the problem of metadata. And, we discuss some problems that occur in progress of binding and the solutions.

Keywords : OWL, RDF, KEM

1. 서론

오늘날 전 교육 분야에서 e-러닝이 활성화됨에 따라 효율적, 효과적인 e-러닝 콘텐츠의 생산, 유통, 활용을 위한 국가 수준의 표준화 필요성이 대두되어, 이러한 사회적 요구를 반영하여 한국교육학술정보원(이하 KERIS)은 2004년 12월 국내 최초 e-러닝 관련 기술표준규격(KS X 7001)으로 제정된 KEM(Korea Education Metadata)를 국내 e-러닝 콘텐츠의 개발, 공유,

관리에 있어서 중요한 가이드라인으로 제공하였다[1].

KEM은 현재 고등교육정보 및 저작권 관리를 위한 메타데이터 언어 분야까지 확장, 발전하여 v3.0에 이르고 있으며, 그 활용성은 각 시도교육청 및 각종 교육기관에서 사용해왔던 상이한 운영환경과 통합되지 못한 교육 자원들의 관리 체계를 상호 운용 가능케 하여 체계적인 교육자원들의 운영 체계의 기반을 다지고 있다. 또한 KEM은 국내 환경에 적합한 메타데이터 구조를 제시하여 학습자를 위한 수준별 학습, 적응적 학습 등 학습자의 수요와 역량에 맞는 맞춤형 서비스의 지향점을 보여주고 있다. 그러므로 KEM의 지속적인 연구와 발전 방향성 제시는 국내 e-러닝 산업의 혁신적 향상을 이루기 위한 필수적 과제라 할 수 있다. 그러나, KEM과 같은 기존의 메타데이터를 이용한 키워드 검색 방법은

* 제일저자(First Author) : 장병철

접수일:2006년 06월 02일, 완료일:2006년 06월 30일

* 한양대학교 정보통신대학원 박사 과정

bcchang@hanyang.ac.kr

** 한양대학교 정보통신학과 교수

*** 한양대학교 정보통신학과 석사과정

다음과 같은 사항에서 부족한 면을 보이고 있다 [2].

- 메타데이터 요소간의 상관관계 부재로 인한 검색의 단순화 예를 들어, 학습자가 수학의 '집합' 단원을 학습하기 위해 필요한 선수학을 검색하기 위해서는 개념간 상관관계가 기술되고 추론을 이용한 검색이 가능해야 하는데 기존의 메타데이터 기반에서는 이러한 관계 기술 및 추론이 어렵다.
- 키워드의 시맨틱 부족으로 인한 검색 효율의 저하 기존의 키워드 검색 시스템에 '배' 라는 단어를 입력하면 너무 많은 자료가 검색되며 사용자는 자신이 의도한 내용을 담고 있는 자료를 직접 분류해야하는 문제점이 발생한다. 또한 너무 많은 문서가 검색될 경우 대부분 필요하지 않은 문서임에도 불구하고 정작 필요한 정보를 찾지 못하는 문제가 발생한다.

본 연구는 현재 시맨틱 웹의 기본 기술 중 하나인 온톨로지 기술을 이용하여 초.중등 교육정보 메타데이터 국가 표준 규격인 KEM을 OWL로 바인딩하는 방법과 그 과정에서 발생하는 문제점과 해결 방안을 제시 한다.

2. 관련 연구

2.1 KEM

KEM v3.0은 IEEE 1484.12.1-2002 LOM을 기반으로 개발 되었다[1]. KEM은 총 9개의 category로 되어 있으며 그 내용은 <표 1>과 같다.

2.2 OWL (Web Ontology Language)[3]

OWL 웹 온톨로지 언어는 단지 사람에게 정보를 표시하는데 그치지 않고 정보의 내용을 직접 처리할 수 있는 어플리케이션을 구현하는데 활용될 수 있도록 설계된 언어이다. OWL은 풍부한 어휘(vocabulary)와 형식적 의미론(formal semantics)을 포함하고 있기 때문에 기계 해석이 가능한 웹 콘텐츠를 저작하는데 있어 XML, RDF[4] 및 RDF 스키마(RDF-S)보다 뛰어나다. OWL은 표현력이 서로 다른 세 개의 하위 언어

- OWL Lite, OWL DL, OWL Full - 로 구성되어 있다. 후자로 갈수록 표현력이 더 크다. 또한 각 하위 언어는 서로 다른 개발자 및 사용자층을 대상으로 한다.

