

문화콘텐츠 통합을 위한 메타데이터 포맷 연구*

A Study on Metadata Formats for Integration of Cultural Contents

조 윤 희(Yoon-Hee Cho)**

초 록

최근 문화콘텐츠 관련 기관들은 분산 네트워크를 통해 이용자에게 문화콘텐츠의 접근을 점차 확대시키고 있다. 그러나 문화콘텐츠의 특성상 일반 콘텐츠와 달라서 많은 문화콘텐츠 객체는 정보의 조직과 검색에 사용할 수 있는 단어를 거의 포함하고 있지 못하다. 이러한 문화콘텐츠 시스템이 서비스 될 경우, 효율적으로 자원을 식별하고 검색할 수 없게 될 뿐 아니라 다양한 문화콘텐츠 메타데이터 요소간의 명칭, 표현, 의미가 상이하여 시스템간 연동이나 정보공유가 어렵게 된다. 이러한 문제를 해결하기 위해 문화콘텐츠의 속성에 적합한 표준화된 메타데이터의 조직과 관리가 필요하다. 본 연구는 문화콘텐츠 분야에서 다양하게 접근되고 있는 메타데이터 포맷인 Dublin Core, EAD, VRA, CDWA, CIMI, Object ID를 중심으로 각 포맷의 데이터 요소를 비교 분석함으로써 상이한 메타데이터 포맷의 상호운용성 확보를 통하여 문화콘텐츠 통합을 위한 기초 자료를 제공하고자 하였다.

ABSTRACT

Recently, the organizations related to cultural contents are gradually expanding access to cultural contents for general public through the distributed network. However, since cultural contents have different characteristics than general contents, the objects of cultural contents seldom contain the words generally used for organization and search of information. If the cultural contents system is created without any consideration of such differences, we cannot effectively identify and search resources. Moreover, because the names, expressions and meanings are different between metadata elements of various cultural contents, it is very difficult to interconnect or share information between different systems. In order to solve these problems, proper organization and management of metadata is vital. In this study, we have comparatively analyzed the data elements of each format based on Dublin Core, EAD, VRA, CDWA, CIMI, and Object ID, the metadata formats approached from various aspects in the cultural contents area. Through this study, we tried to provide the basic materials for integration of cultural contents by securing interoperability of different metadata formats.

키워드: 문화콘텐츠, 메타데이터, 더블린 코어, 데이터 요소, Cultural Contents, Metadata, Dublin Core, EAD, VRA, CDWA, CIMI, Object ID

* 이 논문은 2002년도 한국학술진흥재단의 지원에 의하여 연구되었음(KRF-2002-075-H00002)

** 중앙대학교 문화유산디지털연구센터 전임연구교수(cho519@wm.cau.ac.kr)

■ 논문 접수일 : 2003. 5. 21

■ 게재 확정일 : 2003. 6. 3

1 서 론

1.1 연구의 필요성 및 목적

문화콘텐츠는 텍스트, 이미지, 사운드, 그래픽, 애니메이션, 동영상 등 다양한 형태의 콘텐츠로 구성되어 있다. 이러한 다양한 형태의 디지털 콘텐츠는 시스템에 따라 데이터의 요소, 명칭, 길이, 표현, 의미가 상이하게 구축되어 문화콘텐츠 정보 시스템들간 연동이나 정보공유가 어렵다.

우리 나라는 21세기 핵심산업으로 문화콘텐츠 산업에 투자와 관심이 집중되고 있다. 이에 정부기관, 연구소, 단체 및 업체 등에서는 분야별로 전문적이고 질 높은 문화콘텐츠의 개발과 이를 효율적으로 검색할 수 있는 지식정보시스템 구축을 활발히 진행하고 있다. 그러나 이들 각 문화콘텐츠 시스템에 저장된 디지털 객체의 식별과 접근을 보장하는 메타데이터의 표준화가 선행되지 않아 요소의 명칭이나 길이, 표현, 의미 관계 등이 다양하게 표현되어 향후 시스템간 연동이나 통합에 문제점으로 대두되고 있다.

