

# 웹학술정보 서비스를 위한 데이터베이스 및 인터페이스 모델링에 관한 연구

## A Study on the Database & Interface Modeling for the Information Services of the Scientific Resources on the Web

한상우, 공진현 성균관대학교 대학원 문헌정보학과

Han, Sang Woo & Kong, Jin Hyun

Dept. of Library and Information Science, Graduate School of Sung Kyun Kwan University

요약 : 현재 인터넷 상에는 많은 정보자원이 존재하고 있다. 이처럼 많은 정보자원 중 학술적인 가치가 있는 정보에 대한 요구가 증가하고 있다. 이러한 요구를 바탕으로 웹학술정보에 대한 메타데이터 데이터베이스를 구축함으로써 원정보에 대한 정확한 정보를 알고 웹상의 정보자원을 효율적으로 사용할 수 있다. 본 논문은 현재 진행하고 있는 IFL(Internet Faculty Library, <http://www.ifl.or.kr>)을 사례로 웹학술정보의 메타데이터 데이터베이스 구축을 위한 데이터베이스 및 인터페이스의 설계 및 모델링하는 과정을 제시하도록 하겠다.

### 1. 서론

현재 인터넷 상에는 많은 정보자원들이 존재하고 있다. 이처럼 많은 정보자원 중 학술적인 가치가 있는 정보에 대한 요구가 증가하고 있다. 이러한 요구를 바탕으로 웹학술정보에 대한 메타데이터 데이터베이스를 구축하고 제공함으로써 웹상의 정보를 보다 효율적으로 사용할 수 있다. 본 논문은 현재 성균관대학교 정보관리연구소에서 진행하고 있는 IFL(Internet Faculty Library)을 사례로 하여 웹학술정보의 메타데이터 데이터베이스 구축을 위한 데이터베이스 및 인터페이스의 설계 및 모델링 과정을 제시하도록 하겠다.

### 2. 데이터베이스 모델링

데이터베이스 시스템에서 기본적인 사항은 자원을 보다 효율적으로 관리하고 자원을 공유하는데 그 목적이 있다. 데이터베이스의 모델링은 이러한 목적에 부합할 수 있는 형태로 모델링되어야 하며 이용 및 관리가 편리하면서도 효율적으로 정보를 이용할 수 있는 모델링이 필요하다.

#### 2.1. 관계형 데이터베이스

현재의 많은 데이터베이스 시스템은 관계형 데이터베이스 시스템(RDBMS)을 사용하고 있다. 관계형 데이터베이스 시스템은 다음과 같은 장점을 가지고 있다.

첫째, 관계형 데이터베이스는 만들거나 이용

하기가 쉽지만, 무엇보다 확장이 용이하다.

둘째, 관계형 데이터베이스를 구축할 때 데이터가 가질 수 있는 값의 범위나 제한사항 등을 정의하여 정확한 데이터를 가질 수 있게 한다.

셋째, 관계형 데이터베이스 시스템은 일련의 정형화된 테이블 내에 데이터들의 저장되어 있는 형태이므로, 테이블을 재구성하지 않아도 원하는 정보에 접근할 수 있으며, 서로 관계되어 있는 정보의 참조가 용이하다.

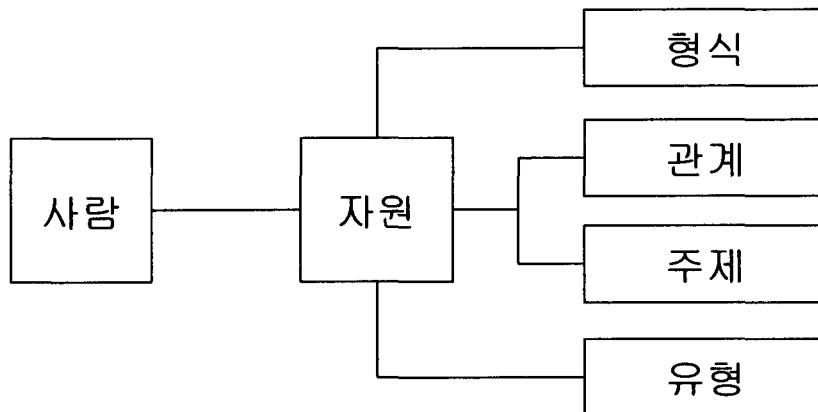
이러한 장점들로 인해 관계형 데이터베이스를 사용하게 되었다.

