

이미지 자료의 디지털 아카이빙과 메타데이터 설계

김효원*, 윤용익

숙명여자대학교 멀티미디어 과학과

e-mail : nnanda@sookmyung.ac.kr, yiyoon@sookmyung.ac.kr

Digital Archiving and Design of Metadata for Image

Hyo-Won Kim, Yong-Ik Yoon

Dept of Multimedia Science, Sook-myung Women's University

요 약

디지털 시대에 들어서면서 폭발적으로 증가하는 정보를 효율적으로 유지, 보관, 이용하기 위해 디지털 아카이빙의 필요성이 제기되었다. 현재 아카이브의 표준으로는 OAIS(Open Archival Information System)이 있지만 주로 포괄적인 콘텐츠를 대상으로 하고 있다. 본 논문에서는 이미지 자료를 대상으로 한 디지털 아카이빙과 메타데이터를 설계하고자 한다. 이미지 자료의 효율적인 디지털 아카이빙 구축을 위해서는 참조 모형을 통해 서지정보와 멀티미디어 자료에 대한 메타데이터 표준이 추가되어야 한다. 구체적인 방안으로 클래스 기반 다단계 메타데이터 모델 관리를 제안한다.

1. 서론

디지털 시대로 접어들면서 정보가 폭발적으로 증가하게 되어 이를 효율적으로 유지, 보관, 이용하기 위해서는 자료들을 체계적으로 분류하여 정리할 필요성이 제기되었다. 이를 위한 디지털 아카이빙은 아날로그 형태의 영상, 음향, 필름 등 다양한 자료들을 디지털 콘텐츠로 변환·저장·관리하는 활동들을 의미한다. 디지털 아카이빙된 자료들의 체계적인 관리와 효율적인 이용을 위해서는 메타데이터 요소 설계가 중요하다.

본 논문에서는 이미지 자료를 대상으로 한 디지털 아카이빙과 메타데이터를 설계하고자 한다. 기존의 연구는 포괄적인 콘텐츠를 대상으로 아카이브를 구축했기 때문에 이미지 자료를 위한 아카이빙과 이미지 자료에 대한 메타데이터의 적용이 부족했다. 아카이빙 시스템의 표준으로는 2002년 확정된 OAIS(Open Archival Information System) 참조 모형이 있다. 기존에는 단순한 사실 기반의 메타데이터만 적용이 되어 있어 디지털 아카이빙의 저장, 검색, 이용 등이 효율적이지 못했다. 그러므로 단순한 사실 기반의 메타데이터 뿐만 아니라 이미지에 대한 내용과 의미 있는 정보를 표현할 수 있는 메타데이터의 추가를 본 논문에서 제안한다. 구체적인 방안으로 클래스 기반 다단계 Metadata 모델 관리를 제안한다.

2. 관련 연구

2.1 디지털 아카이브 시스템의 동향

디지털 아카이빙 시스템에 대한 국제 표준인 OAIS(Open Archival Information System) 참조모형은 2002년 ISO 표준 14721로 확정되었다. OAIS 참조 모형의 확정은 디지털 정보의 영구 보존과 관련하여 각계 각 분야의 전문가들의 서로 다른 시각 속에서 향후 표준과 보존을 위한 기초 모형을 마련했다는 점에서 또한 의의가 있다. 그러나 OAIS 참조모형은 장기보존에만 관심을 가지고 있고, 주로 포괄적인 콘텐츠에만 신경을 쓰고 있다.

2.2 메타데이터

2.2.1 서지정보 표현(DublinCore)

<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

DublinCore는 웹 정보자원의 관리를 위한 메타데이터로 서지 정보에 주로 사용된다. 그 15가지 요소는 표 1에서 나타내고 있다. 그러나 DublinCore는 텍스트 자료의 표현에 주안점을 두고 있다. 사실 정보만 입력하고 있어 이미지에 대한 내용과 의미 기반의 정보들이 표현되어 있지 않다.