현재 문화콘텐츠분야 메타데이터는 표준화된 단일 포맷이 사용되는 것이 아니라 여러 기관들이 개발한 다양한 포맷이 사용되고 있다. 특히, 문화콘텐츠는 특성상 일반적인 콘텐츠와 다르게 모든 메타데이터 포맷의 필수 요소로 채택되어 있는 표제와 저자에 해당하는 내용도 포함하고 있지 않은 경우도 있다.

문화콘텐츠의 효율적인 조직과 검색을 위해서 메타데이터는 반드시 작성되어야 한다. 그러나 이들 메타데이터는 객체가 속한 집단의 목적이나 자원의 속성에 따라 다양하게 표현되고 있다. 이에 따라 공통의 표준을 마련하는 것이 필요하나는 비용과 시간 낭비일 뿐만 아니라 합의가 어려운 문제가 있다.

메타데이터의 정보를 조직하고 관리하는데 시간과 비용을 최소화하면서 이용자에게 광범위한 서비스를 제공하기 위한 방법론으로 다양한 포맷간 메타데이터 요소들을 참조할 수 있는 참조테이블(crosswalk)이 작성되었다.

참조테이블은 하나의 메타데이터 포맷과 다른 메타데이터 포맷을 매핑하기 위한 명세이다. 즉, 하나의 메타데이터 표준에서 정의한 요소가 다른 메타데이터 포맷의 관련 요소로 참조하여 사용할 수 있도록 하는 것이다. 이는 다양한 메타데이터 포맷의 유사성을 분석하고 매핑하는 방법론으로써 메타데이터 포맷간의 차이점을 조정하고 상호운용성을 확보할 수 있도록 하는 중요한 매카니즘이다.

이러한 참조테이블을 통한 매핑이 일관성을 유지하면서 객관적으로 수행된다면, 문화콘텐츠의 효율적인 메타데이터 조직과 이를 통한 검색의 효율성을 제고시킬 수 있을 것이다.

이에 본 연구에서는 다양한 문화콘텐츠 메타데이터의 포맷을 검토하고, 각 포맷의 요소를 비교 분석하고자 한다. 이를

통하여 각 포맷간의 특성을 파악하고, 나아가 다양한 형태의 문화콘텐츠 메타데이터 요소로서 적합한 최소 수준의 메타데이터 요소의 선정과 포맷간의 관련성을 규명하여 문화콘텐츠 통합의 기초자료를 제공하는데 목적이 있다.

1.2 연구의 범위 및 방법

본 연구는 분산 네트워크 환경하에 급격히 증가하고 있는 문화콘텐츠의 메타데이터 포맷에 관한 비교 분석 연구이다. 문화콘텐츠 분야는 관련 기관의 서로 다른 관점으로 여러 개의 메타데이터 포맷이 개발되고 사용되어 왔다. 최근 이들 메타데이터는 상호운용성을 확보하기 위해 참조데이터를 통한 매핑작업들이 활발하게 이루어지고 있다.

본 연구는 세계적으로 널리 사용되고 있는 문화콘텐츠의 메타데이터 포맷으로 적용되고 있는 더블린 코어, EAD, VRA, CDWA, AMICO, CIMI, Object ID를 중심으로 각 포맷에서 제공하고 있는 요소들을 비교 분석하였다. 메타데이터 포맷은 Gilliland-Swetland(1998)가 데이터 요소의 기능적인 측면을 고려하여 범주화한 관리, 설명, 보존, 기술, 이용 등 5가지 유형으로 분류하고, 포맷간 메타데이터 요소의 명칭을 중심으로 비교 분석하였다.

본 연구의 방법은 문화콘텐츠의 효율적인 정보조직과 색인 및 검색에 관련되는 메타데이터 포맷을 자료 조사를 통하여 이론적 검토를 수행하였다. 아울러 문화

콘텐츠의 메타데이터 표준으로 널리 사용되고 있는 다섯 개의 포맷을 선정하여 각 포맷에서 제공하는 요소들을 중심으로 비교 분석을 수행하여, 최소 수준의 메타데이터 요소를 선정하였다.