<표 1> KEM Category

Category	하위 element
General	Identifier, Title, Sub Title, Language, Description, Table of Contents, Keyword, Coverage, Structure, Aggregation Level
Rights	Cost, Copyright&Otherrestrictions, Rights Holder, Right Issuer Grant, Expiry Date, Description
LifeCycle	Version, Status, Contribute
Relation	Kind, Resource
Meta-metadata	Identifier, Contribute, Metadata Scheme, Language,
Annotation	Entity, Date, Description
Technical	Format, Size, Location, Requirement, InstallationRemarks, OtherPlatformRequirement, Duration
Classification	Purpose, Taxon Path, Description, Keyword
Educational	Interactivity Type, Learning Resource Type, Interactivity Level, Semantic Density, Intended End User Role, Context, Typical Age Range, Difficulty, Typical Learning Time, Description, Language, Pedagogy

- OWL Lite는 클래스 분류 계층과 간단한 제약 사항표현을 필요로 하는 사용자들을 위한 언어이다. 예를 들면, OWL Lite는 관계차수 제약 사항의 표현을 지원하지만, 차수의 값으로 0 또는 1만 사용할 수 있도록 제한한다. OWL Lite를 지원하는 도구를 제작하는 것은 다른 하위 언어를 지원하는 도구를 제작하는 것 보다 상대적으로 쉽다. OWL Lite는 유의어 사전이나 여타 분류

체계의 표현 언어를 빠르고 손쉽게 OWL 화하기 위한 용도로 적합하다. OWL Lite 는 OWL DL보다 이론적 복잡도가 낮다. 이와 관련한 더 자세한 사항은 OWL 참고서의 OWL Lite 설명 부분에 제공되어 있다.

- OWL DL은 계산학적 완전성(Computational Completeness)과 결정가능성(Decidability)을 유지하면서 최대의 표현력을 활용하고자 하는 사용자에게 적합하다. 완전성은 모든 결론이 계산될 수 있다는 특성이고, 결정가능성은 모든 계산이 유한한 시간 안에 끝난다는 특성이다. OWL DL은 OWL의 모든 어휘를 포함하고 있다. 그러나, 어휘의 사용에 있어 정해진 제약 사항을 준수해야 한다. 제약 사항의 한예를 들어 보면, 클래스는 다른 클래스들의 하위 클래스가 될 수 있지만 다른 클래스의 인스턴스가 될 수는 없다. OWL DL은 그 이름으로 짐작할 수 있듯이 기술 논리(Description Logic)에 상당한다. 기술 논리는 OWL의 형식적 기반이 된 논리학의 한 분야이다.
- OWL Full은 계산학적인 어떤 보장 없이 최대의 표현력과 RDF의 유연한 문법을 모두 활용하고자 하는 사용자에게 적합하다. 예를 들면, OWL Full에서 클래스는 개체(Individual)의 집합인 동시에 그 자체가 하나의 개체가 될 수도 있다. OWL Full은 사전 정의된 (RDF 또는 OWL) 어휘의 의미를 확장하는 온톨로지를 작성하도록 허용한다. OWL Full의 모든 기능에 대하여 완전한 추론을 지원하는 추론 소프트웨어는 만들 수 없을 것으로 보인다.

올바른 표현의 범위 및 타당한 결론의 영역에 있어서 OWL DL은 OWL Lite의 확장이고 OWL Full은 OWL DL의 확장이다. 각 하위 언어 간에는 다음과 같은 관계들이 성립된다. 각 관계의 역은 성립되지 않는다.

- 본 연구에서는 OWL-DL을 이용하여 KEM 바인딩을 수행 한다.
- 모든 올바른(legal) OWL Lite 온톨로지는

올바른 OWL DL 온톨로지 이다.

