

# 인터넷 정보서비스를 위한 메타데이터 프레임워크 개발에 관한 연구\*

## A Study on Development in Metadata Framework for Internet Information Service

황상규(Sang-Kyu Hwang)\*\*, 윤세진(Se-Jin Yun)\*\*\*, 오경목(Kyung-Mook Oh)\*\*\*\*

### 초 록

웹 정보체계에서 여러 가지 다양한 종류의 정보자원을 보다 효과적으로 관리하기 위하여, 인터넷 정보를 기술하고 있는 서로 다른 메타데이터 표준들간의 상호운용성 문제가 점차 중요한 비중을 차지하게 된다. 이러한 문제점을 해결하기 위하여 IFLA의 FRBR, INDECS, ABC 모델 등의 상호연동 모델들을 비교·분석함으로써, 다양한 형태의 메타데이터간의 상호연동을 위한 새로운 방식의 상호운용성 모델을 개발, 제안하였다. 리소스모델의 기본 메타데이터 항목 요소들로부터 중심이 되는 개체(core entity)와 이와 연관된 주요 사건(core event)들을 식별하고, 이를 리소스중심 중간모델과 이벤트중심 중간모델로 변환(swapping)하는 과정을, 구체적인 방법을 제시함으로써, 이들 모델간 전환 방법을 명확하게 제시하였다.

### ABSTRACT

The Dublin Core, MARC, IFLA-FRBR user communities are developing international standards for describing textual, physical and audiovisual resources to enable their resources discovery over the Internet. Therefore the metadata interoperability problem has been exacerbated by the need for more complex metadata descriptions. In this paper we propose a new mechanism for metadata interoperability based on the new semantic web applications: IFLA-FRBR framework, INDECS metadata and an event-aware ABC models. This study introduces a new approach method which is essential to the generation of interoperable metadata descriptions, particularly in the context of multimedia contents.

**키워드** : 메타데이터, 상호운용성, 프레임워크, 상호연동, 자원발견, metadata, crosswalk, interoperability, framework, resource discovery

\* 본연구는 과학기술부 지원 '01 숙명여자대학교 연구기반 확충사업(01-N6-02-01-A-01)에 의하여 지원되었음.

\*\* 국방연구원 연구원(kid4@naver.com)

\*\*\* 엔피아 시스템즈 기술연구소 연구원(ysj01@enpia.co.kr)

\*\*\*\* 숙명여자대학교 정보과학부 문헌정보학 전공 조교수(kmoh@sookmyung.ac.kr)

■ 논문 접수일 : 2002년 5월 20일

■ 게재 확정일 : 2002년 6월 4일

## 1 서 론

월드 와이드 웹의 창시자인 Tim Berners-Lee(1998)는 1998년 데이터 정보체계 수준에 머물고 있는 기존 웹 정보체계의 한계를 지적하고, 지식(knowledge)수준으로 발전해 가는 새로운 웹 정보체계인 'Semantic Web'이라는 새로운 개념을 제안하였다. 'Semantic Web'의 가장 중요한 개념은 기존의 기계가독형(machine-readable) 웹 정보체계에서 컴퓨터가 스스로 의미를 이해할 수 있도록 디자인된 정의와 규칙들을 추가(machine-understandable)한 정보체계로의 전환을 구현목표로 한다. 이는 사람이 직접 작업명령을 지시할 필요 없이 컴퓨터가 수행 가능한 언어로 웹 상에 존재하는 정보를 해석하여, 스스로 필요로 하는 행동을 수행하게 된다.

웹 정보체계의 발전과정상 새로운 웹 정보체계인 'Semantic Web'은 기존의 웹 정보체계에 있어서 상당부분 큰 변화를 가져오리라 예상된다. 가장 큰 변화가 예상되는 분야로, 현재 한계에 이른 인터넷 정보검색기술을 예로 들 수 있다. 현재의 웹 정보검색기술은 크게 수작업 방식의 디렉토리 검색과 검색로봇을 이용하는 키워드 검색으로 구분될 수 있다. 전자는 현재 웹 정보의 증가속도를 수작업 방식으로는 더 이상 감당할 수 없다는 문제점을, 후자는 낮은 검색 정확률이라는 문제점을 가지고 있다. 그러나, 디렉토리 검색의 경우, 검색시스템을 개발하는 집단이

직접 인터넷상의 정보를 수집, 분류해야 하지만, 새로운 'Semantic Web'에서는 컴퓨터가 스스로 의미를 이해할 수 있도록 디자인된 정의와 규칙을 토대로 콘텐츠 생산자가 정보검색을 위한 기본적인 정보를 구축하게 된다. 콘텐츠 생산자가 정보검색을 위한 기본적인 정보를 구축하기 위한 도구로는 XML(Tim Bray et al. 2000)과 RDF(Lassila and Swick 1999)가 있으며, 이러한 규칙에 따라 웹 문서가 만들어짐으로서 웹 로봇이 검색대상인 웹 정보자원에 대한 정보를 스스로 이해(machine-understandable)할 수 방법을 제시할 수 있다.

이러한 의미정보의 표현규칙이 보다 구체적이고 널리 적용될 때, 웹 로봇은 수집한 정보자원에 대한 명확한 주제정보, 유통기한(제작일과 배포일도 구분할 수 있다), 심지어 "누가, 무엇을 가지고, 무슨 목적으로 만들었다"라는 보다 구체적인 사실까지도 알아낼 수 있다. 이는 현재 전자상거래를 포함하여 인터넷 정보자원 활용에 있어 문제시 되어왔던 여러가지 문제점을 동시에 해소할 수 있는 하나의 방안을 제시하게 된다. 이와 관련된 추가적인 연구사례로는 RDF기반 'ontology' 구축방법인 'DAML+OIL'(Ian, Frank and Peter, 2001), RDF기반 'ontology'와 규칙을 통한 추론의 방법(Guha et al. 1998) 등이 있다. 그러나 아직까지 'Semantic Web'에 관한 연구는 초기단계이며, 앞으로 'Semantic Web'이 널리 실용화되기

위해서는, 그 전제조건으로 전세계적으로 상이한 웹 정보 메타데이터간 의미정보 전달방법과 용어의 표준화에 대한 보다 구체적인 연구가 선행되어야 한다.

전통적인 의미에서 메타데이터란 '데이터를 위한 구조화된 데이터'로서 주로 도서관에서 책을 찾기 위한 목록의 역할을 하였으나, 웹 정보체계에 있어 메타데이터란 기계가 이해할 수 있는 의미태깅이다. 웹 정보체계에서 여러 가지 다양한 종류의 메타데이터가 표준화된 의미정보 전달을 위한 수단으로 사용되기 위해서는, 현존하는 수많은 메타데이터 표준안을 통합, 운영하기 위한 하나의 표준모델의 형태로써 메타데이터 프레임워크 혹은 메타데이터 스키마를 기반으로 서로 다른 메타데이터간의 상호운용성(interoperability)이 보장되어야 한다. 메타데이터 프레임워크 혹은 메타데이터 스키마란 해당 메타데이터의 전반적인 논리 구조, 또는 그러한 구조를 기술한 명세서라 볼 수 있는데, 대상 메타데이터의 체계개념을 명시화한다는 점에서 단순히 메타데이터 표준모델을 제시하는 것과 구분된다. 이러한 메타데이터 프레임워크모델은 일종의 대상개념을 추상화하는 데이터 모델링 기법을 제공하며, 이는 각기 서로 다른 메타데이터 표준의 구성요소(element set)들을 대상으로 통합된 메타데이터 항목을 기술할 수 있게 해 줌으로써 상호운용성(interoperability)에 있어 기본토대를 제공할 수 있다.