2 문화콘텐츠와 메타데이터

2.1 문화콘텐츠

문화는 동일한 공간과 시간 속에 공존하는 공동체 구성원들이 공동생활을 영위하면서 일정한 목적이나 생활의 이상을 구현하려는 활동 및 활동의 결과물로 생산된다. 따라서 역사성과 동시성을 향유하는 공동체 구성원들의 삶을 영위하기 위한 전반적 방식인 동시에 역사적으로 구현된 물질적 정신적 구성물이라고 할 수 있다(황동열 외 2003).

콘텐츠(contents)란 영어 content의 복수형으로 일반적으로 단순히 내용이나 문서의 내용을 소개하는 목차의 의미로 사용되어 왔다. 그러나 최근 디지털콘텐츠의 출현과 함께 텍스트, 이미지, 동영상, 음악파일, 데이터베이스의 정보 등을 통칭하여 사용되고 있다. 또한 그 장르가 영화든 문학이든 학습이든 뉴스든 오락이든 간에 기획이나 창작, 혹은 가공이나 개발을 누가 했는지가 분명하게 나타나서 추후에 저작권을 주장할 수 있는 모든 종류의 원작을 콘텐츠로 정의하고 있다(심상민 2002).

문화산업진흥법(2001)에서는 부호, 문자, 음성, 음향 및 영상으로 표현된 모든 종류의 자료 또는 지식 및 이들의 집합물을 콘텐츠로 정의하고 있다.

문화콘텐츠란 문화가 멀티미디어 정보 기술과 접목되는 현상에 주목한 개념으로 문화산업 가운데 핵심적 내용을 그 대상으로 한다(한상근 외 2002). 문화콘텐츠진흥원(2002)은 음반, 방송, 게임, 애니메이션, 캐릭터, 만화 등의 장르로 문화콘텐츠를 정의하였다. 문화관광부(2002)는 <그림 1>에서 보는 바와 같이 디지털기술의 발전이 기존의 문화콘텐츠(출판, 영상, 게임, 음반 등)의 제작과 유통 측면에 영향을 미치게 되어 새로이 디지털형태로 제작, 유통, 소비되고 있는 것을 문화콘텐츠라고 정의하였다. 황동열 외(2003)는 산업적 측면에서 문화콘텐츠는 문화적 요소가 체화되어 경제적 부가가치를 창출하는 디지털콘텐츠라고 정의하고 있다.

이상과 같이 최근 정보기술의 발전과

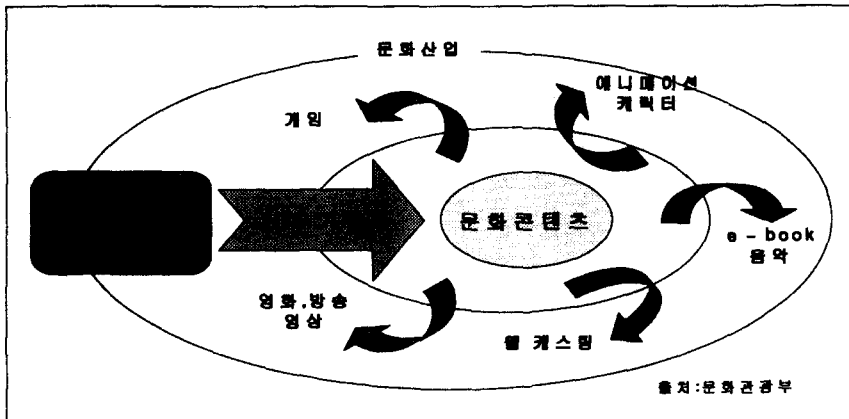
함께 출현한 문화콘텐츠는 기본적으로 문화창의력과 디지털기술의 접목이라는 기본 개념으로 CT(Culture & Contents Technology)로 귀착되고 있다.

종합하여 본 연구는 수 세대를 통해 인류 공동 구성원들의 유산으로 축적된 문화 객체로서 텍스트, 이미지, 오디오, 비디오, 동영상으로 표현된 인류의 물질적 정신적 구성물을 총칭하여 문화콘텐츠로 정의한다.

2.2 메타데이터의 필요성

메타데이터는 데이터에 관한 구조적 데이터로서 자원에 대한 정보를 제공하는 데이터라고 정의할 수 있다. 따라서 네트워크 기반 정보조직의 효율성은 메타데이터의 관리와 조직에 달려있다고 할 수 있다.