* 본 논문은 지식경제부 한국산업기술평가원 지정 한국항공대학교 부설 인터넷정보검색 연구센터의 지원에 의함

2.2.2 멀티미디어 정보 표현(MPEG-7)

<표 2> MPEG-7 part4 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

MPEG-7은 이미지, 오디오 등의 멀티미디어 내용을 표현하기 위한 국제 표준이다. MPEG-7의 part4에서 Visual Descriptor에 대해 정의하고 있다. 표 2에서 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 정보와 의미 기반의 내용이 적다.

3. 디지털 아카이빙 워크플로우

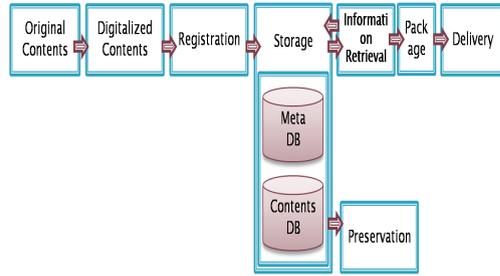
3.1 Digital Object의 정의

다단계 클래스기반 메타데이터 모델에 따른 디지털 아카이빙의 대상으로 Digital Object를 정의한다. Digital Object는 콘텐츠이며 등록, 저장, 유통되는 디지털 파일이다. 본 논문에서는 이미지 콘텐츠를 대상으로 정의하였다.

3.2 Digital Object의 워크플로우

Digital Object의 디지털 아카이빙을 위한 1단계는 Original Contents 정보를 선정하고 수집하는 단계이다. 2단계는 Digital Object의 디지털 아카이브를 위해 Digitalized Contents로 변환시키는 단계이다. 3단계는 Registration으로 메타데이터와 Digital Object가 등록되는 단계이다. 4단계는 Storage로 디지털화된 Digital Object 자체와 메타데이터들을 Database에 저장하는 단계이다. 5단계는 Preservation으로 Digital Object의 장기보존을 위한 단계이다. 6단계는 Information Rerieval으로 Digital Object를 검색하기 위한 단계이다. 7단계는 Package로 Digital Object와 필요 메타데이터들을 패키지

화 시키는 단계이다. 마지막 단계는 Delivery로 Digital Object를 사용자에게 전송하는 단계이다. 그림 1에서 이 워크플로우를 보여주고 있다.



(그림 1) 전체적인 워크플로우

3.3 Digital Object의 메타데이터

Digital Object의 메타데이터는 다단계 클래스를 기반으로 하고 있다. Digital Object는 사실기반, 내용표현기반, 의미기반의 3가지 구조를 가진다. 1단계의 메타데이터는 Digital Object가 가지고 있는 사실 정보에 대한 표현이다. 2단계의 메타데이터는 Digital Object의 내용에 대한 표현이다. 3단계의 메타데이터는 의미 정보를 생성하는 단계이다. 최종 사용자가 등록할 때 생성되는 앞으로의 유통 및 사용자의 이용을 위한 Digital Object의 의미정보를 생성하는 단계이다.

4. 디지털 아카이빙 관리와 메타데이터

4.1 Digital Object Class 정의

Digital Object의 메타데이터를 관리하기 위해 객체 지향 개념을 기반으로 Digital Object Class를 정의한다. Digital Object Class는 Digital Object를 디지털 아카이빙하는데 필요한 전반적인 프로세스들을 담당하는 Class이다. 그리고 다음과 같이 Sub Class를 정의한다. Digital Object Class의 Sub Class는 GetOriginalContents, GetDigitalizedContents, GetRegistration, GetStorage, GetInformationRetrieval, GetPackage 그리고 GetDelivery이다.

