

이미지 자료의 관리를 위한 효율적인 디지털 아카이브 워크플로우와 메타데이터 표현에 관한 연구

김효원*, 윤용익**

요약

사회의 발달에 따라 정보의 양이 폭발적으로 증가하게 되어 이를 효율적으로 유지, 보관하기 위해 디지털 아카이브의 필요성이 대두되었다. 현재 아카이브의 표준으로는 OAIS(Open Archival Information System) 참조 모형이 있다. 이 모형은 장기보존에만 관심이 있고, 주로 포괄적인 콘텐츠에만 신경을 쓰고 있다. 본 논문에서는 이미지 자료를 대상으로 장기보존만이 아닌 효율적인 활용을 위한 디지털 아카이브 워크플로우와 메타데이터에 대한 모델을 제시하고자 한다. 이미지 자료 디지털 아카이브 구축을 위해서는 OAIS 참조 모형을 통해 서지 정보와 미디어 형태의 메타데이터 표준이 추가되어야 한다는 필요성이 제기된다. 이미지 정보의 효율적인 디지털 아카이빙 및 활용을 위한 클래스 기반 다단계 메타데이터 모델 관리를 제안한다.

Study on Efficient Digital Archive Workflow and Metadata Representation for Image Data Management

Hyo-Won Kim*, Yong-Ik Yoon**

Abstract

As the amount of information increases explosively with the development of society, the need for a digital archive has emerged in order to maintain information efficiently. There is a current standard of digital archive is OAIS(Open Archival Information System) reference model. OAIS reference model is mainly interested in long-term preservation and concerned about comprehensive content. This paper propose digital archive workflow that is used not only for long-term preservation but also for the efficient utilization with the image data. The OAIS reference model and the metadata standard should be added surge information to build a digital archive. Therefore we propose metadata model should be managed based on multi-level classes not only for effective digital image archive but also for utilization.

Keywords : 디지털 아카이브, 메타데이터, 아카이브 워크플로우, 이미지

1. 서론

아카이브의 시작은 인간의 끊임없는 의사전달 활동을 통해서 이루어지는 정보와 그것을 기록하기 시작하면서 부터라고 할 수 있다. 사회가 발달함에 따라 정보의 양은 폭발적으로 증가하

게 되어 이를 유지, 보관하고 효율적인 이용을 위해 자료들을 체계적으로 분류하여 정리할 필요성이 제기되었다.

디지털 시대 이전의 아카이브는 별도의 보존 공간에서 관리되어, 상호간에 필요한 정보를 열람하거나 교환하기 위해서는 도서, 사진, 그림 등의 실물 자체의 정보로만 가능했다. 그러나 디지털 시대로 접어들면서 통신기술의 발달과 매체의 변화로 이용자가 필요한 정보를 즉각적으로 이용 가능한 상태로 서비스가 가능해졌다. 또한 그 정보를 서비스하기 위해서는 디지털로 변환되어 아카이브화 되어야만 서비스가 가능하게 되었다[1].

※ 제일저자(First Author) : 김효원
접수일:2008년 09월 17일, 완료일:2008년 12월 23일
* 숙명여자대학교
nnanda@sm.ac.kr
** 숙명여자대학교

디지털 아카이브란 아날로그 형태의 영상, 음향, 필름 등 다양한 자료들을 디지털 콘텐츠로 변환·저장·관리하여 편리한 검색 환경에서 콘텐츠 접근성을 획기적으로 향상시키며, 여러 전송매체에 손쉽게 제공할 수 있는 콘텐츠 보존 및 유통 기반 시스템이다[2].

기존의 연구는 포괄적인 콘텐츠를 대상으로 아카이브를 구축했기 때문에 이미지 자료를 위한 아카이브와 이미지 자료에 대한 메타데이터의 적용이 부족했다. 아카이브의 표준으로는 2002년 확정된 OAIS(Open Archival Information System) 참조 모형이 있는데 이것은 단계별로 체계화되어 정보가 저장되는 구조를 가지고 있다. 그 구조를 통해 이미지 자료의 디지털 아카이브를 위해서는 서지정보와 미디어 형태의 메타데이터 표준이 추가되어야 한다는 필요성을 발견하였다. 기존의 모형은 단순한 사실 기반의 메타데이터만 적용이 되어 있어 디지털 아카이브의 저장, 검색, 이용 등이 효율적이지 못했다. 기존의 “이미지 검색시스템을 위한 메타데이터 구축에 관한 연구”에서도 콘텐츠의 관리를 위한 방안은 언급하지 않고 있으며, 의미 기반의 메타데이터들이 표현되어 있지 않다. 단순한 사실 기반의 메타데이터 뿐만 아니라 이미지에 대한 의미 있는 정보를 표현할 수 있는 메타데이터의 추가를 본 논문에서 제안한다. 그러므로 본 논문에서는 클래스 기반 다단계 메타데이터 모델 관리를 제안한다. 1단계는 DC(DublinCore) 기반의 메타데이터, 2단계는 이미지에 대한 MPEG-7 기반 메타데이터, 3단계는 온톨로지 기반 시맨틱 메타데이터 사용을 제안한다. 단순한 텍스트 정보나 이미지 정보의 표현 뿐만 아니라 온톨로지 기반의 의미 있는 정보를 메타데이터에 추가한다[3][4][5]