이러한 메타데이터 프레임워크모델의 종류로는 크게 전통적인 서지정보 체계상에서 대상 리소스를 중심으로 관련 속성들을 기술하는 리소스중심의 메타데이터 모델과 웹 정보체계상에서 사건(event)을 중심으로 메타데이터 요소들을 기술하는 이벤트중심 모델로 구분할 수 있다. 이는 전통적인 서지정보 체계의 텍스트 위주 정보관리와는 달리, 웹 정보체계상에서는 상대적으로 단순한 정보관리업무에서 보다 복잡한 전자상거래 등으로 그 사건과 행위의 종류가 다양해지고 복잡해짐에 따라, 대상 자원보다는 실제 발생하는 다양한 사건 위주로 메타데이터요소를 기술하는 이벤트중심 모델을 중심으로 새로운 메타데이터 표준들이 개발되고 있다. 본 연구에서는 자원중심 모델인 FRBR 프레임워크 모델과 이벤트중심 모델인 INDECS 프레임워크 모델을 분석하였으며, 이를 기반으로 자원중심의 메타데이터 모델과 이벤트중심의 메타데이터 모델간의 상호운용성을 제공하기 위한 새로운 모형인 하이브리드 메타데이터 프레임워크 모델(Hybrid Metadata Framework)을 개발하였다.

## 2 상호운용성 모델 분석

### 2.1. IFLA의 FRBR 프레임워크 모델

전통적인 서지정보체계에서 가장 널리 쓰여진 MARC표준은 현 시점에서 데이터의 다양성 및 물리적 저장매체의 기술

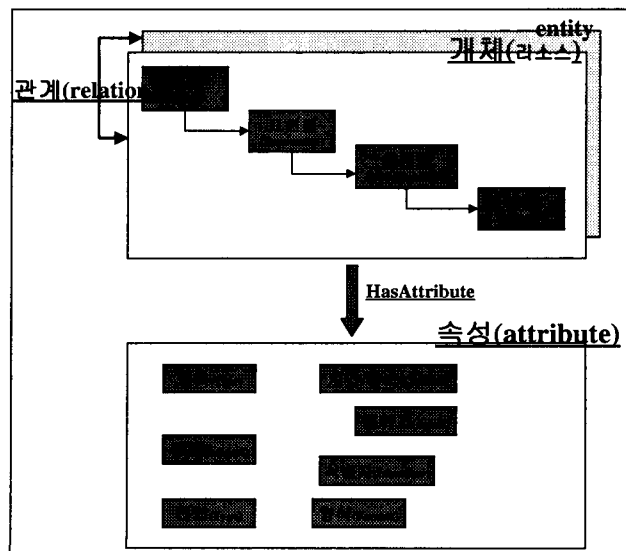
변화를 반영하는데 한계를 보여, 이로 인해 새로운 웹 정보체계 환경에 적용하는데에는 유연성이 부족하다는 단점을 지니고 있다. 따라서 메타데이터간의 상호운용성 측면에서 MARC을 기본으로 하는 단일형식으로서의 통합보다는, 메타데이터의 구조를 특정 형식으로 제한하지 않는 다양성을 인정하는 추세이다. 그 이유로는 기본적으로 하나의 메타데이터 형식이 전세계적으로 널리 사용될 수 없고, 또한 특정 형식이 요구되는 새로운 추세에 맞춰 메타데이터항목 요소들이 무한정 확장되는 것도 불가능하기 때문이다. 이는 이용자의 수준과 응용분야마다 요구되는 메타데이터요소들의 항목수와 전문성의 수준이 각기 다르기 때문에, 어떤 단일형식의 메타데이터로는 모든 조건을 충족시킬 수 없고, 따라서 메타데이터의 다양성을 인정하고 이를 수용할 수 있는 포괄적인 구조가 요구되기 때문이다(Lagoze 1996). 이러한 요구에 부응하기 위해서 다양한 메타데이터를 인정하고 이들을 하나의 통합된 틀 안에서 운용하기 위한 관점에서는 컨테이너(container)와 패키지(package) 개념을 이용해 메타데이터를 한 곳에 모으는 워릭구조(Warwick Framework)가 제안되었으며, 데이터와 메타데이터의 구별을 두지 않고 이 모두를 개별적인 객체로 보고 하나의 자원과 관련된 여러 자원을 모으고 연결시키려는 RDF(Resource Description Framework)가 제안된 바 있다(정상원, 안계성 1998).

이러한 추세의 하나로, IFLA 연구 그룹에서는 서지레코드를 대상으로 정보 검색을 지원하기 위한 기능적 요구사항을 정의하는 도구로서 FRBR 프레임워크 모델(Functional Requirements for Bibliographic Records)을 개발하였다(IFLA Study Group, 1998). 기본적으로 FRBR 프레임워크 모델은 관계형 데이터베이스 설계에 사용되는 개체-관계 모델(entity-relationship model)을 기초로 하고 있는데, 이는 FRBR 프레임워크 모델이 데이터베이스 시스템 자체를 설계하는 것을 목적으로 하고 있지는 않지만, 대상을 분석하는데 있어 체계적 분석을 가능하게 하면서 구조적 접근이 가능하도록 해주기 때문이다. 그러나 이러한 개체-관계 모델의 분석방법은 비논리적이거나 비정형화된 부분이 많은 주제를 다루기에는 부적합하며, 이러한 이유에서 기존의 이미 널리 사용되고 있는 메타데이터 표준들과는 서로 부합되지 않은 부분이 많다. 하지만, FRBR 프레임워크 모델 이외에 전통적인 서지정보 기반의 메타데이터를 대상으로 하는 표준모델을 찾아보기 힘든 관계로, FRBR 프레임워크 모델을 비교, 분석함으로써 자원 중심의 메타데이터 모델의 특징을 살펴보기로 한다.

FRBR 프레임워크 모델은 크게 개체(entity), 관계(relation), 속성(attribute)의 3부분의 요소(element)로 구성되는데, 먼저 대상 개체를 모두 식별하고, 다음 단계로 개체간의 관계를 확인하며, 마지

막으로 각각의 개체의 고유한 속성을 식별하게 된다. 예를들면 식별된 개체 중 하나가 저자(author)일 때, 저자의 주소(address)나 직위(position)등이 저자에 대한 고유한 속성으로 식별되어질 수 있으며, 저자는 저작물(creation)에 대한 관계로서 소유권(rights)을 가지게 된다. FRBR 프레임워크 모델에서 개체는 크게 다음 3가지 그룹으로 나뉘어지는데, 첫 번째 그룹으로 지적활동의 산물(product)로는 크게 저작(work), 표현형(expression), 구현형(manifestation), 개별자료(item)라는 4가지 요소가 존재한다. 저작은 추상적 개념의 창작물(creation)이며, 표현형은 저작이 문자 혹은 숫자, 이미지, 음악 등의 매체로 기록되어진 형태를 말한다. 구현형은 서지 혹은 오디오테이프, CD-ROM

등으로 제작된 출판형태를 말하며, 개별자료(item)은 ISBN등의 식별자가 부여된 각각의 개체를 지칭한다. 개체의 두 번째 그룹으로는 지적활동의 주체(acts)로서 'person'과 조직이나 단체를 가리키는 'corporate body'가 있다. 마지막으로 세 번째 그룹은 기타 주제(subject)와 관련된 항목으로 개념(concept)과 객체(object)가 있으며, 그밖에 나머지 항목들로 사건(event)과 장소(place)가 있다. 개체-관계 모델은 모든 항목요소가 개체, 관계 혹은 속성 중의 하나로 식별되어야 하며, 이러한 이유에서 사건과 장소 역시 개체라는 범주로 속하게 되었다. FRBR 프레임워크 모델에서 3가지 구성요소중 하나인 속성(attribute)은 개체의 항목요소별로 다양한 속성이 존재하며, 나머지 구성요



<그림 1> 자원중심의 메타데이터 모델

소인 관계(relation) 역시 개체들 간의 다양한 관계들이 존재하게 된다.

FRBR 프레임워크 모델을 토대로 자원 중심의 메타데이터 모델의 특징을 살펴보면 있어 가장 중요한 항목은 개체이며, 각각의 3가지 그룹의 개체들 가운데 가장 중심이 되는 부분은 저작, 표현형, 구현형, 개별자료로 구성되는 첫 번째 그룹의 개체(entity)이다. 이는 FRBR 프레임워크 모델의 첫 번째 그룹의 개체는, 더블린 코어와 같은 자원중심의 메타데이터 모델에서 대상 자원(resource)에 해당하는 부분이며, 대부분의 기존 메타데이터의 항목요소들은 대상 자원과 각 자원의 특징적 요소인 속성(attribute)이라는 두 부분으로 구분될 수 있기 때문이다. 결론적으로, 더블린코어로 대표되는 자원중심의 메타데이터 모델은 FRBR모델에서 첫 번째 그룹의 개체(entity)로 대표되는 개체와 관계, 속성의 3부분의 요소(element)로 구성되며 그 구체적 모습은 <그림 1>과 같다.