- 모든 올바른(legal) OWL DL 온톨로지는 올바른 OWL Full 온톨로지 이다.
- 모든 타당한(valid) OWL Lite 결론은 타당한 OWL DL 결론이다.
- 모든 타당한(valid) OWL DL 결론은 타당한 OWL Full 결론이다.

3. Binding Principal

3.1 메타데이터 메타모델 (Metadata Metamodels)

KEM과 같은 메타데이터 모델을 특정 형태, 즉 xml을 이용하는 owl로 표현하기 위해서는 먼저 KEM의 메타데이터 메타모델을 알아야 한다. 메타데이터 메타모델은 일종의 개념 스키마 이다.

본 연구에서 사용하는 KEM은 그 메타모델이 다른 것과 다르다. Mikael Nilsson등이 [5]의 연구에서 LOM을 RDF로 바인딩하는 과정에서 언급하고 있듯이, LOM을 기반으로 하는 KEM을 OWL로 바인딩하는 과정에서도 KEM의 메타모델은 다루기 힘든 내용이다. KEM의 메타모델을 예를 들어 설명하자면, Dublin Core(이하 DC)[6]의 메타모델의 경우 DC의 각 요소가 표현하고 자하는 어떤 resource의 property이다. 따라서 DC의 모든 요소는 OWL로 바인딩 할 때, property로 처리하면 쉽게 바인딩 된다. 그러나 2.1에서 살펴보았듯이 KEM의 각 요소는 모두가 어떤 resource를 설명하는 property가 될 수 없다. 예를 들어 <annotation>의 <date> 요소만 해도 어떤 resource의 date가 아니라 그것의 주석을 단 날짜를 표현하는 것이므로 직접적으로 resource의 property라 볼 수 없다.

전술한 DC와는 달리 KEM은 Container 기반 메타모델이기 때문에 DC처럼 쉽게 바인딩할 수 없다. 이 metamodel 문제가 본 연구의 가장 큰 해결 과제이며 궁극적으로 본 연구에서는 KEM의 모든 요소들을 하나씩 고려하면서 바인딩 하는 방법을 취하였다.

3.2 Metadata Frameworks

기본적으로 OWL은 RDF와 동일하게 기본 형

태가 문장(statement)으로 구성되어 있다. 문장은 하나의 resource에 대하여 하나의 property가 하나의 value를 가지도록 되어 있다. 따라서 OWL 한 문장으로는 KEM에서 지칭하는 resource이 모든 내용을 담을 수가 없다. 따라서 KEM OWL Binding은 KEM의 XML 표현과는 달리 self-contained 방식으로 기술하지 않으며 각각의 문장이 network 형태로 유기적으로 연결되어 KEM에서 지칭하는 하나의 resource를 표현하는 방식으로 표현하도록 했다.

3.3 Dublin Core의 표현

많은 수의 KEM 요소들은 의미적으로 Dublin Core의 요소들과 유사하다. 이는 KEM 표준에서 밝히고 있듯이 KEM이 LOM에 기반을 두고 있고 LOM은 Dublin Core 요소들을 대부분 수용하였기 때문이다. 따라서 [5]에서 Dublin Core의 요소들을 수용한 방식과 유사하게 본 연구에서도 Dublin Core의 요소들을 모두 수용하였다. Dublin Core의 요소들과 관련있는 KEM의 요소들을 rdfs:subPropertyOf를 이용하여 Dublin Core 요소의 하위 요소들로 정의 하는 방식으로 구성하였다. 따라서 본 연구의 KEM OWL은 Dublin Core와 완전히 호환된다.

3.4 바인딩 기본 원칙

본 연구의 KEM 온톨로지 바인딩의 기본적인 원칙은 다음과 같다.