2. 관련연구

아카이브에 대한 연구와 노력은 지금도 끊이지 않고 계속 되고 있다. 사회의 발달에 따라 정보의 양이 폭발적으로 증가하고 디지털화됨에 따라 디지털 아카이브의 필요성이 대두되었다. 본 장에서는 디지털 아카이브의 표준으로 자리 잡아 가는 OAIS(Open Archival Information Sy-

stem) 모델과 GPO(Government Printing Office)의 Future Digital System을 통해 디지털 아카이브 시스템의 현황을 파악한다. 또한 디지털 아카이브 구축 시 필수 요소인 메타데이터에 대해 살펴본다. 도서관에서 사용되는 서지 정보 표현을 위한 DublinCore와 멀티미디어 정보를 표현하기 위한 MPEG-7 메타데이터를 중심으로 살펴본다.

2.1 디지털 아카이브 시스템의 동향

디지털 아카이빙 시스템에 대한 국제 표준인 OAIS(Open Archival Information System) 참조 모형은 2002년 ISO 표준 14721로 확정되었다. OAIS 참조 모형의 확정은 디지털 정보의 영구 보존과 관련하여 각계 각 분야의 전문가들의 서로 다른 시각 속에서 합의점을 도출해 냈다는 데에 의의가 있으며, 향후 표준과 보존을 위한 기초 모형을 마련했다는 점에서 또한 의의가 있다. 그러나 OAIS 참조모형은 장기보존에만 관심을 가지고 있고, 주로 포괄적인 콘텐츠에만 신경을 쓰고 있다. 이러한 OAIS 참조모형을 기반으로 미국의 GPO(Government Printing Office)에서는 2005년 Future Digital System을 제안하여 사용하고 있다. 이것은 미국 내의 디지털 아카이빙 표준 스펙이지만 주로 동영상 기록물 관리에 중점을 두고 있다[6][7][8].

2.2 메타데이터

2.2.1 서지정보 표현(DublinCore)

DublinCore는 웹 자원을 위한 메타데이터이다. 주로 서지 정보에 사용되며 15가지 메타데이터 요소로 이루어져있다. 15가지 요소는 <표 1>에서 보여주고 있다.

<표 1> DublinCore의 15가지 요소

요소	설명	요소	설명	요소	설명
Title	제목	Contributor	공헌자	Source	소스
Creator	작가	Type	유형	Language	언어
Subject	주제	Date	날짜	Relation	관계
Description	기술	Format	포맷	Coverage	커버
Publisher	출판사	Identifier	식별자	Rights	권리

그러나 DublinCore는 텍스트 자료의 표현에 주안점을 두고 있다. 사실 정보만 입력하고 있어 의미 기반의 정보들이 표현되어 있지 않다 [9][10].

2.2.2 멀티미디어 정보 표현(MPEG-7)

MPEG-7은 이미지, 오디오 등의 멀티미디어 내용을 표현하기 위한 국제 표준이다. 표준화 대상은 Descriptor, Description Scheme, Description Decinition Language 그리고 기술 데이터의 효율적인 압축과 전송이다. MPEG-7의 part3에서 Visual Descriptor에 대해 정의하고 있다. <표 2>에서 Visual Descriptor의 요소에 대해 나타내고 있다. 그러나 MPEG-7의 메타데이터는 너무 방대하여 모든 요소 적용이 불가능하다. 또한 MPEG-7은 내용 표현에 치중하고 있어 이미지에 대한 사실 정보가 적고, 의미 기반의 내용이 없다[11][12].