## 2.2. INDECS 프레임워크 모델

INDECS(INteroperability of Data in E-Commerce Systems)는 1998년 유럽연합이 지원하는 Info2000 프로젝트의 지원을 받아, 국제적인 저작권 소유자들의 주도로 시작된 프로젝트이다. INDECS 메타데이터 모델의 목표는 상이한 메타데이터 스키마가 상호연동(Crossworking)되고, 특정 분야에서 개발된 메타데이터가 다른 분야

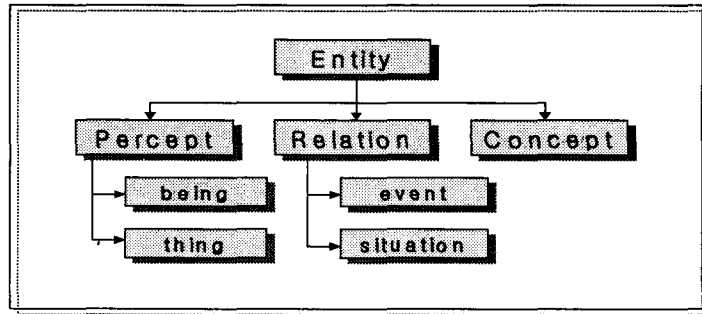
에서도 이용 가능하도록 하는데 있다 (Rust and Bide, 2000). INDECS 메타데이터 모델은 사건(event)을 중심으로 메타데이터 요소들을 기술하는 이벤트중심 모델로, INDECS 메타데이터 모델을 구성하는 주요항목은 다음과 같다(황상규, 윤세진, 오경목 2001).

### 1) 개체(entity)

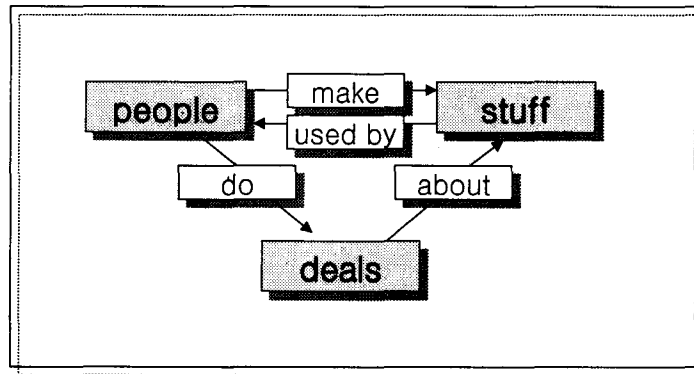
INDECS모델에서 개체는 식별되는 어떤 대상이다. 이것은 월드와이드웹 컨소시움(W3C)에 의해 채택된 자원(resource)와 그 의미가 비슷하다. 그러나 개체는 대상을 메타데이터 스키마 설계자가 보는 관점에 따라 서로 다른 메타데이터요소로 식별할 수 있다는 문제점을 가지고 있다. INDECS모델에서는 설계자 관점의 차이에서 발생할 수 있는 의미상 모호함(ambiguation)을 해결하기 위하여, 메타데이터 설계자 관점을 다음과 같은 3가지 관점(view)으로 구분하고 있다.

#### ① 일반적 관점(The general view)

개체가 식별하는 과정에 있어 일반적 관점은, <그림 2>와 같이 개체는 3개의 기본 유형들로 구분된다. 지각(percept)은 감각으로 받아들여지는 것이고 개념(concept)은 마음으로 생각되는 것이다. 그리고 관계(relation)는 두 가지 또는 세 가지 이상의 것들의 관계를 나타내는 것이다. 이러한 관계는 다시 사건(event)과 상태(situation)로 나누어진다. 사건은 두 개



<그림 2> 일반적 관점(general view)



<그림 3> 상업적 관점(commerce view)

이상의 개체간의 동적인 관계(dynamic relation)로서 발생하는 사건들을 말하며, 상태는 개체간의 정적인 관계(static relation)로서 '소유(possessing situation)'와 같이 시간이 흘러감에 따라 쉽게 변하지 않는 관계를 의미한다. 일반적 관점에서 사건은 개체와 개체를 연결하는 중요한 고리 역할을 하고 있다. 모든 메타데이터의 관계는 그 자체가 사건이거나 그것들을 구축하기 위해 사건에 의존하기 때문이다.

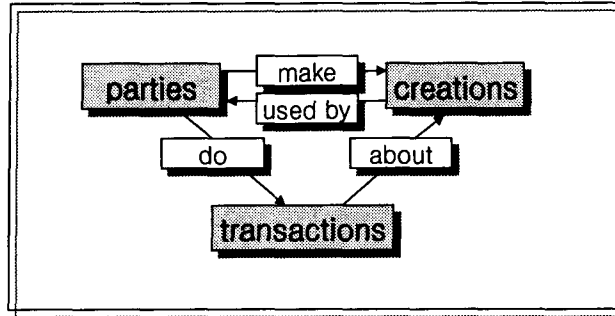
② 상업적 관점(The commerce view)

상업적인 관점은 실제 전자상거래 행위

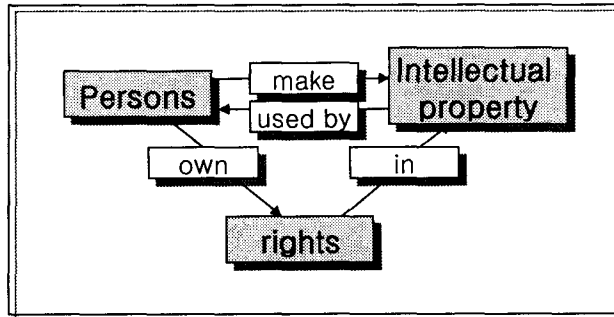
가 발생하는 과정에서 어떤 대상이 어떤 행위를 통해 만들어 졌는가에 초점을 두고 있다.

위의 <그림 3>에서 행위자인 사람(people)은 작품(stuff)를 만들고 사용하며, 이러한 이유로 사람은 창작자(creator)인 동시에 사용자(user)가 될 수 있다. INDECS 모델은 이러한 상업적인 개체를 표현하기 위하여 아래 <그림 4>와 같은 3가지 기본 타입들을 제공한다.

<그림 4>에서 역할(parties)은 <그림 3>의 행위자로서 상거래를 하는데 있어서



〈그림 4〉 상업적인 개체(entity)의 3가지 기본 유형



〈그림 5〉 지적재산권적 관점(legal view)

주체가 되는 사람이나 단체를 가리킨다. 또한 창작물(creations)은 작품으로서 행위자간에 이루어지는 상거래(transactions)라는 사건의행위의 대상(input)을 가리킨다.

③ 지적재산권적 관점

(The intellectual property view)

마지막으로 전자상거래 행위가 발생하는 과정에서 행위자인 사람(people)은 지적재산이라는 관점에서 저작에 대한 권리(right)를 만들고 사용한다.

전자상거래 행위가 발생하는 과정에서 행위자 A를 개체로 식별하고자 할 때 일

반적 관점에서는 사람(people)으로, 상업적 관점에서는 역할(party)로, 지적재산권적 관점에서는 지적재산권 소유자(person)로 식별될 것이며, 각각은 서로 다른 속성을 갖는다.

일반적으로 기존의 메타데이터는 하나의 관점 밖에 갖지 못함에 비하여, INDECS 모델에서는 대상을 보는 관점을 일반적 관점, 상업적 관점, 지적재산권적 관점이라는 서로 다른 관점에 따라 다양하게 표현되며, 이는 개체간의 관계를 표현함에 있어 다양성을 제공해 주기 때문에 기존의 메타데이터가 서로 각기 다른 관점을



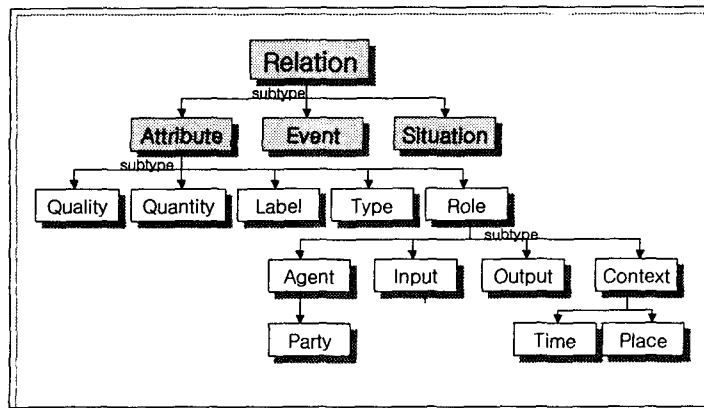
가짐에 따라 서로 다른 메타데이터간의 상호운용시 발생하는 상호연동(crosswalking)의 문제를 해결할 수 있다(윤세진 2002).