특히 도서관, 도큐멘테이션 센터, 박물관, 미술관 등은 문화 유산을 보존하고 연구하며, 정보를 제공하는 중심적인 역



<그림 1> 문화콘텐츠의 개념도

할을 하는 문화콘텐츠 관련 기관이다. 이러한 기관에서는 자원 자체에 대한 정보 수집과 저장 및 도큐멘테이션 과정을 거쳐 보존의 기능을 수행한다. 즉, 문화콘텐츠란 도큐멘테이션에서부터 연구에 이르기까지의 과정에서 파생되는 결과로 콘텐츠와 본래의 객체(original object)가 담고 있는 특별한 정보를 의미한다. 이러한 과정의 결과로서 생성되는 콘텐츠는 다양한 형식으로 표현된다. 목록과 같은 리스트, 사진과 회화와 같은 이차원 영상, 그리고 영화나 비디오와 같은 동영상 등의 콘텐츠가 그 예가 된다.

최근 문화콘텐츠 관련 기관들은 분산 네트워크를 통해 문화콘텐츠의 접근을 점차 확대시키고 있다. 그러나 문화콘텐츠의 특성상 일반적인 콘텐츠와 다르게 접근과 검색에 여러 문제점이 나타나고 있다. 이에 Zang(1999)은 박물관 객체와 일반 서지 자료를 비교하였다. 그는 연구의 결과에서 박물관 객체는 대부분 정보의 조직과 검색에 사용할 수 있는 단어를 거의 포함하고 있지 않음을 발견하였다. 또한 일반적으로 모든 메타데이터 포맷의 필수 요소로 채택되어 있는 표제와 저자에 해당하는 내용도 포함하고 있지 않은 경우가 많다고 지적하였다. 아울러 객체의 이름, 제작 이유, 시기, 제작자에 대한 모든 정보는 객체에 수반되는 자료, 참고 자료, 전문가 및 목록작성자와 같이 외부 자료나 관련 전문가를 통해서 얻고 있음을 지적하였다.

문화콘텐츠의 한 유형인 미술자료의 경우에도 이러한 문제점이 나타났다(Stephenson 1999). 미술자료의 이미지 데이터베이스는 이미지를 관리하고 기술하며, 식별하기 위한 도큐멘테이션 과정이 필요하다. 그러나 이러한 도큐멘테이션 수준은 이용자의 필요와 기관 자체의 정책에 따라 매우 다양하게 나타난다. 이러한 경우에 가장 중요한 것은 이미지 데이터베이스에 대한 도큐멘테이션 과정에서 원작품(original work)과 그 작품의 사진 복제물(photographic representation) 및 디지털화된 이미지의 기술적인 특성을 구별할 수 있어야 한다는 점이다(Besser 1995).

최근 문화콘텐츠와 관련하여 디지털콘텐츠의 양적 성장은 기하급수적으로 증가하고 있다. 그러나 정보의 일관성은 그 정보의 양에 비례하여 감소하고 있다. 대부분의 문화콘텐츠 정보는 도서관 목록, 박물관 소장품관리시스템과 같은 구조화된 데이터베이스내에 존재한다. MARC 형식을 이용하는 도서관의 상황과는 다르게 문화콘텐츠 관련 기관들은 비일관적인 데이터 구조와 값에 대한 특별한 합의가 이루어지지 못한 채 다양하게 사용하고 있다(Busch 1995).

문화콘텐츠는 양적으로 방대하고, 유일본이라는 작품의 특성상 다른 작품과 공유될 수 없는 성격을 가지고 있다. 그리고 개별적이고 단순한 정보가 반복되므로 데이터 요소는 양적으로 방만해지고, 이로 인해 색인과 검색이 어려워지는 결과

를 초래한다. 따라서 문화콘텐츠 정보서비스를 효율적으로 제공하기 위해서는 이들의 특성을 잘 반영한 표준화된 메타데이터 포맷의 개발이 필요하다.