<표 2> MPEG-7 part 3 Visual Descriptor

Visual Descriptor	요소
Color	Color space, Color Quantization, Dominant Color, Scalable Color, Color Layout, Color Structure, GoF/GoP Color
Texture	Homogenous Texture Descriptors, Homogenous Texture Descriptors, Homogenous Texture Descriptors, Edge Histogram
Shape	Region Shape, Contour Shape, Shape 3D
Motion	Camera Motion, Motion Trajectory, Parametric Motion, Parametric Motion

3. 디지털 아카이브 워크플로우

디지털 아카이브를 위해서는 먼저 원본 콘텐츠를 디지털로 바꾸는 과정이 요구된다. 또한 콘텐츠를 표준화 시켜서 저장해야 한다. 그리고 정보가 방대하므로 저장, 관리하기 위해 표준화 전

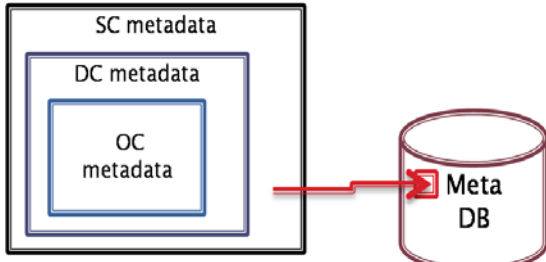
략이 요구되는데 OAIS와 GPO 시스템은 단계별로 체계화되어 정보가 저장되는 구조이다. 이미지 자료의 디지털 아카이브를 위해서는 그 구조에 도서관 서지정보와 멀티미디어 형태의 메타데이터 표준이 추가되어야 하는 필요성이 있다. 그러므로 디지털 아카이빙을 위한 새로운 메타데이터 모델이 필요하다. 본 논문에서는 다단계 클래스기반 메타데이터 모델 관리를 제시한다 [13].

3.1 Media Object의 정의

다단계 클래스기반 메타데이터 모델에 따른 디지털 아카이빙의 대상으로 Media Object를 정의한다. Media Object는 디지털화 시킬 콘텐츠 자체이며 등록, 저장, 보존, 유통되는 디지털 파일이다. Media Object는 텍스트, 이미지, 비디오, 사운드 등의 모든 미디어 콘텐츠를 포함할 수 있지만, 본 논문에서는 이미지 콘텐츠를 중심으로 정의하였다.

3.2 Media Object의 메타데이터

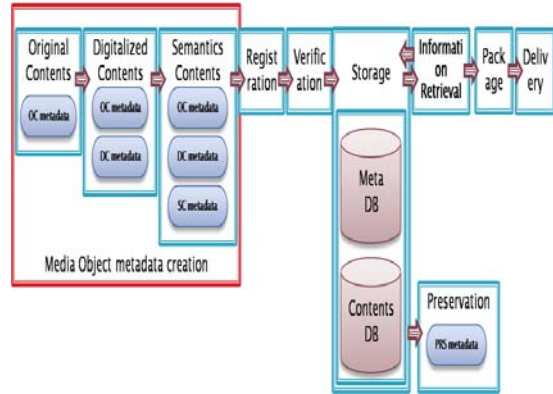
Media Object의 메타데이터는 다단계 클래스를 기반으로 하고 있다. Media Object는 사실기반, 내용표현 기반, 의미 기반의 3가지 구조를 가지고 있다. 1단계의 메타데이터는 Media Object가 가지고 있는 사실 정보에 대한 표현이다. 그리고 이것을 OC(Original Contents) 메타데이터라고 한다. 2단계의 메타데이터는 Media Object의 내용에 대한 표현으로써 DC(Digitalized Contents) 메타데이터라고 한다. 이는 Media Object를 디지털화 시킬 때 저장된다. 3단계의 메타데이터는 의미 정보를 생성하는 단계로써 SC(Semantics Contents) 메타데이터라고 한다. 최종 사용자가 등록할 때 생성되는 앞으로의 유통 및 사용자의 이용을 위한 Media Object의 의미정보를 생성하는 단계이다. (그림 1)에서는 Media Object의 ODS 메타데이터를 표현하고 있다. ODS 메타데이터는 누적메타데이터를 표현하는 것으로 1단계에서 부가한 메타데이터가 2단계에 누적되어 1단계 메타데이터를 가지고 있는 형태를 말한다. 그러므로 3단계 메타데이터는 1단계와 2단계에서 부가한 메타데이터를 가지고 있다[14].