자원중심의 메타데이터 모델과 이벤트 중심의 메타데이터 모델에서 대상 개체(entity)를 보는 관점에서 차이를 살펴보면, 자원중심의 메타데이터 모델에서 지적활동의 주체(acts)가 하나의 개체임에 비해, INDECS와 같은 이벤트중심의 메타데이터 모델에서는 별도의 행위자(agent)라는 요소(element)로 분리된다. 또한, 개체의 범주에 속하는 '사건'과 '장소'의 경우에는, INDECS와 같은 이벤트 중심의 메타데이터 모델에서는 개체(entity)가 아닌 별도의 사건, 시간, 장소라는 3가지 구성요소(element)로 분리된다. 마지막으로, 더블린코어 메타데이터의 경우를 살펴보면, 두 번째 그룹의 개체에서, 행위자의 하나인 'person'은 INDECS와 같은 이벤트 중심의 메타데이터 모델에서는 별도의 행위자(agent)로서 그 역할(role)에 따라 창작

자(creator) 혹은 편집자(editor)등으로 세분화된다.

## 2) 관계(Relation)

관계는 INDECS모델에서 가장 중요한 구조로 사건(event), 상태(situation), 속성(attribute) 3가지로 구분된다. 이러한 관계의 구성요소는 <그림 6>과 같다. 각각의 구성요소에 대해 살펴보면, 사건은 개체간의 동적인 관계를 나타내며, 상태는 개체간의 정적인 관계를 나타낸다. 속성은 개체가 가지는 특성으로 INDECS모델에서는 어느 개체든 라벨(labels), 양(quantities), 질(qualities), 형식(types), 역할(roles)이라는 5가지 속성을 갖는다. 라벨은 사람이름(name)과 DOI같은 식별자(identifier)를 나타내는 것이며, 양(quantities)은 객체(object)의 물리적인 형태를, 질(qualities)은 객체(object)를 창작했는지, 번역했는지 등의 내용과 관련된 정보를 나타낸다. INDECS모델에서 형식은 개체



<그림 6> 관계(relation)의 구성요소

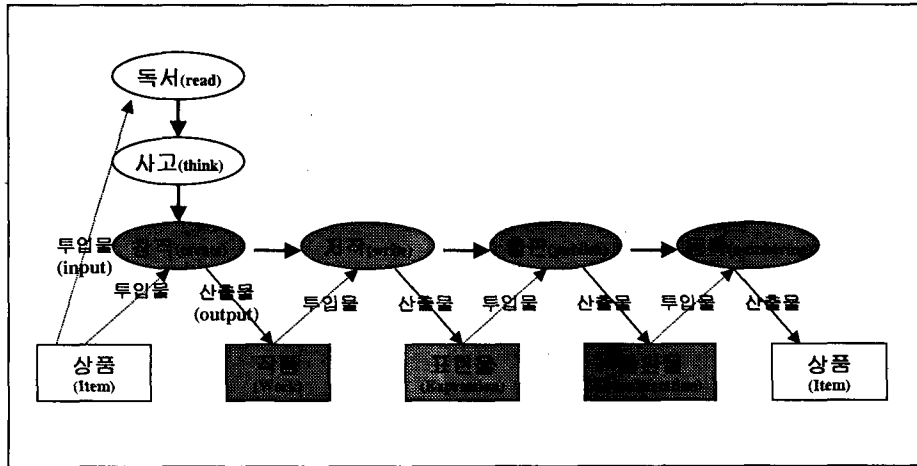
(entity)그룹에 속하는 개체를 하나 또는 그이상의 특징들로 분류한 것이라고 정의하고 있으며, 더블링크어의 형식과는 그 용도가 서로 다르다. INDECS모델에서의 형식은 더블링크어의 형식에 해당하는 장르(genre)뿐만 아니라 이용자(audience), 형식(format), 저작권 유형(IPtype) 등을 포함하는 것으로 더블링크어에서 보다 더 포괄적인 의미로 사용된다는 것을 제시해주고 있다.

실제 사건이 발생하는 과정에서 둘 또는 그 이상의 개체는 서로 간의 관계를 형성하며, 그 과정에서 하나의 개체는 서로 다른 개체에 대해 어떠한 역할(roles)을 하게 되며, INDECS모델에서는 행위자(agent), 투입물(input), 산출물(output), 시간/장소(context)라는 4개의 일반적인 역할을 제공한다. 행위자는 일반적으로 사람이나 다른 존재, 그 존재가 속한 기관 등의 주체를 나타내며, 사건과 상태의 특성을 결정한다. 행위자의 서브타입(sub-type)으로 역할(party)이 존재하는데 INDECS 모델에서 역할은 “어떤 사건이 발생하는 과정에서 행위자의 역할은 무엇인가?”에 의해 정의되는 것으로 창작행위의 주체로서 행위자는 창작자(creator)의 기능을 수행하게 된다. 역할(party) 또한 앞의 관계(relation)와 마찬가지로 모든 일반적 속성(labels, quantities, qualities, types, roles)을 포함한 구조를 갖는다. 역할(party)에서 라벨은 이름, 양(quantities)은 나이, 질(qualities)은 성, 형식은 개인과 기관, 역

할(role)에서는 탄생·죽음·결혼 등이 존재하게 된다. 'input role'은 사건(event)이 발생하는 과정을 적합하게 하거나 그것에 종속되는 수동적인 역할(role)로 투입물을 의미한다. 'output role'은 사건이 발생하는 과정으로부터 기인된 개체로 산출물이라고 할 수 있다. 마지막으로 'context role'은 사건이 발생하는 시간과 장소를 기술한다. 이러한 INDECS모델의 특징은 이벤트 기반 상호운용성 모델인 다음의 ABC모델에도 잘 나타나 있다.

### 2.3. ABC모델

ABC모델은 서로 다른 형식의 메타데이터간 1:1매핑을 통한 메타데이터간의 상호운용성(interoperability)의 한계를 극복하고, 보다 보편적인 상호연동(crosswalk)을 제공하기 위한 상호운용성 모델로서 제안되었다(Lagoze, Hunter and Brickley 2000). ABC모델의 가장 큰 특징은 이벤트 기반의 상호운용성 모델이라는 점이며, 이는 IFLA, CIDOC, INDECS, MPEG-7, DC등의 기존 메타데이터 모델을 비교, 분석함으로써, 이벤트중심의 분석방법이 가장 효과적인 모델링 방식임을 ABC 워크숍을 통해 제시하였다(ABC Workshop 2000). 예를 들어 IFLA FRBR모델에서 대상이 되는 단일 저작물은 개체(entity)의 관점에서 볼 때, 작가의 추상적 개념의 창작물(creation)에서 출발하여, 이를 문자 혹은 숫자, 이미지, 음악 등의 매체



〈그림 7〉 창작에서 출판까지 이벤트 흐름도

로 표현(expression)하고, 이는 다시 서지 혹은 오디오테이프, CD-ROM 등의 매체로 구현(manifestation)하며, 구현된 각각의 개별자료(item)에는 ISBN 등의 식별자가 부여되어 각각의 개체를 식별하게 된다. 이러한 일련의 과정(life-cycle)에는 각 단계마다 어떠한 사건(event)이 발생하며, 그 산출물로서 새로운 개체가 생성되기 때문이다. 또한, 한 단계의 산출물(output)은 또 다른 사건의 투입물(input)로서 사용된다. 창작에서 출판으로 이어지는 이러한 일련의 흐름을 이벤트 중심의 관점에서 다음 〈그림 7〉과 같이 표현할 수 있다.