- Container로써 지칭되는 General, LifeCycle 등의 정보 모델은 경우에 따라 각각 class 또는 Property로 바인딩 하였다. [1]의 연구에서는 LOM을 바인딩 할 때 container를 모두 네임스페이스로 바인딩 하였는데, 본 연구에서는 container라 할 지라도 충분한 의미를 가지는 것이 있다고 판단하였다. 예를 들어
- Container가 아닌 경우 독립적인 개념으로 존재할 수 있는 것은 클래스로 그렇지 않으면 프로퍼티로 구축한다.
- KEM 정보 모델이 표현하는 교육컨텐츠는 LearningResource(이하 LR) 라는 클래스로 정의하며 개별 컨텐츠는 LR 클래스의 instance로서 간주하도록 한다.

- Class 명명은 첫 문자를 대문자로 한다.
- Property의 명명은 ObjectProperty 인 경우 has~ 로 시작하며, DatatypeProperty인 경우 첫문자를 소문자로 한다.
- KEM 에서는 정보 모델마다 정보의 최대 크기를 지정하나, 바인딩 과정에서는 이것을 생략한다.
- owl-DL의 기능을 최대한로 이용하여 각 클래스에 owl:oneOf 및 각종 cardinality 관련 서술어를 사용하여 KEM에서 표현하는 의미를 최대한 표현 하였다.

상기 기본 원칙 중 마지막이 [5]의 연구와 가장 대별되는 본 연구의 특징 중 하나이다. RDF를 이용한 LOM의 바인딩에서는 RDF 자체가 가지는 표현력의 한계로 인하여 각종 cardinality를 표현할 수 없다. 따라서 LOM 이 가지는 각종 제한 요소를 표현할 수가 없었으며, 정확한 LOM 모델링을 할 수 없었다. 그러나 OWL-DL은 풍부한 표현력을 가지고 있으므로 KEM이 가지는 각종 제한 요소를 그대로 표현할 수 있다. 따라서 본 연구에서는 OWL-DL의 표현력을 최대한 이용하여 KEM이 담고 있는 내용을 그대로 표현하였다.

3.5 Langstring

KEM 정보 모델을 OWL을 활용한 온톨로지로 변환하는 과정에 있어, KEM 정보 모델의 실제 값(value)을 변환해야 하는 경우가 발생한다. 일반적인 경우 정보 모델의 값을 의미하는 명칭을 태그로 작성하여 그 태그 사이에 값을 입력하면 되나, 그 값의 의미를 태그로서 선언하기 어려우며, 일반 명사를 지칭하는 경우 langstring 속성을 선언하여 활용할 수 있도록 한다.

langstring 속성을 활용하여 KEM 정보 모델을 지칭하는 데이터는 다음과 같은 유형이 있다.

- Identifier
- Restricted or Best-Practice Vocabulary
- Vocabulary Source

4. The KEM OWL Binding

본 절에서는 KEM의 각각의 요소들을 실제 OWL로 encoding한 결과를 설명한다. 그러나 2.1에서 전술한바와 같이 KEM 요소가 9개에 이르고 그 하부요소가 많은 관계로 모든 내용을 설명하지 않으며 모든 내용의 Binding 결과는 참고문헌 [2]에 설명되어 있다.

4.1 MetametaData

해당 교육 콘텐츠보다 메타데이터 기록자체에 대한 특정의 정보를 기술한다. 해당 내역으로는 메타데이터를 기록한 사람, 기록한 방법, 기록 일자, 참조 자료에 대한 정보를 설명한다. 다음은 MetaMetaData중에서 Identifier에 대한 바인딩 방법에 대하여 설명한다.

4.2 Contribute

자원의 발전(evolution)과정(개발, 편집, 출판 과정을 포함)에서 영향을 끼친 사람 혹은 조직들을 설명한다. 단 추후 언급될 MetaData의 Contribute와는 그 의미가 다르다.

본 항목에서는 콘텐츠의 발전 과정에 기여한 종류와 그에 참여한 사람과 날짜를 모두 기술해야 한다. 앞서 언급한 많은 내역이 해당 정보 모델은 하나의 값 혹은 하나의 값과 그에 대한 근원(source)에 대해서 기술하였으나, 본 내역에서는 해당 정보 모델 즉 자원의 발전 과정에 기여한 종류(Role)에 기여한 복수의 사람을 기술하는 방법을 표현하여야 한다. 본 연구의 KEM Core 온톨로지에서 완전히 구현하지 않았으며, 현재 단계에서는 contribution instance의 여러 기술 내역으로 Role과 Person의 instance가 참조한다.