(그림 1) Media Object의 ODS 메타데이터

3.3 Media Object의 워크플로우

3.2절에서 정의한 메타데이터의 생성은 (그림 2)에서 정의한 워크플로우를 따라 이루어진다. 1 단계는 Media Object를 위한 사용자가 정의한 Original Contents 정보를 선정하고 수집하는 단계이다. 그리고 이 단계에서 Media Object의 사실 정보에 관한 메타데이터가 생성된다. 2단계는 Media Object의 디지털 아카이브를 위해 Digitalized Contents로 변환시키는 단계이다. 디지털 변환시에 Media Object의 내용 표현에 관한 메타데이터가 생성된다. 3단계는 Semantics Contents로 Media Object에 대해 온톨로지의 특징에 따라 의미 기반의 메타데이터가 생성되는 단계이다. 4단계는 Registration으로 1단계, 2단계, 3단계의 메타데이터와 Media Object가 전부 등록이 되는 단계이다. 5단계는 Verification으로 Media Object를 선정, 수집, 디지털화, 등록 과정에서 발생한 메타데이터를 검수하고 품질평가가 이루어지는 단계이다. 6단계는 Storage로 Media Object의 메타데이터들과 디지털화된 Media Object 자체를 Database에 저장하는 단계이다. 7단계는 Preservation으로 Media Object의 장기보존을 위해 단계이다. 8단계는 InformationRetrieval으로 Media Object를 검색하기 위한 단계이다. 9단계는 Package로 Media Object와 필요 메타데이터들을 패키지화 시키는 단계이다. 마지막 단계는 Delivery로 Media Object를 사용자에게 전송하는 단계이다. 이러한 내용들이 디지털 아카이빙의 단계들이다[15].



(그림 2) 전체적인 워크플로우

4. 디지털 아카이브 관리와 메타데이터

4.1 Media Obejct Class 정의

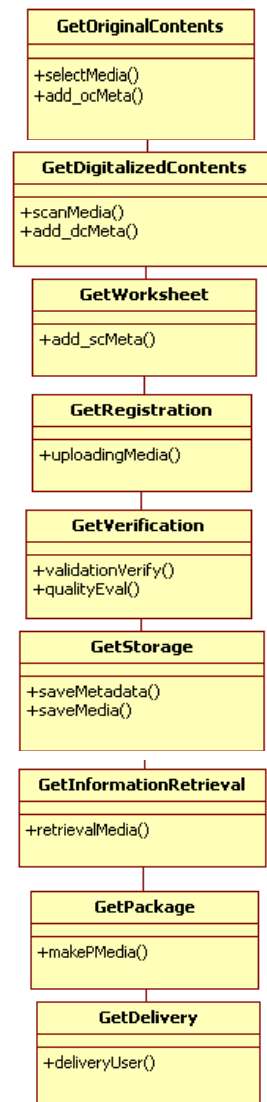
Media Object의 메타데이터를 관리하기 위해 객체 지향 개념을 기반으로 하여 Media Object Class를 정의한다. 여기서의 Class는 Media Object에 대해 일반적으로 적용할 수 있는 변수와 메소드를 정의하고 있는 소프트웨어적인 설계도 또는 프로토타입을 말한다. Media Object Class는 Media Object를 디지털 아카이빙 하는데 필요한 전반적인 프로세스들을 담당하는 Class이다. 객체 지향 개념으로 클래스를 이용하여 다른 클래스를 생성 또는 정의가 가능하기 때문에 Sub Class를 다음과 같이 정의한다. Media Object Class의 Sub Class는 GetOriginalContents, GetDigitalizedContents, GetSemanticsContents, GetRegistration, GetVerification, GetStorage, GetInformationRetrieval, GetPackage 그리고 GetDelivery 이다.

4.2 각 단계별 Class 정의

Media Object의 Sub Class들과 세부 메소드들에 대한 정의는 (그림 3)에서 보여주고 있다. Media Object를 수집, 선정하는 클래스는 GetOriginalContents로 selectMedia와 add_ocMeta 메소드를 가진다. selectMedia 메소드를 통해 이미 지 자료가 Media Object로 만들어지고 정보가 생성된다. 그리고 add_ocMeta을 통해 Media Ob

ject에 대한 사실 정보인 Title, Date, TopCategory, SubCategory, Actor, Producer, RegistWorker, RegistDate 그리고 RegistCom의 1단계 메타데이터들이 생성된다. GetOriginalContents에서 선정된 Media Object를 디지털화 시키는 GetDigitalizedContents 클래스는 scanMedia와 add_dcMeta 메소드를 가진다. scanMedia는 생성된 Media Object를 스캐닝하거나 이미지화 등을 통한 디지털변환 메소드이다. 이 과정에서 이미지 Media Object의 포맷, 해상도, 크기 등의 메타데이터가 추출된다. 그리고 add_dcMeta을 통해 scanMedia 과정에서 추출되어진 메타데이터요소와 결합이 되어서 Format, Resolution, Size, Color, Text 등의 2단계 메타데이터들이 생성된다. Media Object에 대해 워크시트에 메타데이터를 정의하고 1, 2, 3단계의 누적 메타데이터들을 관리하는 클래스는 GetSemanticsContents로 add_scMeta 메소드를 가진다. add_scMeta를 통해 사용자가 워크시트에 정의한 내용을 가져오고 Subject, MainDetail, SubDetail, Lesson의 시맨틱 3단계 메타데이터들이 생성된다. 디지털화된 Media Object를 디지털 아카이브 시스템에 등록하는 클래스는 GetRegistration으로 uploadingMedia 메소드를 가진다. 이전 단계 까지 생성되어 누적된 1, 2, 3단계의 메타데이터들과 Media Object를 업로드하는 메소드가 uploadingMedia 이다. 메타데이터를 검수하고 품질평가가 이루어지는 클래스는 GetVerification이고 validationVerify, qualityEval 메소드를 가진다. 선정, 수집, 디지털화, 등록 과정에서 발생한 메타데이터와 Media Object의 유효성을 검증하기 위한 메소드는 validationVerify이고, 메타데이터와 업로드된 Media Object의 품질을 평가하는 메소드는 qualityEval이다. Media Object의 저장을 위한 클래스는 GetStorage로 saveMetadata와 saveMedia 메소드를 가진다. 1, 2, 3단계에서 생성된 메타데이터들이 saveMetadata를 통해 Meta Database에 저장되고, 유효성 검증과 품질 평가를 마친 이미지 Media Object가 saveMedia를 통해 Contents Database에 저장된다. DB에 저장된 Media Object를 검색 하기 위한 클래스는 GetInformationRetrieval로 retrievalMedia 메소드를 가진다. 사용자가 입력한 메타데이터들을 통해 1, 2, 3단계의 누적 메타데이터들을 기반으로 Media Object를