이와 같이 ABC모델에서 가장 중요한 요소는 사건이며, ABC모델은 다음과 같은 총 6개의 항목들로 구성되어 있다.

① 자원(Resources)

이 세상에서 존재하는 물리적, 추상적인 모든 것을 지칭하는데, FRBR모델에서

개체와 그 개념이 유사하다.

② 사건(event)

사건이란 어떠한 시간과 장소에서 발생하는 행위를 지칭하며, 사건의 투입물(Input)과 사건의 산출물(Output)을 가지게 된다.

③ 투입물(Input) / 산출물(Output)

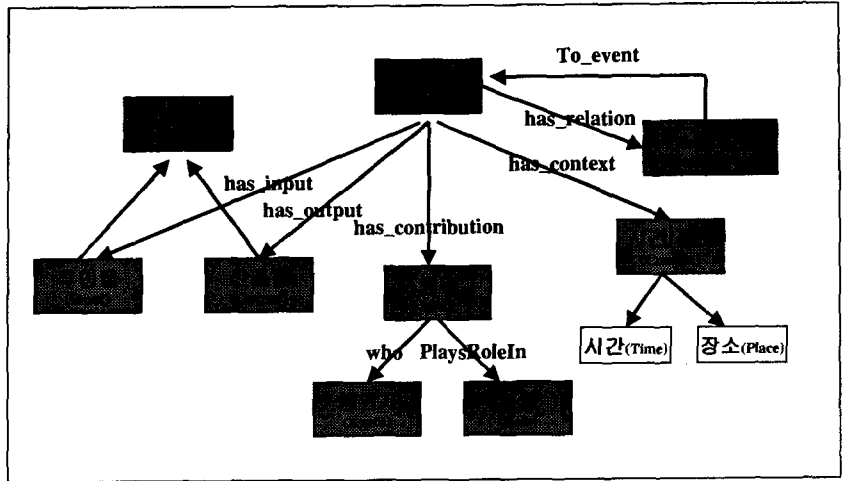
어떠한 사건이 발생하는 과정에는 투입물(Input)과 사건의 산출물(Output)이 존재한다.

④ 행위자(Acts) / 역할(role)

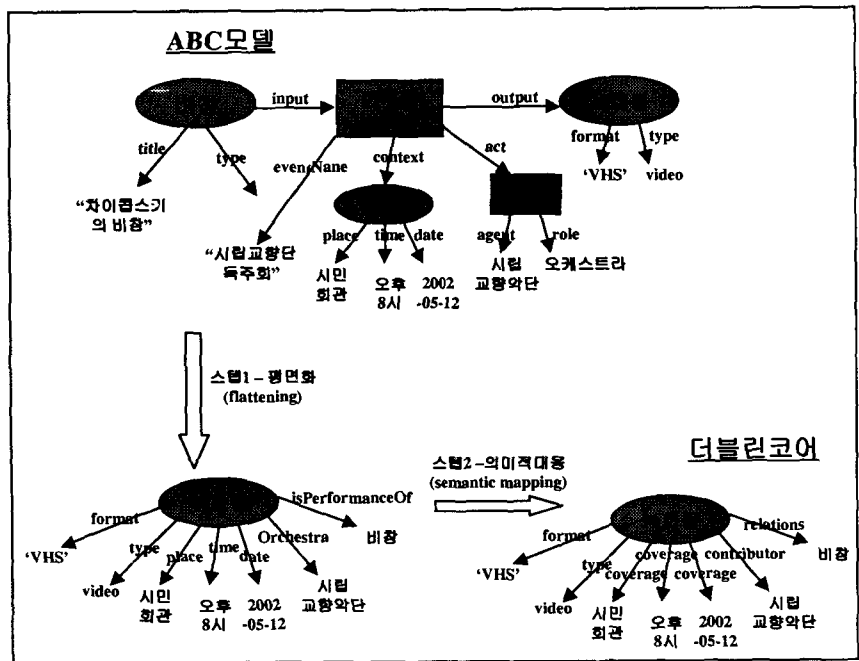
어떠한 사건이 발생하는 과정에는 사건의 주체로서의 행위자(Acts)와 그 사건에서 행위자의 역할(role)이 존재한다.

⑤ 시간/장소(context)

어떠한 사건이 발생하는 시간(time)과 장소(place)를 기술한다.



<그림 8> ABC모델



<그림 9> ABC모델과 더블린코어간의 상호연동

⑥ 관계(Event Relations)

발생하는 사건들간의 관계를 기술한다. 자원중심의 메타데이터 모델과의 가장 큰 차이점은 FRBR모델에서는 개체간의 관계가 성립됨에 비하여, ABC모델에서는 이벤트 중심의 관점에서 대상을 모델링함으로써 사건들간의 관계를 식별하게 된다.

실제 상황으로 적용함에 있어 ABC모델은, 전통적인 메타데이터표준들의 특성을 고려하지 않고, 항상 이벤트 중심의 관점에서 대상을 분석하고자 하는 문제점을 가지고 있다. ABC모델에서는 <그림 9>와 같이 사건(event) 중심의 관점에서 대상을 분석하고, 이를 토대로 메타데이터간의 상호운용성을 제공하게 된다.

<그림 9>에서 사건(event) 중심의 메타데이터를 더블린코어와 같은 자원 중심의 메타데이터 모델로 전환하는 과정은 사건(event)모델로부터 사건 흐름의 중심이 되는 개체를 식별하는 과정으로부터 시작된다. 사건(event)중심의 메타데이터모델을 자원중심모델로 전환하는 과정에 있어서, 먼저 중심이 되는 개체를 식별하고, 각각의 대표개체별로 관련된 나머지 구성요소(element)들을 나열하는 평면화(flattening)라는 작업을 수행하게 된다. 그 다음 단계로, 대표개체를 중심으로 나열된 구성요소들을 실제 더블린코어의 메타데이터 항목(element)으로 대응(semantic mapping)시킴으로서 자원중심모델로의 전환이 이루어진다. 하지만, 이러한 ABC모델의 방식은 역으로 자원중심모델로부터

사건(event)중심의 메타데이터모델로의 전환 과정에 있어, 중심이 되는 사건을 추출하는 명확하고 구체적인 방안을 제시하지 못하고 있다. 이는 결국, ABC모델의 방법을 통해 더블린코어와 같은 자원 중심의 메타데이터 모델을 INDECS와 같은 사건(event) 중심의 메타데이터 모델로 전환하는 과정에는 상당부분 모호한 부분이 존재할 수 있음을 보여주게 되는 것이다.

### 3 하이브리드 메타데이터 프레임워크 개발

앞에서 살펴본 바와 마찬가지로, ABC 모델의 방법을 통해 더블린코어와 같은 자원 중심의 메타데이터 모델을 사건(event) 중심의 메타데이터 모델로 전환하는 과정에 있어 중심이 되는 사건(event)요소를 식별하는 구체적인 방법이 명확히 제시되지 못함을 확인한 바 있다. 이러한 ABC모델의 문제점을 해결하는 방안으로, 사건(event) 중심의 분석방법이 아닌 자원(resource) 중심의 분석방법이 보다 효과적인 해결방안이 될 수 있음을 발견할 수 있다.