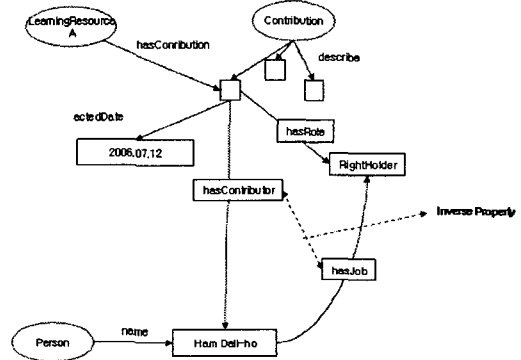
4.3 Identifier

본 연구에서 메타데이터 식별자를 위한 카탈로그는 현재 본 Identifier는 KEM에서 활용하는 목록 생성시스템이 일반적으로 생성하는 값으로, 현재의 단계에서는 KerisMetadataEntry 1개라고 가정하고 클래스를 생성하였다.

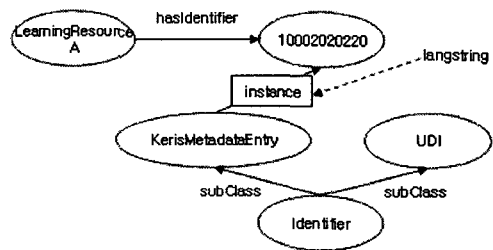
4.4 Classifications

Classifications 요소는 OWL로 바인딩하기 가

장 어려운 요소이다. Classifications 요소 내에 taxon 요소가 있는데, 이는 계층적 요소이기는 하지만 계층간에 중복성도 많고 또 KEM에서 지원하는 taxon에도 제한이 없기 때문이다. taxon요소를 binding하는 방법은 여러 가지가 있을 수 있으나 크게, 각각에 존재하는 taxon을 새로운 온톨로지 로 바인딩 하고 그것을 KEM과 연결하는 방법과 taxon을 property로 취급하는 방법이 있다. 또한 taxon정보 자체를 String으로 처리하는 방법이 있는데, 본 연구에서는 구현상의 편의를 위하여 마지막의 방법을 취하였다.



(그림 1) hasContribution property의 구성도



(그림 2) Identifier 요소의 클래스 관계도



(그림 3) taxon element 표현 방법

4.5 Vocabulary

Vocabulary는 여러 가지 방법으로 바인딩 될 수 있다. [5]의 연구에서는 Vocabulary를 namespace 또는 property 활용하는 여러 가지 방법을

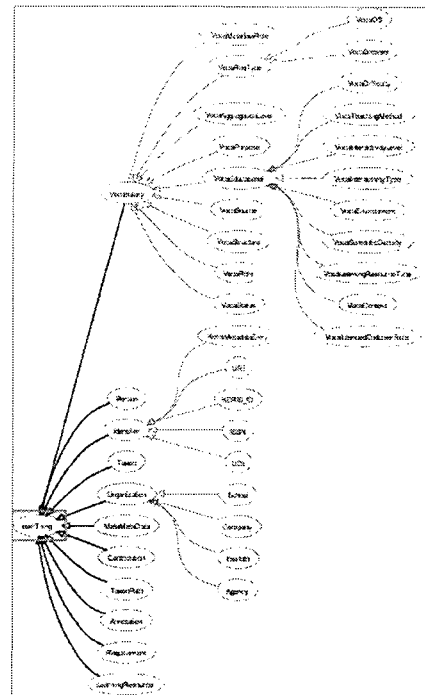
이용하였지만, 본 연구에서는 Vocabulary클래스를 독립적으로 구성하고 Vocabulary를 사용하는 다른 요소에서 Vocabulary 클래스의 내용들을 value이용하는 방법을 취하였다. 다음 그림은 본 연구에서 구축한 Vocabulary 클래스의 구성도이다. 이렇게 Vocabulary를 구성하면 OWL 인스턴스 생성시에 다소간의 자료 입력의 번거로움이 발생할 수도 있으나 KEM에서 사용하고 있는 모든 Vocabulary를 표현할 수 있을 뿐 아니라, owl:oneOf 같은 property를 사용하지 않아도 되기 때문에 OWL로 바인딩된 KEM 온틀로지를 추론하는데도 용의하기 때문에 본 연구에서는 Vocabulary를 이와 같이 별도의 클래스로 구성하였다.