검색하는 메소드가 retrievalMedia이다. 검색 방식은 일반검색, 통합 검색 모두 가능하다. Media Object의 패키지를 만들기 위한 클래스는 GetPackage로 makePMedia 메소드를 가진다. 요청한 사용자에게 의 필요에 따라 makePMedia를 통해 Media Object와 필요 메타데이터들이 사용자 배포용으로 패키지가된다. Media Object를 전송하는 클래스는 GetDelivery로 deliveryUser 메소드를 가진다. deliveryUser를 통해 GetPackage 클래스에서 만들어진 패키지 파일을 사용자에게 전달한다.



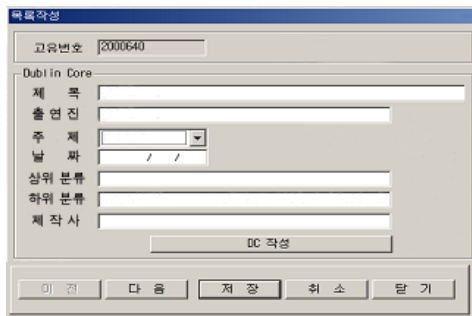
(그림 3) Media Object의 Sub Class 정의

5. 메타데이터 입력과 예제, 분석, 활용

본 장에서는 Media Object의 메타데이터 입력 화면과 디지털 아카이브 예제를 포스터와 사진으로 나누어 각각의 단계마다 생성되는 메타데이터 요소를 살펴보고, 이를 분석, 활용한다.

5.1 메타데이터 입력화면

DublinCore를 기반으로 한 1단계 메타데이터 입력화면은 (그림 4)와 같다. 제목과 출연진, 주제, 날짜, 상위 분류, 하위 분류, 제작사가 기본 입력이다. ‘다음’ 버튼을 클릭하면 2단계 메타데이터를 입력할 수 있는 단계가 나온다. 그리고 ‘저장’ 버튼을 클릭하여 각각의 단계마다 메타데이터를 저장 하고, ‘취소’ 버튼을 클릭하면 현재 창의 메타데이터가 초기화 되어 다시 입력 할 수 있다. ‘닫기’ 버튼을 통해 현재 창을 닫는다.



(그림 4) 1단계 메타데이터 입력 화면

5.2 Media Object의 아카이브 예(포스터)

본 절에서는 Media Object의 아카이브 예로 해리포터 포스터를 기반으로 살펴본다. Media Object의 메타데이터는 다단계 클래스기반 메타데이터라고 앞서 제시하였다. 1단계는 DublinCore를 기반으로 한 메타데이터이고, 2단계는 이미지 내용 표현을 위해 MPEG-7을 기반으로 한 메타데이터이고, 3단계는 의미 있는 정보 표현을 위한 온톨로지를 기반으로 한 메타데이터이다.