지막으로 각 속성(Attribute)으로 제대로 표현된 엔터티(Entity)를 관계(Relationship) 지어야 한다. 관계(Relationship)는 각 엔터티(Entity)를 정확하게 맺어주는 역할을 하며, 데이터베이스를 사용하고 유지하는데 있어 질의(Query)를 수행하는데 있어 중요한 역할을 한다.

이러한 개념적인 부분을 통한 데이터베이스의 개념적인 설계는 [그림 1]과 같다.

### 2.3. 데이터베이스 모델링

위에서 진행한 개념적인 설계를 바탕으로 본



[그림 1] 데이터베이스 개념 설계

### 2.2. 데이터베이스의 개념적 설계

본 논문에서 사례로 하는 IFL은 웹상의 학술 정보에 대한 메타데이터 데이터베이스의 구축을 기본으로 하고 있다. IFL의 데이터베이스를 구축하는데 가장 먼저 고려한 사항은 엔터티(Entity)와 속성(Attribute), 관계(Relationship)의 정의이다.

엔터티(Entity)는 각 속성들의 집합으로 각 속성간에 관련되어 있는 항목들이 가장 기초적이고 명확하게 표현되어야 하고, 데이터베이스의 모델링에서 가장 기반이 되는 것이다. 또한 속성(Attribute)은 엔터티(Entity)의 특성을 제대로 표현할 수 있는 요소를 선정해야 한다. 마

격적인 데이터베이스의 모델링을 진행하였다. IFL 데이터베이스의 엔터티(Entity)는 자원, 주제, 유형, 저자 등으로 크게 분류되었으며, 각 속성(Attribute)은 각 메타데이터 요소를 매핑시켜 각 엔터티(Entity)를 정확하게 표현할 수 있도록 하였다. 또한 각 엔터티(Entity)간의 관계를 명확하게 파악하여 관계(Relationship)는 1 : M 또는 M : N 로 관계지어 관련되어 있는 속성들이 질의(Query)를 통하여 정확하게 표현될 수 있도록 하였다.

일례로 IFL의 데이터베이스 모델링에서 가장 문제가 되었던 부분은 개념설계 부분에서 볼 수 있는 "사람" 항목이다. 현재 웹상의 자원에서 가장 큰 문제점은 컨텐츠의 복사가 용이하

다는 것이다. 따라서 사용자가 보고 있는 자원의 원저자 및 다른 형태로의 자원의 가공자, 현 자원이 작성되는 데에 기여한 기여자 등의 역할 부분이 가장 어려운 부분이다. 또한 동명이인이 있으므로 이미 등록되어 있는 인물과 새로 등록되어야 하는 인물의 구분도 어려운 부분이다. 이러한 문제점을 해결하기 위한 방안으로 IFL에서는 Person, Role, Type, Work이라는 4개의 테이블을 이용하여 자원에 대한 사람의 역할, 유형, 그리고 동명이인의 문제는 E-Mail의 고유성을 이용하여 중복을 최소화하도록 설계하였다. 또한 Work이라는 테이블을 이용하여 Person과 Resource가 M : N의 관계를 갖도록 함으로써 하나의 자원에 여러 사람이 참여하는 부분에 대한 문제를 해결하였다.

앞의 연구에서 진행되었던 메타데이터의 요소 추출등의 연구를 통한 모델링의 결과로 [그림 2]와 같은 ERD(Entity Relationship Diagram)를 구성하게 되었다.