(그림 4) Vocabulary 클래스 구성도

5. 결론

XML을 기반으로 한 KEM이 데이터를 표현하는 방식과 OWL이 데이터를 표현하는 방식 간에 기본적으로 많은 차이가 있어 KEM을 OWL로 Binding하는데는 많은 어려움이 있었다. 본 연구에서는 기본적으로 KEM의 최대한 그대로 OWL로 Binding하는데 중점을 두었다. 그러나 OWL로 바인딩하는 것의 궁극적인 이유가 OWL로 구축된 정보를 이용하여 추론을 통한 검색효과를 향상 시키기 위한 것이므로 현재의 추론 엔진에 적합하게 KEM OWL정보를 수정하는 작업이 추가로 요구된다. 예를 들어 owl:OneOf등을 이용하여 기술한 Cardinality 정보는 KEM의 내용을 정확하게 OWL로 Binding하였다는 점에서는 의의가 있겠으나 실제 이를 처리할 수 있는 엔진이 극히 드물기 때문에 추론 엔진을 고려한 owl binding에 대한 연구가 필요하다.



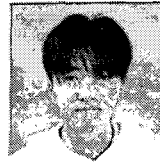
(그림 5) KEM OWL 클래스 계층도

참고 문헌

[1] 손진곤 외, 고등교육정보 및 저작권 관리를 위한 메타 데이터(KEM v3.0)연구, 연구보고 KR-2005-27, 한국교육학술정보원, 2005.

- [2] 차재혁 외, KEM 고도화를 위한 시맨틱웹 기반 온톨로지 연구, 한국교육학술정보원, 2006, 출간예정.
- [3] OWL Overview, <http://www.w3.org/TR/owl-features/>
- [4] RDF Primer, <http://www.w3.org/TR/rdf-primer/>
- [5] The RDF binding of LOM, <http://kmr.nada.kth.se/el/ims/metadata.html>
- [6] The Dublin Core Metadata Initiative, <http://dublincore.org/>
- [7] IMS Global Learning Consortium <http://www.imsproject.org>
- [8] IMS Resource Description Framework(RDF) Bindings, <http://www.imsproject.org/rdf/>
- [9] IMS Learning Resource Meta-data Specification, <http://www.imsproject.org/metadata/index.cfm>

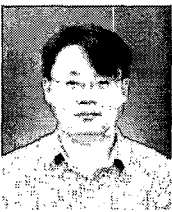
함 달 호



2006년: 한양대학교
컴퓨터교육과(학사)
2006년~현재: 한양대학교
정보통신학과 석사과정

관심분야 : e-Learning, 시맨틱웹, 웹 서비스

장 병 철



1996년 안동대학교
컴퓨터공학과(학사)
2001년 한양대학교
컴퓨터교육과(석사)
2003년~현재 한양대학교
정보통신학과 박사과정

관심분야 : 데이터마이닝, 시맨틱웹, e-Learning

차 재 혁



1987년 서울대학교
계산통계학과(학사)
1991년 서울대학교
컴퓨터공학과(석사)
1997년 서울대학교
컴퓨터공학과(박사)

1997년~1998년: 한국학술진흥재단 부설
첨단학술정보센터선임연구원

1998년~2001년: 한양대학교 사범대학
컴퓨터교육과 조교수

2001년~현재: 한양대학교 정보통신대학
정보통신학부 교수

관심분야 : XML, 데이터베이스, 플래시 메모리 기반
저장시스템, 멀티미디어 콘텐츠 적용화,
e-Learning