2.3 문화콘텐츠의 메타데이터

문화콘텐츠 서비스를 제공하고 있는 가상박물관과 디지털도서관 및 디지털 아카이브와 같은 통합정보시스템은 소장 자원에 대한 기술에서 디지털콘텐츠에 이르기까지 텍스트에서 멀티미디어에 이르는 다양한 형식의 자원을 포괄하고 있다.

박물관과 도서관 및 아카이브에 수집된 문화콘텐츠의 가치는 세계로 연결되는 인류의 생명 기록이다. 이것은 세대를 통해 축적된 지식으로 인류의 진실을 나타낸다. 아울러 기록 기관으로서 인류의 경험과 표현을 미래 세대를 위해 기록, 보존하여 교육과 학습을 지원하도록 지식을 공유하고 있다.(Gill 2002).

문화콘텐츠의 메타데이터 작성자인 박물관 기록자(registrar), 도서관 목록작성자, 아카이브 기록관리사(archivist)들은 콘텐츠 객체의 식별이나 배열 및 접근성을 강화하기 위해 작성하는 목록이나 색인과 같은 맥락에서 메타데이터를 이해하고 있는 것으로 조사되었다 (Gilliland-Swetland 1998).

문화콘텐츠 관련 통합정보시스템에서 다양한 유형의 메타데이터의 통합은 자원의 관리와 통합을 유지하는데 매우 중요한 문제로 제기되고 있다.

Gilliland-Swetland(1998)는 <표 1>과 같이 메타데이터의 유형은 디지털콘텐츠 객체의 관리에서 자원의 설명, 보존, 기술, 이용 등으로 기능적 측면을 고려하여 범주화하고 있다.

오늘날 박물관, 도서관, 아카이브의 메타데이터 작성방법은 매우 다양하다. 이에 다양한 유형의 메타데이터를 조정하고 통합하는 다양한 노력을 기울이고 있다. 현재 문화콘텐츠 메타데이터 표준 포맷으로 널리 사용되는 것은 Dublin Core, EAD, VRA, CDWA, CIMI, Object ID 등이 있다.

3 문화콘텐츠 메타데이터 포맷

3.1 더블린 코어

더블린 코어(Dublin Core)는 네트워크 기반 인터넷 자원의 필수적인 특징을 기술하고 메타데이터간의 호환성을 제공하도록 설계된 포맷이다.

1995년 NCSA와 OCLC주최 워크샵으로 출발한 더블린 코어 메타데이터 포맷은 현재 50개국에서 사용되고 있으며, 20개이상의 언어로 정의되어 제공되고 있다.

더블린 코어는 원래 디스크립터들의 소규모 집합이었지만 전 세계의 예술, 과학, 교육, 경영, 정부부문에서의 다양한 정보제공자들로부터 즉각적인 관심을 불러일으켰다. 이에 따라 수 차례의 더블린 코

〈표 1〉 기능적 측면의 메타데이터 유형

유형	정의	사례
관 리 (Administrative)	자원의 관리 및 운영 메타데이터	<ul style="list-style-type: none"> · 수집정보 · 권리와 재생산 · 법적 접근 요구사항 문서화 · 위치정보 · 디지털화 선정 기준 · 유사 객체간 버전 통정과 차이 · 시스템에 의해 생성된 감사결과
설 명 (Descriptive)	자원의 설명과 식별 메타데이터	<ul style="list-style-type: none"> · 목록 레코드 · 탐색 보조도구 · 특수화된 색인 · 자원간의 연결 관계 · 이용자에 의한 주석 · 레코드 제작자에 의한 메타데이터
보 존 (Preservation)	자원의 보존 관리 메타데이터	<ul style="list-style-type: none"> · 자원의 물리적 상태 문서화 · 자원의 물리적이거나 디지털버전의 보존 조치에 대한 도큐멘테이션
기 술 (Technical)	시스템운용 및 작동 메타데이터	<ul style="list-style-type: none"> · 하드웨어와 소프트웨어 문서화 · 디지털화 정보(예:포맷, 압축비율) · 시스템 반응시간 추적 · 승인과 보안데이터(예:비밀번호)
이 용 (Use)	자원 이용의 수준과 유형 메타데이터	<ul style="list-style-type: none"> · 전시 레코드 · 이용과 이용자 추적 · 콘텐츠 재사용, 멀티브러전정보