### 3. 인터페이스 모델링

정보시스템에서의 인터페이스는 이용자의 만

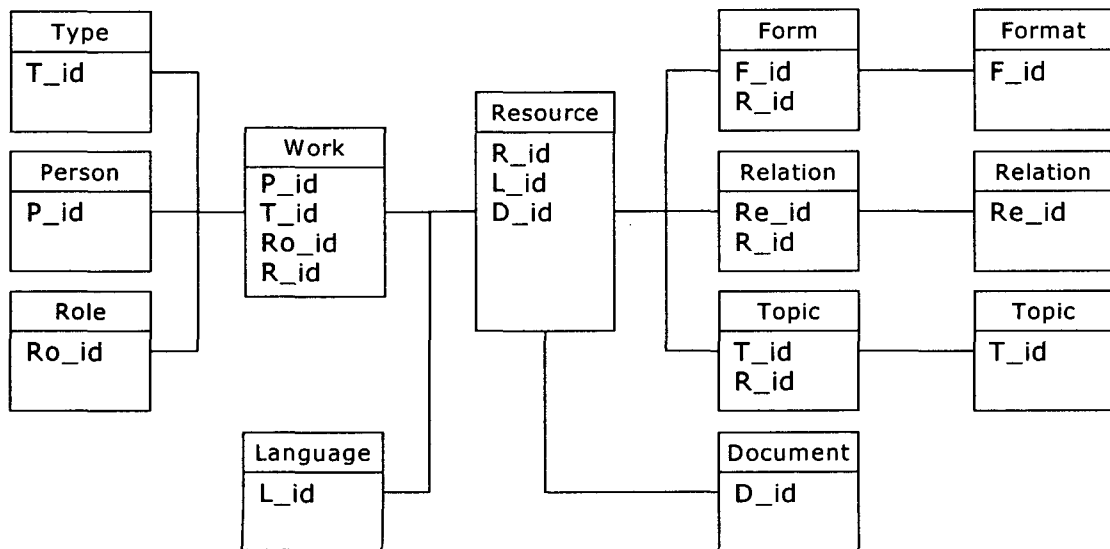
족도 증가와 사용의 용이성 때문에 지속적인 관심의 대상이 되어왔다. 인터넷의 등장과 정보기술의 발전으로 인터페이스도 점차 세분화, 다양화되고 있다. 따라서 본 논문의 사례인 IFL의 경우에도 웹학술정보 서비스에 적합한 인터페이스의 모델링이 필요하다고 하겠다. 이에 크게 정보탐색 관련 인터페이스와 검색결과 인터페이스로 나누어 IFL의 인터페이스 모델링을 살펴보도록 하겠다.

#### 3.1 정보탐색 관련 인터페이스

정보탐색 관련 인터페이스는 키워드 탐색과 브라우징 탐색으로 나누어 살펴보도록 하겠다.

##### 3.1.1 키워드 탐색

IFL은 대상이 되는 정보의 종류와 이용자 계층에 있어서 검색엔진을 이용한 일반적인 검색 서비스와는 다르다. 검색서비스의 경우 일반 이용자가 웹상의 모든 정보를 대상으로 무차별적인 수집을 하지만, IFL의 경우 해당 학문분야의 연구자를 중심으로 학술정보만을 그 대상으로 하고 있다. 이는 이용자가 탐색 과정에서

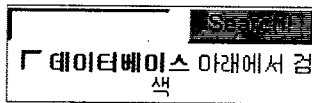


[그림 2] 데이터베이스 ERD

질의어 선정의 정확성을 높일 수 있는 요인이다. 이에 따라 IFL의 키워드 탐색 인터페이스는 [그림 3]과 같이 초기 화면에서 '주제'와 '제목' 필드 중에 하나를 선택하여 검색하도록 하였다. 주제와 제목은 이용자가 가장 접근하기 쉬운 필드이기 때문이다. 또한 검색 결과 화면에서 결과내 재검색 인터페이스 [그림 4]를 제공하여 많은 양의 정보가 검색되었을 경우 이를 다시 필터링할 수 있도록 하였다.

주제 제목 | 검색

[그림 3] 키워드 탐색 화면



[그림 4] 결과 내 재검색

### 3.1.2 브라우징 탐색

IFL은 브라우징 탐색을 주된 탐색 인터페이스로 제공하고 있다. 그 이유는 IFL이 한정된 학문분야를 대상으로 하고 있으며 기존에 비교적 명확한 학문분류 체계들이 만들어져 있고 브라우징 과정을 통하여 잠재적 정보요구도 해결할 수 있는 장점이 있기 때문이다. 특히 주이용자층인 연구자의 경우 연구의 시작 단계에서의 잠재적 정보요구를 해결하는 것은 매우 중요하다. 브라우징 인터페이스는 주제 분류와 유형 분류의 두 가지 유형으로 나누어져 있다. 주제 분류는 DDC, KDC, 학진학문분류표 등의 기존의 분류 체계와 인터넷이라는 매체의 특성을 고려하여 관련 분야의 포털사이트의 분류 체계를 분석, 매핑하여 새로운 분류 체계를 만들었다. 새로운 분류 체계는 3가지의 단계를 가지게 된다. 예를 들어 '문헌정보학 → 정보조직

→ 분류' 와 같이 3단계를 거치면 된다. 이러한 분류 체계는 브라우징 탐색 인터페이스에서 상위항목에서 하위항목으로 이동하는 방식으로 구성되어 있다[그림 5].