4.2 Digital Obejct의 Sub Class

Digital Object를 수집, 선정하는 클래스는 GetOriginalContents로 selectDigital과 add_ocMeta 메소드를 가진다. selectDigital 메소드를 통해 이미지 자료가 Digital Object로 만들어지고 1단계 메타데이터들이 생성된다. GetOriginalContents에서 선정된 Digital Object를 디지털화 시키는 GetDigitalizedContents 클래스는 scanDigital와 add_dcMeta 메소드를 가진다. scanDigital는 생성된 Digital Object를 디지털로 변환하는 메소드이다. 이 과정에서 이미지 Digital Object의 포맷, 해상도, 크기 등의 메타데이터가 추출된다. 그리고 add_dcMeta을 통해 scanDigital 과정에서 추출되어진 메타데이터요소와 결

합이 되어서 2단계 메타데이터들이 생성된다. 디지털화된 Digital Object를 디지털 아카이브 시스템에 등록하는 클래스는 GetRegistration으로 add_scMeta와 uploadingDigital 메소드를 가진다. add_scMeta를 통해 3단계 메타데이터들이 생성된다. 그리고 이전 단계 까지 생성된 1, 2, 3단계의 메타데이터들과 Digital Object를 업로드하는 메소드가 uploadingDigital 이다. Digital Object의 저장을 위한 클래스는 GetStorage로 saveMetadata와 saveDigital 메소드를 가진다. 1, 2, 3단계에서 생성된 메타데이터들이 saveMetadata를 통해 Meta Database에 저장되고, 이미지 Digital Object가 saveDigital를 통해 Contents Database에 저장된다. Database에 저장된 Digital Object를 검색하기 위한 클래스는 GetInformationRetrieval로 retrievalDigital 메소드를 가진다. 사용자가 입력한 메타데이터들을 통해 1, 2, 3단계의 메타데이터들을 기반으로 Digital Object를 검색하는 메소드가 retrievalDigital이다. Digital Object의 패키지를 만들기 위한 클래스는 GetPackage로 makePDigital 메소드를 가진다. 요청한 사용자의 필요에 따라 makePDigital를 통해 Digital Object와 필요 메타데이터들이 사용자 배포용으로 패키지화된다. Digital Object를 전송하는 클래스는 GetDelivery로 deliveryUser 메소드를 가진다. deliveryUser를 통해 GetPackage 클래스에서 만들어진 패키지 파일을 사용자에게 전달한다.

5. 디지털 아카이빙 예

5.1 포스터 예제

5.1.1 1단계 정보

본 절에서는 Digital Object의 아카이브 예로 “드리머 포스터”를 기반으로 살펴본다.

<표 3> 1단계 메타데이터

1단계 add_ocMeta	
Title	Dreamer
Date	2006-43-13
TopCategory	영화
SubCategory	드라마
Actor	다코다 패닝, 키트 러셀, 엘리자베스 슈, 크리스 크리스토퍼슨, 데이비드 모스, 프레디 로드리게즈, 루이스 구즈만
Producer	Hyde Park Entertainment, DreamWorks productions LLC
Nation	America
RegistWorker	김효원
RegistDate	2008-09-20
RegisterCom	숙명여자대학교

1단계 메타데이터는 DublinCore를 기반으로 한 사실기반의 메타데이터로 add_ocMeta 메소드를 통해 생성된다. 제목, 날짜, 상위 분류, 하위 분류, 출연진, 제작사, 국가, 등록자, 등록일 그리고 등록기관 메타데이터이다. 그림 2의 드리머 포스터를 바탕으로 부가한 1단계 메타데이터는 표 3과 같다.

5.1.2 2단계 정보

<표 4> 2단계 메타데이터

2단계 add_dcMeta	
Format	JPG
Resolution	60dpi
Size	350 x 560 pixel
Color	main color(green), subcolor(yellow, black, white, red, blue)
Text	다코다패닝, 11살, 소녀가, 만드는, 위대한, 감동, 실화, 드리머, 2006년, 봄, 꿈의, 절주가, 시작된다
Layout	오른쪽, 위
Object	사람, 말, 나무, 울타리, 텍스트
MCount	9개