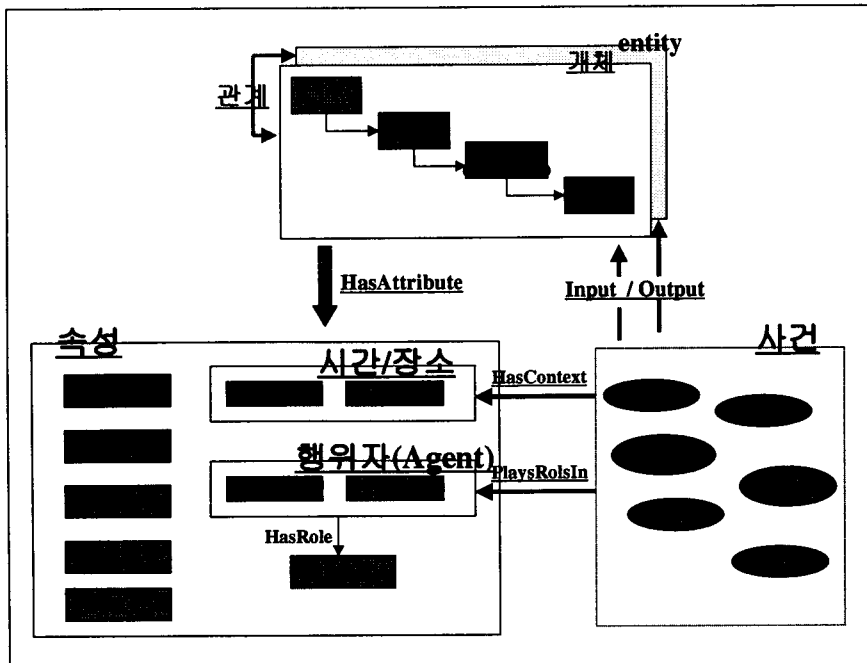
<그림 7>에서 창작(creation)에서 저술(write)로, 저술에서 출판(publish)으로, 출판에서 식별자부여(numbering)라는 일련의 사건(event) 흐름에서 '저작(work)', '표현형(expression)', '구현형(manifestation)'이라는 개체가 일련의 사건흐름의 투입물(input)/산출물(output)이 됨으로서 사건

과 사건간의 연결고리 역할을 수행함을 확인할 수 있다. 이처럼 하나의 사건과 사건 간에는 일련의 사건 흐름에 중심이 되는 개체(core entity)가 존재하며, 이는 중심개체가 자원 중심의 메타데이터 모델을 사건 중심의 메타데이터 모델로 전환하는 과정에서 중심연결고리 역할을 제공할 수 있음을 발견하였다. 이를 토대로 본 연구에서는, 자원 중심의 메타데이터 모델과 사건 중심의 메타데이터 모델간을 통합, 분석할 수 있는 기반을 제공하기 위한 새로운 방식의 하이브리드 메타데이터 프레임워크 모델을 개발하였다. 하이브리드 메타데이터 프레임워크 모델은 사건의 흐름에 중심이 되는 중심개체(core

entity)와 중심사건(core event)간의 변환(swapping)을 수행함으로써, 사건 중심의 메타데이터 모델을 자원 중심의 메타데이터 모델로 전환하는 과정에 있어, 자원 중심의 메타데이터 모델과 사건 중심의 메타데이터 모델간의 중간연결고리(중간 모델)를 제공한다. <그림 10>은 하이브리드 메타데이터 프레임워크 모델 중 첫 번째 단계인 사건 모델로, INDECS와 같은 기존 사건 중심의 메타데이터 모델과 같이 4가지 구성요소로 재구성하게 된다.

① 개체(entity)

하이브리드 메타데이터 프레임워크 모델의 중심은 하나의 사건과 다른 사건 흐름



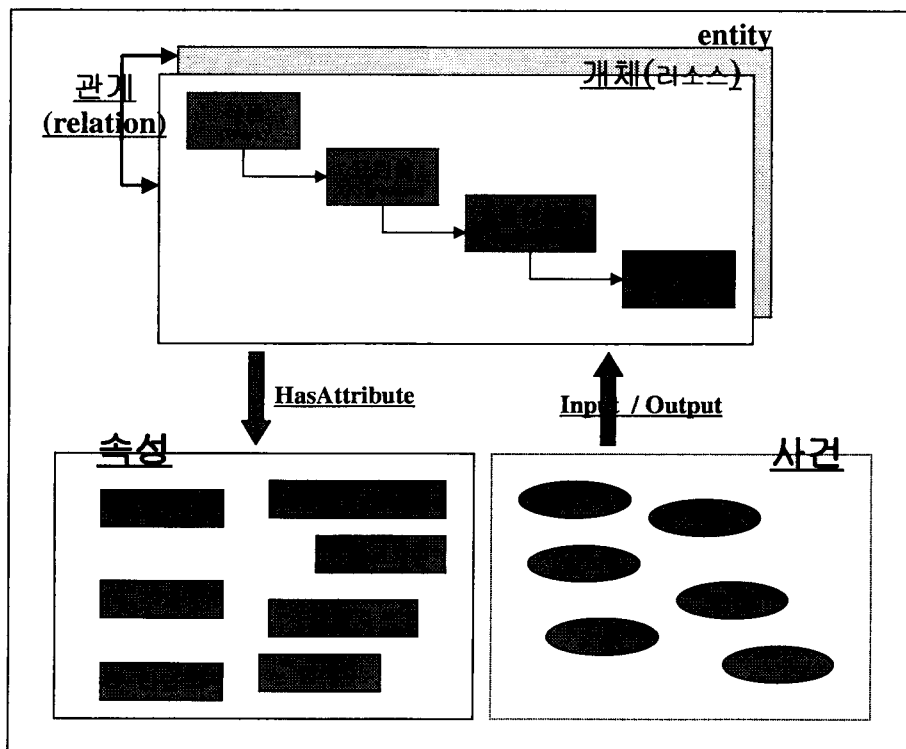
<그림 10> 하이브리드 메타데이터 프레임워크(사건 모델)

간에 중심이 되는 개체이며, 주요 사건들 간의 연결고리 역할을 수행하는 자원만을 중심개체(core entity)로 식별하게 된다. 하나의 예로, 원저작물인 구현형(manifestation) 'The Ant'는 출판(publish)이라는 행위과정에서 하나의 중심개체로 식별될 수 있는데, 이는 'The Ant'가 출판이라는 행위의 산출물(output)인 동시에, 원저작물의 한국어판인 '개미'를 만드는 번역(translate) 과정에서 투입물(input)로 사용됨으로써, 사건과 사건간의 연결고리 역할을 수행하는 하나의 중심개체로 식별될 수 있기 때문이다. 그러나, 'The Ant'를 각색하여 새

로 기획하는 영화의 소재로 삼은 경우에는, 단순한 'The Ant'는 영화를 제작하는 과정에서 단순한 투입물(subject)로서의 기능만을 수행하므로 하나의 중심개체로 식별되지는 않으며, 속성(attribute)의 한 요소로만 분류된다.

② 관계(relation)

하이브리드 메타데이터 프레임워크 모델에서 관계는 개체와 개체간의 관계로서, 기존 FRBR모델에서 관계(relation)와 그 의미가 동일하다.



<그림 11> 하이브리드 메타데이터 프레임워크(중간 모델)

③ 사건(event)

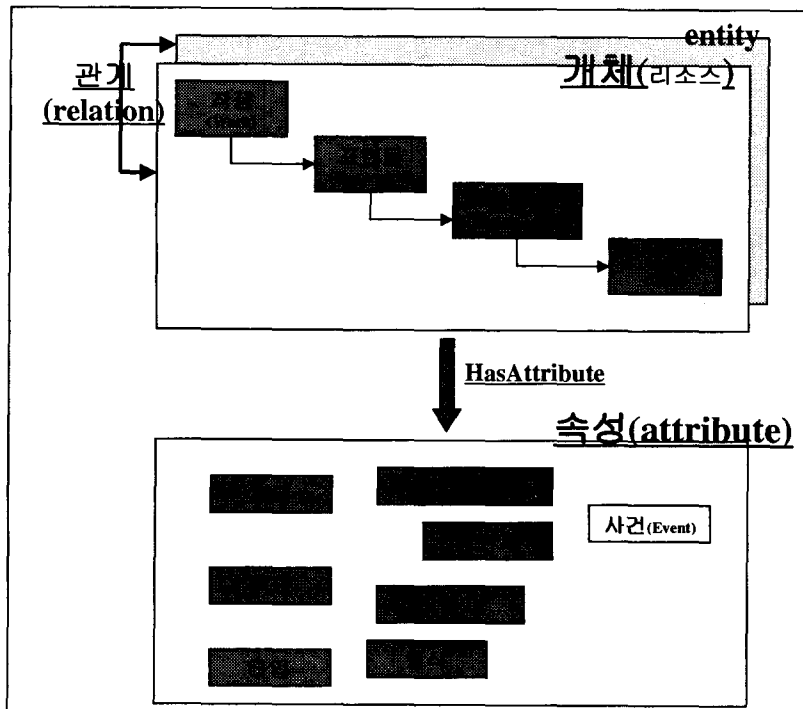
하이브리드 메타데이터 프레임워크 모델에서 사건은 일련의 행위의 흐름에 중심이 되는 개체(core entity)와 관련된 중심 사건으로서, 최종적으로 자원 중심의 메타데이터 모델로 전환 시에는 속성(attribute)의 한 요소로 포함되게 된다. 앞 사례에서 'The Ant'라는 원저작물을 토대로 이루어지는 출판(publish), 편집(edit), 번역(translate)등의 행위가 사건으로 식별될 수 있다.