5.2.1 1단계 정보(DublinCore 기반)

<표 3> 포스터 1단계 메타데이터

1단계 add_ocMeta	
Title	HarryPotter
Date	2008-08-12
TopCategory	영화
SubCategory	가족영화
Actor	해리포터, 다니엘 제이콥 레드클리프, 헤르미온느 그레인저, 엠마 샬롯 듀어 왓슨, 론 위즐리, 루퍼트 알렉산드라 그린트
Producer	Warner Bros
RegistWorker	김효원
RegistDate	2008-08-13
RegisterCom	숙명여자대학교

1단계 메타데이터는 DublinCore를 기반으로 한 사실기반의 메타데이터로 add_ocMeta 메소드를 통해 생성된다. 포스터 예제에서는 Title, Date, TopCategory, SubCategory, Actor, Producer, RegistWorker, RegistDate, 그리고 RegisterCom가 1단계 메타데이터이다. 제목, 날짜, 상위 분류, 하위 분류, 출연진, 제작사, 등록자, 등록일 그리고 등록기관 메타데이터이다. Title, Date, TopCategory, SubCategory, RegistWorker, RegistDate, RegistCom 메타데이터는 포스터, 개인 사진, 행사 사진 등의 모든 Media Object에 적용될 수 있으므로 필수 메타데이터 요소로 지정한다. Actor, Producer는 포스터 이미지에만 필요하기 때문에 선택 메타데이터 요소로 지정한다. 그리고 포스터, 사진, 책표지 등의 이미지의 종류에 따라 메타데이터를 사용자가 정의해 추가시킬 수 있다. (그림 5)의 해리포터 포스터를 바탕으로 부가한 1단계 메타데이터는 <표 3>과 같다.

5.2.2 2단계 정보(MPEG-7 기반)

<표 4> 포스터 2단계 메타데이터

2단계 add_dcMeta	
Format	JPG
Resolution	50dpi
Size	349 x 519 pixel
Color	main color(blue), subcolor(black, yellow, white)
Text	12월, 14일, 마법이, 시작된다

Layout	가운데
Object	사람, 성, 텍스트
MCount	12개

2단계 메타데이터는 MPEG-7을 기반으로 한 효과적인 이미지 내용 표현을 위한 메타데이터이다. Media Object의 디지털화 과정에서 scan Media 메소드를 통해 메타데이터가 추출되고, add_dcMeta 메소드를 통해 추출된 메타데이터 요소와 결합하여 2단계 메타데이터가 생성된다. 포스터 예제에서는 Format, Resolution, Size, Color, Text, Layout, Object 그리고 MCount가 2단계 메타데이터이다. Format, Resolution, Size, Text는 Media Object의 디지털화 과정에서 추출 가능한 메타데이터이지만 Color, Layout, Object, MCount는 직접 부가해서 생성해야 하는 메타데이터이다. Color는 주요 칼라와 하위 칼라로 나누어지고, Object는 포스터의 사물, MCount는 주요 사물과 텍스트의 개수이다.

5.2.3 3단계 정보(온톨로지 기반)

<표 5> 포스터 3단계 메타데이터

3단계 add_scMeta	
Subject	마법, 동화, 꿈, 환상, 정의
MainDetail	주인공 부각, 왼쪽 시선, 노란색 텍스트
SubDetail	갈색 머리, 안경, 흰색 깃, 파란색 옷, 마법사 모자, 책, 긴머리, 기차
Lesson	상상력
Relationship	해리, 마법사 이야기, 호그와트, 마법나라, 지팡이, 망토

3단계 메타데이터는 온톨로지를 기반으로 한 메타데이터로 add_scMeta 메소드를 통해 메타데이터를 부가한다. Subject, MainDetail, SubDetail, Lesson 그리고 Relationship이 3단계 메타데이터이다. 해리포터의 주제를 Subject 메타데이터 요소에 부가한다. MainDetail 메타데이터 요소는 주인공이 부각되고 있는 모습, 주인공이 왼쪽으로 시선을 향하고 있는 점 등의 전체적인 이미지를 설명해준다. SubDetail 메타데이터 요소는 해리포터 주인공의 머리 색깔, 안경, 착용한 옷 등의 세부적인 이미지를 설명해준다. 사용자에게

교훈을 줄 수 있는 부분을 Lesson 메타데이터 요소에 부가하고, 사용자가 잘못 기재한 내용이나 유사, 관련 키워드는 Relationship 메타데이터 요소에 부가한다. 다음 (그림 5)가 해리포터 포스터 예제를 보여주고 있다.



(그림 5) 예제 1 포스터(harrypoter.jpg)

5.3 Media Object의 아카이브 예(사진)

본 절에서는 Media Object의 아카이브 예로 하이 서울 페스티벌 사진을 기반으로 살펴본다.

5.3.1 1단계 정보(DublinCore 기반)

<표 6> 행사사진 1단계 메타데이터

1단계 add_ocMeta	
Title	하이 서울 페스티벌 2008
Date	2008-08-17
TopCategory	행사
SubCategory	페스티벌
RegistWorker	김효원
RegistDate	2008-08-13
RegisterCom	국내 유일 민영통신사 NEWSIS

1단계 메타데이터는 DublinCore를 기반으로 한 사실기반 메타데이터로 add_ocMeta 메소드를 통해 메타데이터가 생성된다고 하였다. 사진 자료는 포스터와 달리 Actor, Producer 메타데이터 요소가 필요하지 않다. Title, Date, TopCategory, SubCategory, RegistWorker, RegistDate, 그리고 RegisterCom이 하이 서울 페스티벌 사진의 1단계 메타데이터 요소이다. (그림 6)의 하이 서울 페스티벌 사진을 바탕으로 부가한 1단계 메타데이터는 <표 6>과 같다.