2단계 메타데이터는 MPEG-7을 기반으로 한 이미지 내용 표현을 위한 메타데이터이다. Digital Object의 디지털화 과정에서 scanMedia 메소드를 통해 메타데이터가 추출되고, add_dcMeta 메소드를 통해 추출된 메타데이터 요소와 결합하여 2단계 메타데이터가 생성된다. 포스터 예제에서는 포맷, 해상도, 사이즈, 칼라, 텍스트, 위치, 포스터에서 표현하는 사물을 나타내는 Object, 주요 사물과 텍스트의 개수를 나타내는 MCount가 2단계 메타데이터이다. 그림 2의 해리포터 포스터를 바탕으로 부가한 2단계 메타데이터는 표 3과 같다.

5.1.3 3단계 정보

<표 5> 3단계 메타데이터

3단계 add_scMeta	
Subject	가족, 꿈,
MainDetail	주인공 부각, 왼쪽 시선, 녹색 텍스트, 말의 옆모습, 나무, 노을, 들판, 걸여가는 소녀
SubDetail	갈색 머리, 파란 눈, 흰색 옷, 턱을 괴고 있는 모습
Lesson	가족의 사랑
Relationship	실화, 아름다운 소녀, 가족애, 감동 있는 영화

3단계 메타데이터는 온톨로지를 기반으로 한 메타데이터로 add_scMeta 메소드를 통해 메타데이터를 부가한다. Subject, MainDetail, SubDetail, Lesson 그리고 Relationship이 3단계 메타데이터이다. 그림 2의 해리포터 포스터를 바탕으로 부가한 3단계 메타데이터는 표 3과 같다. MainDetail 메타데이터 요소는 주인공의 모습과 전체적인 이미지를 설명해준다. SubDetail 메타데이터 요소는 주인공의 세부적인 이미지를 설명해준다. 사용자에게 주는 교훈을 Lesson 메타데이터 요소에 부가하고, 사용자가 잘못 기재한 내용이나 유사, 관련 키워드는 Relationship 메타데이터 요소에 부가한다. 다음 그림 2가 드리머 포스터를 보여주고 있다.



(그림 2) Dreamer

6. 결론

기존의 디지털 아카이브 시스템의 표준인 OAIS 참조 모형은 장기 보존에만 관심이 있고, 주로 포괄적인 콘텐츠에만 신경을 쓰고 있다고 했다. 본 논문에서는 이미지 자료의 디지털 아카이빙을 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 Metadata 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. 또한 사회전반에 폭넓은 양질의 이미지 정보를 많은 사용자에게 서비스하여 자체 정보력의 역량을 강화하고 나아가 콘텐츠의 경쟁력 향상에 있어서 중요한 역할을 할 수 있게 되었다. 또한 디지털 아카이빙 데이터의 활용 빈도 증가, 보관 공간 및 관리 비용감소, 디지털 콘텐츠에 대한 신속하고도 완전한 공유가 가능하게 되었다.

참고문헌

- [1] 이수상, “디지털 아카이빙의 워크플로우와 보존처리 기술에 관한 연구”, 한국도서관 정보학회지 제35권 제3호, 2004
- [2] 박은경, “디지털라이제이션의 디지털이미지와 메타데이터 표준화작업”, 한국기록관리학회지 제5권 제2호, 2005
- [3] 조소연, “영상문화콘텐츠 디지털 아카이빙 도입 전략 방안”, 영상문화콘텐츠 디지털 아카이빙 컨퍼런스, 2005
- [4] 오삼균, 권기성, “전통예술자료 아카이빙을 위한 보존 메타데이터 설계에 관한 연구”, 한국문헌정보학회지 제41권 제4호, 2007
- [5] CCSDS, Reference Model for an Open Archival Information System(OAIS),” Red Book CCSDS, 650.0-R-1, 2001
- [6] B.S. Manjunath, Phillipe Salembier and Thomas Sikora, “Introduction to MPEG-7”, Wiley, 2002
- [7] <http://uk.dublincore.org>