에서 속성은 하이브리드 메타데이터 프레임워크 모델에 중심이 되는 개체와 관련된 모든 항목요소들을 기술하게 된다. 개체에 속한 속성들 가운데 발생한 사건이나 행위의 주체인 행위자와 사건이나 행위가 발생한 시간 혹은 장소(Context)를 식별하면 사건 중심의 메타데이터 모델이 만들어진다.

ABC모델과 비교할 때 하이브리드 메타데이터 프레임워크의 사건 모델의 가장 큰 특징은, ABC모델이 사건(event)중심으로 모델링이 이루어지는 것과는 달리 일련의 사건 흐름의 중심이 되는 개체(core entity)를 중심으로 사건모델이 표

④ 속성(attribute)

하이브리드 메타데이터 프레임워크 모델



<그림 12> 하이브리드 메타데이터 프레임워크(자원 모델)

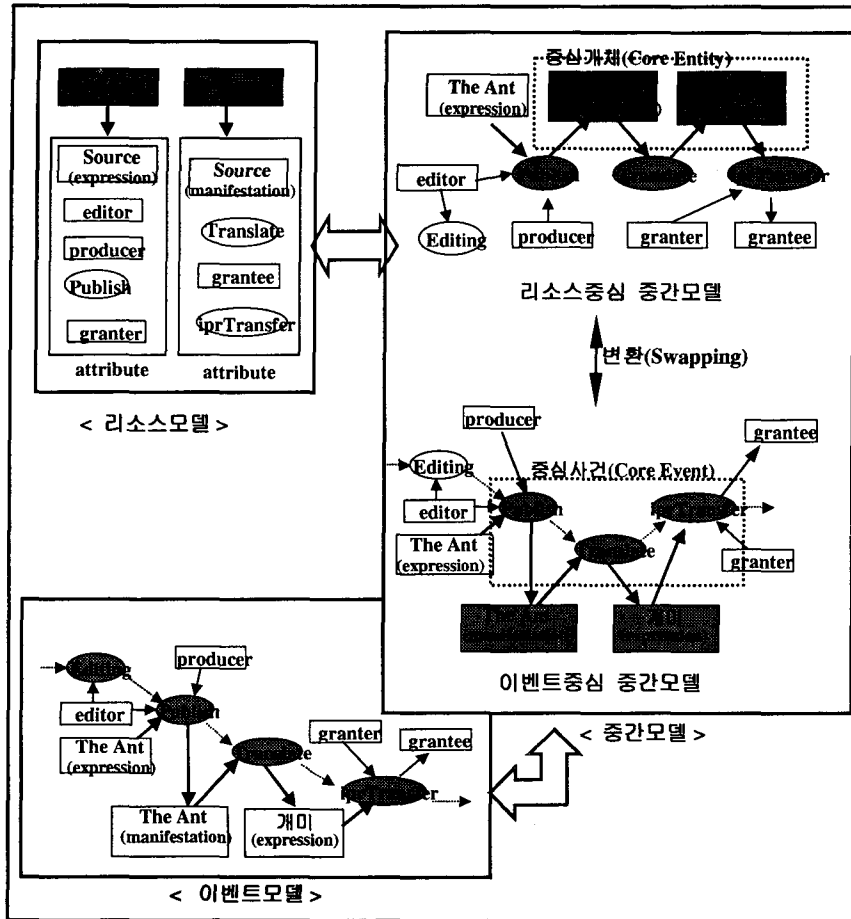


현되어 진다는 점이다. <그림 10>의 사건 모델로부터 단순히 행위자(agent)와 시간/장소(context)를 단순히 속성(attribute)의 하나로 재분류함으로써 자원 중심의 메타데이터 모델로 전환하기 위한 중간연결고리로서 중간모델을 생성하며, 구체적인 모습은 <그림 11>과 같다. 하이브리드 메타데이터 프레임워크의 중간모델은 사건 모델로부터 사건(event)흐름 간에 중심이 되는 개체(core entity)와 이와 대응되는 주요 사건(core event)간의 연관관계를 식별함으로써 사건 모델과 자원모델 간의 중간 연결고리를 확인하게 된다. <그림 7>을 살펴보면, 식별된 모든 행위들 가운데 'create', 'write', 'publish', 'numbering'이라는 중심 사건(core event)을 식별할 수 있는데, 이는 일련의 사건의 흐름에 있어 한 행위의 결과(output)가 다음 행위의 투입물(input)로 이어짐으로써 일련된 흐름의 연결고리 역할을 수행하기 때문이다. 이와 마찬가지로, 주요 사건간의 연결고리 역할을 수행하는 중심이 되는 개체로서 '저작(work)', '표현형(expression)', '구현형(manifestation)'을 식별할 수 있는데, 식별된 각각의 중심개체는 자원모델로 전환 시 각각 하나의 개체(자원)로 식별되어 진다. 결국 중간모델에서의 중심개체를 자원모델에서 각각의 개체(자원)로 식별함으로써, 최종적으로 <그림 10>의 사건모델에서 <그림 12>와 같은 자원모델로의 전환이 이루어지게 된다.

이와 반대로 자원모델을 사건모델로 전

환하는 과정에 있어서는, 자원모델로부터 각각의 개체(자원)마다 관련된 모든 사건을 식별하고, 이를 토대로 사건모델로 전환하기 위한 중간모델을 생성하게 된다. 중간모델을 생성하는 과정은 앞에서 설명한 바대로 <그림 7>에서처럼, 식별된 모든 사건들 가운데 일련의 흐름에 중심이 되는 주요 사건(core event)들을 식별함으로써 최종적으로 사건모델로 전환하기 위한 중간모델을 생성하게 된다. 구체적인 예로 <그림 13>을 살펴보면, 중간모델을 만드는 과정에 있어서 발생한 사건의 순서에 따라 주요 사건(core event)을 기준으로 모델링이 이루어지는 사건중심 중간모델과 사건 흐름간에 중심이 되는 개체(core entity)를 기준으로 모델링이 이루어지는 자원중심 중간모델이 각각 존재함을 확인할 수 있다. 중심사건과 중심개체는 이미 앞에서 언급한 바대로 <그림 7>에서처럼 서로간의 투입물(input)과 산출물(output)이 되는 밀접한 연관관계를 가지게 되며, 자원 모델에서 사건모델로의 전환과정에 있어 중간 연결고리 역할을 제공하게 된다.

이와 같이 중간모델의 기준점을 중심개체로부터 중심사건으로 전환함에 따라, 자원중심 중간모델로부터 사건중심 중간모델로의 변환(swapping)과정이 이루어지게 된다. 이러한 전환과정을 통하여, 하이브리드 메타데이터 프레임워크 모델은 자원중심의 메타데이터 모델에서 사건중심의 메타데이터 모델로의 전환방법이 명



<그림 13> 자원모델과 사건모델 간의 변환과정(사례)

확히 제시되기 어렵다는 기존 ABC모델의 문제점에 대한 구체적 해결방안을 제시하게 된다. 또한, 하이브리드 메타데이터 프레임워크 모델의 변환(swapping) 과정은 사건모델에서 자원모델로의 전환과정에 있어서도 마찬가지로 동일하게 적용됨에 따라 기존 ABC모델의 평면화(flattening)작업을 대신할 수 있게 된다.