5.3.2 2단계 정보(MPEG-7 기반)

<표 7> 행사사진 2단계 메타데이터

2단계 add_dcMeta	
Format	JPG
Resolution	50dpi
Size	533 x 361pixel
Color	main color(green, purple), subcolor(light green, gray, blue, yellow)
Text	newsis
Layout	양쪽
Object	사람, 풀, 강
Mcount	6개

2단계 메타데이터는 MPEG-7을 기반으로 한 효과적인 이미지 내용 표현을 위한 메타데이터로 add_dcMeta 메소드를 통해 메타데이터가 생성된다고 하였다. 사진 예제에서는 Format, Resolution, Size, Color, Text, Layout, Object 그리고 MCount가 2단계 메타데이터이다. 이것을 바탕으로 부가한 메타데이터는 <표 7>과 같다.

5.3.3 3단계 정보(온톨로지 기반)

<표 8> 행사사진 3단계 메타데이터

3단계 add_scMeta	
Subject	전통 행사, 폐막 행사
MainDetail	남자 주인공 두명, 야외, 의장대 행사, 창 찌르기, 구경꾼, 깃대, 나무, 아파트, 한강, 시민공원
SubDetail	서 있는 사람, 굽히고 있는 사람, 녹색옷, 붉은색 옷, 상투, 전통 옷
Relationship	행사 자료, 휴일 행사, 한강 근처 행사, 불거리, 이색 행사

3단계 메타데이터는 온톨로지를 기반으로 한 시맨틱 메타데이터로 add_scMeta 메소드를 통해 메타데이터를 부가한다고 하였다. Subject, Main Detail, SubDetail 그리고 Relationship이 3단계 메타데이터이다. 하이 서울 페스티벌 사진에 대한 사진의 주제 메타데이터 요소는 전통 행사, 폐막 행사를 부가한다. MainDetail 메타데이터 요소는 남자 주인공 두명, 야외, 의장대 행사, 창 찌르기, 나무, 아파트 등의 전반적인 이미지에

대한 설명이다. 그리고 SubDetail 메타데이터 요소는 서 있는 사람, 굽히고 있는 사람 등 이미지에서 나타나는 사람에 대한 행동들이나 옷 색깔 등의 세부적인 이미지에 대한 설명이 해당된다. 해당 사진과 연관 있는 키워드는 Relationship 메타데이터 요소에 부가한다. 다음 (그림 6)이 하이 서울 페스티벌 사진 예제이다.



(그림 6) 예제 2-2 사진(hiseoul.jpg)

5.4 분석 및 활용

기존의 디지털 아카이브 시스템의 표준인 OAIS 참조 모형은 장기 보존에만 관심이 있고, 주로 포괄적인 콘텐츠에만 신경을 쓰고 있었다. 본 논문에서는 이미지 자료의 디지털 아카이브를 위해 DublinCore 기반의 1단계 메타데이터, 이미지 내용 표현을 위한 MPEG-7기반의 2단계 메타데이터, 온톨로지 기반 시맨틱 정보 표현을 위한 3단계 메타데이터를 가진 다단계 클래스 기반 메타데이터 모델 관리를 제안하였다. 단순한 텍스트 정보나 이미지 내용 표현뿐만 아니라 온톨로지 기반의 의미 있는 정보를 메타데이터에 추가함으로써 디지털 아카이브의 관리가 용이하게 되고 사용자들을 위한 다양한 조건 검색이 가능하게 되었다.

예를 들어, “두 사람이 창을 가지고 찌르는 장면, 배경은 한강 시민 공원, 구경꾼들이 있는 전통행사 사진을 찾고 싶다.” 라고 했을 때, 이전 디지털 아카이브 시스템에서는 의미 정보에 관한 메타데이터가 부족하여 사용자가 원하는 사진을 검색하기 어려웠다. 하지만 본 논문의 디지털 아카이브 시스템은 사용자가 워크시트에 정의한 시맨틱 메타데이터를 중심으로 단순 나열식의 결과가 아닌 “두 사람”, “창 찌르기”, “구경꾼”, “한강 시민 공원”, “전통”, “행사”를 각각의 메타데이터로 인식하여 보다 구체적이고 사용자의 필요에 맞는 사진을 얻을 수 있게 되었다. 위

의 질의로 얻은 검색 결과는 (그림 7)과 같고 메타데이터를 표시하고 있다.