어 메타데이터 워크숍이 개최되었고, 이 과정에서 더블린 코어 내용의 지속적인 보완이 이루어졌으며 동시에 미국 국립정보표준화기구(NISO: National Information Standards Organization)에 의해 표준화가 추진되었다. 2001년 9월 미국표준협회(ANSI: American National Standards Institute)에 의해 ANSI/NISO Z39.85-2001 표준으로 확정되었다. 또한 ISO 기술프로그램 TC 46(Information and documentation)

에서 ISO/DIS 15386 표준화 활동의 제4 단계인 질의단계(40.60)에서 모든 국가회 원기관에 투표를 위한 질의안이 배포된 상태이다(조윤희, 황동열 2003). 더블린 코어 메타데이터 포맷은 15개의 요소 집 합으로 이루어져 있으며, 크게 콘텐츠 기술요소, 지적 속성요소, 물리적 기술요소 등으로 구분된다.

더블린 코어의 목표는 첫째, 데이터의 형식과 구조를 단순화하여 원문의 저자나

발행자가 메타데이터를 직접 작성하고 둘째, 네트워크 출판을 위한 저작도구의 개발자가 이 정보에 대한 템플릿을 직접 해당 소프트웨어에 포함할 수 있도록 하며 셋째, 작성된 데이터를 기초로 특정 분야에서 요구되는 상세한 수준으로 확장하여 사용할 수 있도록 하는 데 있다. 또한 더블린 코어는 다른 메타데이터 표준을 대체하는 것을 목표로 하지 않는 대신에 동일한 정보자원을 기술하는데 있어서 다른 의미 정보를 나타내는 메타데이터 표준과 공존하는 것을 목표로 한다.

이상에서와 같이 더블린 코어의 가장 큰 장점은 단순성(simplicity)과 범용성(common)이다. 그러나 단순성은 효율적인 정보검색을 어렵게 만들어 상호운용성을 저하시킬 수 있는 문제점을 안고 있다(Weibel 1997). 이러한 문제점을 해결하기 위해 각 요소에 하위요소인 한정어를 마련하여 세분화, 상세화를 통해 요소의 단순성 문제를 해결하고 검색효율성을 향상시키고자 하였다.

3.2 EAD

EAD(Encoded Archival Description)는 1993년 미교육부에서 탐색보조도구 전산화의 표준을 마련할 것을 제안하여 버클리대학에서 수행된 Berkeley Finding Aids Project에서 출발하였다

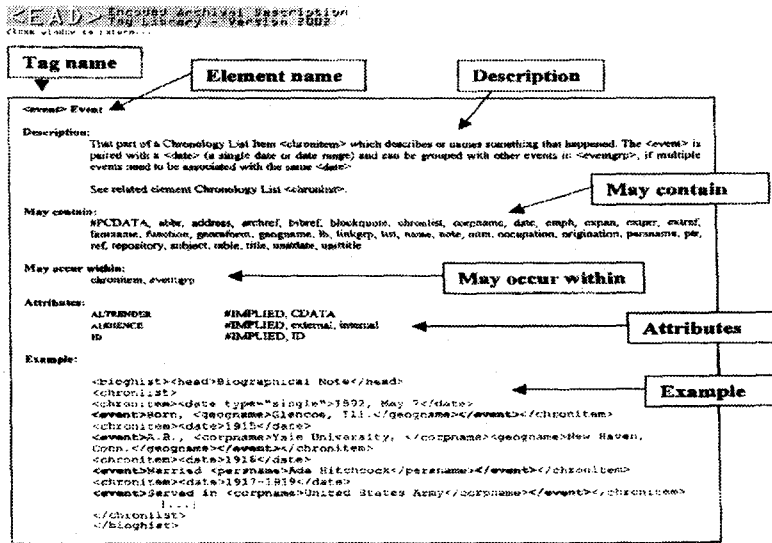
SAA(Society of American Archivists), CAIE(Committee on Archival Information

Exchang), RLG(Research Libraries Group), OCLC 등이 추가로 참여하였고, 국회도서관의 네트워크개발마크표준위원회에서 EAD의 관리를 책임지고 있다.