마지막으로, <그림 13>의 예제를 통해

자원모델과 사건모델간의 구체적인 전환 과정은 다음과 같다. 먼저 자원모델에서 사건모델로의 전환과정을 살펴보면, 자원모델에서 'The Ant'와 '개미'라는 2가지 중심이 되는 개체(자원)와 관련된 속성들 가운데 식별 가능한 모든 사건들을 식별하는 과정에서, 사건들 서로간의 투입물(input)과 산출물(output)이 이어지는 연관관계를 확인함으로써, 출판(publish), 번

역(translate), 지적재산권양도(iprTransfer)라는 중심사건(core event)을 식별해내게 된다. 그 다음 단계로 출판, 번역, 지적재산권양도라는 중심사건들과 관련 있는 추가적인 개체(editor, producer, grantor, grantee)들을 식별하고, 이들 중심 사건과 개체들 간의 관계를 그림으로 표현하면, 중심개체(core entity)위주의 자원중심 중간모델이 만들어지게 된다. 이를 다시 중간모델의 기준점을 중심개체로부터 중심사건으로 전환함에 따라, 개체 중심에서 사건 중심으로 변환(swapping)과정을 통하여 사건중심 중간모델을 만들어내게 된다. 마지막으로 사건 중심의 중간모델의 개체들로부터 행위자(Agent)와 시간/장소(context)를 식별해내면 자원모델에서 사건모델로의 전환이 완료되어 진다. <그림 13>에서 식별된 행위자는 총 4명으로, 이들은 각각 사건(event)이 발생하는 시점에서 편집자(editor), 프로듀서(producer), 양도인(granter), 양수인(grantee)라는 역할(role)을 수행하고 있다.

사건모델에서 자원모델로의 변환과정은, 자원모델에서 사건모델로의 변환과정에서 이루어졌던 사건(event)과 행위자(agent), 시간/장소(context)를 식별해내는 과정이 생략됨으로서 보다 쉽게 자원중심의 메타데이터 모델로의 전환이 이루어질 수 있다. 먼저 사건모델로부터 사건 중심의 중간모델을 만드는 과정에서, 하나의 사건과 다른 사건흐름간에 연결고리 역할을 수행하는 사건만을 중심사건(core

event)으로 식별하게 되는데, <그림 13>에서는 'Publish', 'Translate', 그리고 'iprTransfer'만이 중심개체로서 식별되어 진다. 이를 다시 중심개체를 중심으로 변환(swapping)과정을 거치면 자원중심의 중간모델이 만들어지게 된다. 마지막으로 식별된 중심개체 'The Ant'와 '개미'를 개체(자원)로하여, 각각의 개체(자원)와 관련된 나머지 부수적인 항목들을 그 개체에 속한 속성(attribute)으로 나열함으로써, 최종적으로 더블링크어 형태의 자원모델이 만들어지게 된다.

## 4 결론 및 제안

본 연구에서는 다양한 종류의 메타데이터를 통합, 분석할 수 있는 기반을 제공하기 위한 상호운용성 모델로서 하이브리드 메타데이터 프레임워크 모델을 개발, 제안하였다. 하이브리드 메타데이터 프레임워크 모델은 기존 사건 중심의 상호운용성 모델인 ABC모델의 문제점을 극복하고 보다 발전시킨 모델로서, 기존 자원 중심의 메타데이터 모델과 사건 중심의 메타데이터 모델간의 연동 기반(crosswalk framework)을 제공하게 된다.

사건 중심의 상호연동모델인 ABC모델은 자원중심의 메타데이터 모델에서 사건 중심의 메타데이터 모델로 전환하기 위한 명확한 방법을 제시하지 못하고 있다. 이러한 문제점을 해결하기 위하여, 하이브리드 메타데이터 프레임워크 모델은 자원

중심의 메타데이터 모델을 사건 중심의 메타데이터 모델로 전환하는 과정에서 변환(swapping)과정을 거치는 중간단계모형을 새로 제시함으로써, 사건 중심의 상호운용성 모델인 ABC모델의 문제점을 보완하였다. 기존 ABC모델에서는 모델링에 기준이 되는 사건(event) 요소를 더블 링크어의 기본 메타데이터 항목요소들 가운데에서 직접 추출함으로써 내용이 복잡할 경우에는 추출과정자체가 상당히 어려워지는 문제점이 발생하였으나, 하이브리드 메타데이터 프레임워크 모델에서는 자원모델의 기본 메타데이터 항목요소들로부터 중심이 되는 개체(core entity)와 이 와 연관된 주요사건(core event)들을 식별하고, 이를 자원중심 중간모델과 사건 중심 중간모델로 도식화하는 변환(swapping)과정의 구체적인 방법을 제시함으로써, 자원모델에서 사건모델로의 보다 명확한 전환방법을 제시하고 있다.

본 연구에서는 웹 정보체계인 'Semantic Web'에서 여러 가지 다양한 종류의 메타데이터가 표준화된 의미정보전달을 위한 수단으로 활용되어질 수 있도록, 메타데이터간의 연동 기반을 제공할 수 있는 자원중심의 상호운용성 모델을 제안하였다. 그러나, 상호운용성 모델을 토대로 실제 서로 다른 메타데이터간의 표준화된 의미정보전달이 이루어지기 위해서는 서로 다른 메타데이터간의 용어 표준화라는 또 다른 문제도 해결되어야 한다. 따라서, 서로 다른 메타데이터간의 메타데이터 항목

을 중심으로 어휘(dictionary)수준이 아닌 한 차원 위에 지식베이스(Knowledge Base) 수준에서 용어 표준화가 이루어질 수 있는 방안에 대한 구체적인 연구가 진행되어야 할 것이다.

## 참고문헌

- 정상원, 안계성. 1998. 디지털도서관의 표준화 과제. 『제3회 과학기술 정보 워크샵: 사용자 중심의 과학기술 정보 서비스 발표 논문집』.
- <[http://www.metadb.net/~bstone/vitae/DL\\_Standard.html](http://www.metadb.net/~bstone/vitae/DL_Standard.html)>
- 황상규, 윤세진, 오경목. 2001. 디지털 콘텐츠를 위한 메타데이터에 관한 연구:INDECS를 중심으로. 『제 8회 한국정보관리학회 학술대회 논문집』, 2001년 8월 23-24일. [대전: 한국과학기술정보연구원].
- 윤세진. 2002. 『디지털자원 관리를 위한 메타데이터간의 상호운용성에 대한 비교연구』. 석사학위논문. 숙명여자대학교 대학원, 문헌정보학과.
- Tim Berners-Lee, 1998. "Semantic Web Road map" [cited 2001.5.13]. <<http://www.w3.org/DesignIssues/Semantic.html>>

- Tim Bray, Jean Paoli, C. M. Sperberg-McQueen, 2000. *Extensible Markup Language (XML) 1.0*. [cited 2001.4.10].  
 <<http://www.w3.org/TR/REC-xml>>
- Ora Lassila, Ralph R. Swick, 1999. *Resource Description Framework (RDF) Model and Syntax Specification. REC-rdf-syntax-19990222*. [cited 2001.3.16].  
 <<http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>>
- Ian Horrocks, Frank van Harmelen, Peter Patel-Schneider, 2001. *DAML+OIL*. [cited 2002.2.3]  
 <<http://www.daml.org/2001/03/daml+oil-index>>
- R.V. Guha, Ora Lassila, Eric Miller, Dan Brickley, 1998. Enabling Inferencing. W3C Query Languages meeting in Boston. [cited 2001.7.23].  
 <<http://www.w3.org/TandS/QL/QL98/pp/enabling.html>>
- Lagoze, Carl. 1996. "The Warwick framework: A container architecture for diverse sets of metadata". *D-Lib Magazine*(July/August 1996), [cited 2001.5.13].  
 <<http://www.dlib.org/dlib/July06/lagoze/07lagoze.html>>
- IFLA Study Group, 1998. *Functional Requirements for Bibliographic Records: final Report*. [cited 2001.8.5].  
 <<http://www.ifla.org/VII/s13/frbr/frbr.htm>>
- Rust, Godfrey and Mark Bide. 2000. *The INDECS metadata framework*. [cited 2002.5.13].  
 <<http://www.indecs.org>>
- ABC Workshop, 2000. [cited 2001.4.17].  
 <[http://www.ilrt.bris.ac.uk/discovery/harmony/abc\\_workshop.htm](http://www.ilrt.bris.ac.uk/discovery/harmony/abc_workshop.htm)>
- Lagoze, J., D. Hunter & Brickley, 2000. An Event-Aware Model for Metadata Interoperability, ECDL 2000. [cited 2001.5.13].  
 <[http://archive.dstc.edu.au/RDU/staff/jane-hunter/harmony/harmony\\_ECDL2000.zip](http://archive.dstc.edu.au/RDU/staff/jane-hunter/harmony/harmony_ECDL2000.zip)>