한편, 유형 분류는 동일한 데이터에 대하여 위의 주제 분류와 함께 각각의 데이터를 유형에 따라 분류함으로써 이용자의 편의를 돕도록 하였다. 예를 '정보정책'이라는 주제로 분류된 데이터를 유형 분류에서는 '논문', '강의계획서' 등으로 분류하는 것이다. 이는 주제 분류만으로는 구별할 수 없는 각 데이터의 유형을 분류해 제공함으로써 이용자의 편의를 높이는 인터페이스이다.

#### 정보학 일반

정보학 이론, 정보사회, 정보정책, 지식관리, 정보커뮤니케이션

#### 정보검색

데이터베이스, 검색모형/기법, 자동색인/자동요약, 자동분류/클러스터링

#### 정보서비스/이용자연구

이용자연구, 이용자교육, 주제별 정보원관리, 정보서비스, 상호대차

#### 정보공학/시스템

네트워크, 정보시스템(론), 이용자 인터페이스, 프로그래밍

#### 디지털도서관

도서관자동화/전산화, 디지털도서관 정책, 디지털도서관 관련기술, 디지털도서관 구축사례

#### 정보교육

독서지도, 독서실태, 정보교육

#### 저작권

저작권법, 디지털자료관련 저작권

[그림 5] 브라우징 탐색 화면

### 3.2 검색결과 인터페이스

IFL의 검색결과 인터페이스는 크게 간략 정보 인터페이스와 상세 정보 인터페이스로 나뉘게 된다. 간략 정보 인터페이스는 다량의 결과가 출력되었을 때 저장하고 있는 메타데이터의

정보검색 >> 데이터베이스

■ 국가자료목차정보데이터베이스 구축 추진 현황 및 계획 [상세정보] [URL 바로가기]

URL : [http://www.nl.go.kr/pds/research\\_data/text/newsletter/200106/data2\\_200106.htm](http://www.nl.go.kr/pds/research_data/text/newsletter/200106/data2_200106.htm)  
 간략설명 : 국립중앙도서관의 도서관계 2001년 6월호에 실린 자료이다.  
 주제 : 데이터베이스 I  
 자료유형 : 정책자료  
 언어 : 한국어

■ 국내 인터넷 법률정보 데이터베이스의 이용자 인터페이스에 관한 연구 [상세정보] [URL 바로가기]

URL : <http://www.nanet.go.kr/whatsnew/library/file/nal99083.htm>  
 간략설명 : 국회도서관보 제36권 4호에 실린 논문으로 국내 인터넷 법률정보 데이터베이스에서의 이용자 인터페이스에 대한 설명을 담고 있다.  
 주제 : 이용자 인터페이스 I 데이터베이스 I  
 자료유형 : 일반논문  
 언어 : 한국어

[그림 6] 간략정보 화면

전체를 보여주는 것보다는 중요한 일부 정보만을 보여주어 이용자가 쉽게 다량의 정보를 검색할 수 있도록 한다. 현재 간략 정보 인터페이스에서 보여지는 항목은 '제목', 'URL', '간략 설명', '주제', '자료유형', '언어'의 6항목이다 [그림 6]. 각 데이터마다 위의 5항목을 나타내주는 간략 정보는 한 화면에 10개의 데이터를 보여주고 있다. 한편 모든 항목을 전부 보여주는 상세 정보 인터페이스는 Pop-up 창으로 제시하고 있다. 그 이유는 전체 간략 정보 화면과 선택한 상세 정보 화면을 동시에 볼 수 있고, 여러 상세 정보를 동시에 띄워놓고 비교할 수 있는 장점이 있기 때문이다. Pop-up 창의 크기를 고정시켜 놓음으로써 다른 결과 화면을 동시에 볼 수 있다 [그림 7]. 간략 정보 인터페이스와 상세 정보 인터페이스에서 나타나는 메타데이터 항목은 [표 1]과 같이 정리할 수 있다.