(그림 7) 사진 검색 결과

6. 결론

본 논문에서는 디지털 아카이브 시스템의 ISO 국제 표준인 OAIS 참조 모형을 기반으로 디지털 아카이브 워크플로우를 정리하였다. 또한 디지털 아카이브의 대상으로 Media Object를 정의하고 Media Object 중에서 이미지 자료를 중심으로 콘텐츠를 정의하였다. 이미지 자료의 디지털 아카이브를 위해서는 도서관 서지 정보와 멀티미디어 표준의 두 가지 전략이 합쳐져야 하는데, 단계별 메타데이터가 필요함을 제시하고, 구체적인 방안으로 클래스 기반 다단계 메타데이터 모델 관리를 제안하였다. 향후 디지털 아카이브 시스템은 콘텐츠 정보의 가치 상승, 필요한 정보의 상호 교환 등이 활발히 전개되는 창구 역할을 담당할 것이다. 그리고 더욱 유용하고 가치 있는 정보를 필요로 하는 모든 개인 및 단체에 이용될 것이다[16].

참 고 문 헌

[1] 현종철, "민주화운동 사진 디지털 아카이브즈 구축을 통한 역사정리", 한국기록학회 학술심포지엄 자료집, 123~142p, 2005
 [2] 조소연, "영상문화콘텐츠 디지털 아카이빙 도입 전략 방안", 영상문화콘텐츠 디지털 아카이빙 컨퍼런스, 2005
 [3] 남승희, 문성빈, "이미지 검색 시스템을 위한 메타데이터 구축에 관한 연구", 한국정보관리학회 제8회 학술대회 논문집, pp. 105 ~ 110, 2001
 [4] 박은경, "북미지역 디지털 아카이브즈 구축의 현황과

전망", 제7회 기록관리 워크샵, 2005
 [5] 최호남, 이응봉, "해외 전자저널의 디지털 아카이브 구축 전략에 관한 연구", 한국문헌정보학회지, 제39권 제 2호, 2005
 [6] 김희정, "디지털 아카이빙 최근 연구동향 및 OAIS 참조모형에 관한 연구", 한국기록관리학회지 제3권 제1호, 2003
 [7] CCSDS, "Reference Model for an Open Archival Information System(OAIS)," Red Book CCSDS, 650.0-R-1, 2001
 [8] Concept of operations for the future digital system(F Dsys) Final, May 16, 2005
 [9] 오동근, 메타데이터의 이해, 태일사, 2004
 [10] <http://uk.dublincore.org>
 [11] B.S. Manjunath, Phillipe Salembier and Thomas Sikora, "Introduction to MPEG-7", Wiley, 2002
 [12] 조윤희, "디지털콘텐츠 메타데이터 포맷의 비교 연구", 한국문헌정보학회지 제37권 제2호, 2003
 [13] 양영주, 윤용익, "모바일 서비스 적응성지원을 위한 미들웨어시스템 연구", 한국정보과학회 학술발표논문집, 제32권 제2호, pp.199-201, 2005
 [14] OkJoo Choi and YongIk Yoon, "A Meta Data Model of Context Information for Dynamic Service Adaptation on User Centric Environment", Interactive Multimedia & Intelligent Services in Mobile and Ubiquitous Computing, 2007
 [15] 이수상, "디지털 아카이빙의 워크플로우와 보존처리 기술에 관한 연구", 한국도서관 정보학회지 제35권 제 3호, 2004
 [16] 김동진, 최운호, "관소리 자료의 디지털 아카이브 구축 현황과 방안, 그리고 전망", 관소리 연구 제25집, 2008
 [17] Y. Yoon, S. Kim, and J. W. Lee, "Universal Video Adaptation Model for Contents Delivery in Ubiquitous Computing", The 8th Asia-Pacific Conference on Computer Human Interaction, 2008

김 효 원



2007년 : 숙명여자대학교 문헌정보
학과 학사

2007년~현재 : 숙명여자대학교
멀티미디어 콘텐츠 석사
재학중

2007년~2008년 : 숙명여자대학교 도서관 연구조교
관심분야 : 디지털 아카이빙, 디지털 라이브러리,
멀티미디어 시스템, 멀티미디어 콘텐츠 저장
시스템

윤 용 익



1985년 : 한국과학기술원 전산학과
석사(전산학)

1994년 : 한국과학기술원 전산학과
박사(전산학)

1985년~1997년 : 한국전자통신연구원 책임연구원
2004년~2007년 : 미국 University of Colorado Visiti
ng Professor
1997년~현재 숙명여자대학교 정보과학부 교수
관심분야 : 미들웨어, 멀티미디어 시스템, 영상 아카
이빙 시스템, 콘텐츠 전달 시스템, 디지털 라
이브러리