■ 제목	누리미디어 한국학 DB시리즈		
■ 부제목			
■ 작성자	누리미디어 [Detail]		
■ 주제	한국서지 I 데이터베이스 I		
■ Keyword	한국학, DB		
■ 간략설명	한국학에 관련된 사전, 도감, 문헌등의 DB를 검색하여 볼수 있는 서비스를 제공하는 홈페이지이다.		
■ 목차			
■ 발행자	누리미디어 [Detail]		
■ 기여자			
■ Date	작성일	발행일	최종수정일
■ 자료유형	기업 홈페이지		
■ 자료형식	HTML(.html) I		
■ URL	<a href="http://www.krpia.co.kr/">http://www.krpia.co.kr/</a>		
■ 대체URL			
■ 원자원정보			
■ 원정보처자			
■ 언어	한국어		
■ 관계정보	Krpia는 현재 대학 도서관 등의 기관에서만 사용할 수 있다. 따라서 개인이 Krpia 서비스에 들어오면 현재 서비스하고 있는 내용 가운데 생를 데이터만 볼 수 있다. 만약 전체 사용을 원하시면 소속 학교나 도서관에 구입 신청을 하셔서 해당 기관에서 이용하면 된다.		
■ 평가정보			

[그림 7] 상세정보 화면

구분	항목
간략 정보 인터페이스	제목, URL, 간략설명, 주제, 자료유형, 언어 (6개 항목)
상세 정보 인터페이스	제목, 부제목, 작성자, 주제, 키워드, 간략설명, 목차, 발행자, 기여자, 날짜, 자료유형, 자료형식, URL, 대체 URL, 원자원정보, 원정보저자, 언어, 관계정보, 이용조건, 평가정보 (20개 항목)

[표 1] 간략/상세 정보 인터페이스 항목

#### 4. 결론

IFL을 사례로 데이터베이스 및 인터페이스 모델링에 대한 부분을 살펴보았다.

데이터베이스 모델링의 경우 IFL에서 제공하는 메타데이터를 효율적으로 관리하고 서비스할 수 있는 관계형 데이터베이스를 사용함으로써 데이터베이스의 활용 및 유지에 적합하도록 하였다. 향후 추가적으로 개발되어야 할 부분은 웹상의 정보라는 특성상 발생할 수 있는 추가요소에 대한 확장성을 고려하여야 할 것이다. 또한 자원의 수가 증가함에 따라 데이터베이스의 효율적인 관리를 위한 관리체계와 운용계획에 대한 자세한 연구가 필요할 것이다.

인터페이스의 경우 향후 키워드 탐색의 기능을 향상시킬 필요가 있다. 현재 제공되고 있는 기본적인 기능 이외에 보다 다양한 검색 필드를 제공해 주어야 한다. 또한 다양한 이용자 수준을 고려하여 여러 수준의 키워드 탐색 인터페이스를 제공할 필요성이 있다. 이외에 탐색 과정이나 결과에서 이용자와 피드백을 할 수 있는 인터페이스가 추가되어야 할 것이다.

#### 참고문헌

- 김남호, 이봉형, 김일남, 황의천, 왕수현, "인터넷을 통한 연구센터 데이터베이스 시스템 구축", 연구보고서, Vol.2 No.1, 1997.
- 김미현, "학술정보시스템의 온라인 인터페이스에 관한 연구 : 적응형 인터페이스를 위한 정보조직 및 활용", 한국문헌정보학회지 32(4) pp. 259-276, 1998.
- 김형래, 임기영, "관계형 데이터베이스 관리시스템을 사용한 지식베이스시스템 설계에 관한 연구", 논문집, 1991.
- 서은경, "정보검색시스템에서의 이용자 인터페이스 기능에 관한 분석적 고찰", 정보관리학회지 16(4) pp. 125-150, 1999.
- 주경희, 김홍식, 서재현, "웹과 DB상의 효율적인 정보검색을 위한 데이터 통합 시스템 설계", 자연과학 Vol.4 No.1, 2000.
- 코리아인터넷 닷 컴 (<http://korea.internet.com>)
- Fox et al., "Users, user interface, and objects : envision, a digital library", Journal of the American Society for Information Science 44(8) pp. 480-491, 1993.
- Hendry, D.G. and Harper, D.J., "An information-seeking environment", Journal of American Society for Information Science 48(11) pp. 1036-1048, 1997.
- Peter Rob, Carlos Coronel, "Database Systems - Design, Implementation, and Management, Thomson Learning", 2000.