박물관, 도서관, 아카이브에 소장된 정보에 대한 접근수단으로 네트워크를 기반으로 하는 정보서비스의 중요성이 점차 커지고 있다. 아울러 MARC가 제공하는 정보의 범위와 내용을 확장시키려는 노력이 증가됨에 따라 EAD 프로젝트는 기계가독형 탐색보조도구의 개발가능성을 시험하기 위해 진행되었다. 여기에서 탐색보조도구란 박물관, 도서관, 기록보존소에서 작성한 인벤토리, 색인, 지침과 같은 형식의 기술적 서지 혹은 메타데이터를 의미한다(Cover 1999).

EAD는 아카이브의 필사본, 공문서, 사료, 보존문서를 기술하는 데이터를 부호화하기 위해 여러 아카이브와 필사본 도서관에서 표준으로 사용하고 있다. <그림 2>는 EAD 메타데이터 요소를 기술하고 있는 사례이다.

EAD의 목적은 많은 아카이브 기관에서 소장하고 있는 자원을 이용자가 용이하게 접근하도록 하는데 있다. 아울러 이들의 주요 설계 원칙은 특정 아카이브나 매체기록의 범위를 넘어 유기적 연결 필요성을 명확하게 반영한다. 요소와 속성명은 언어와 응용측면에서 모두 국제 교환에 적합하도록 설계하였으며, 동시에 특정언어나 매체 출력 요구를 반영한 메카니즘을 제공하고 있다.



〈그림 2〉 EAD 메타데이터 구조

EAD는 공적으로 공유된 기록보존 자원에 대한 정보를 다루며, 소유권, 보존, 전시, 이용, 축적이나 자료의 기술적 처리와 같은 수집관리활동을 위한 시스템이 아니다. EAD는 데이터 구조로서 표준을 의미하지는 않는다. 이에 어떤 데이터 요소가 출현하도록 데이터를 어떻게 형성하는가를 기술하는 것이 아니라 외부국가나 국제적 데이터 내용 표준으로서의 역할을 한다.

EAD는 SGML/XML 구문론에 기반한 데이터 커뮤니케이션 포맷으로서 원칙적으로 전송 메카니즘으로 사용된다. 이에 SGML/XML기반 시스템에서 직접 기술 데이터가 관리된다.

EAD는 변화하는 환경을 수용하도록 설계되었다. 이에 내용 표현이 아닌 기록 기술의 구조적 내용에 초점이 맞추어져

있다. 아울러 다양한 형식으로 출력을 지원하도록 충분한 메카니즘을 제공한다. 새로운 출력형식인 웹 표시나 출력을 비롯하여 다양한 형식의 목록과 같은 탐색 보조도구의 전통적 형식을 지원한다

3.3 VRA의 Core Category

시각자료협회의 데이터표준위원회(Visual Resources Association Data Standards Committee)는 네트워크 환경에서 시각 컬렉션 관리에 대한 요구가 증가함에 따라 정보의 관리, 조직, 교환 표준을 위해 1993년에 조직되었다.

위원회에서는 CDWA가 박물관 객체 기술에 대한 광범위한 요소를 수록하고 있으나 이미지 기술에는 미흡하고, 특히 건축물 등의 기술에 필요한 요소가 포함

되지 않음을 발견했다. 이에 건축물을 포함하여 예술 작품, 공예품 등을 묘사한 시각적 문헌을 기술하기 위한 지침을 마련하게 되었다.

1995년 메타데이터 요소의 범주 검토를 시작하여 최소 수준의 기술과 완전 수준의 기술 중간 단계로서 만족할 만한 수준인 "Core"개념을 도입했다. 이에 21개의 요소를 포함한 "The Core Category for Visual Resources" 버전 1.0을 개발하였다. 객체 범주(Object Categories) 13개 요소와 제작자 범주(Creator Categories) 3개 요소 및 대용물 범주(Surrogate Categories) 5개 요소로 정의하였다.

버전 2.0에서는 작품설명범주 (Work Description Categories) 19개 요소와 시각 문헌범주 (Visual Document Description Categories) 9개 요소인 두 범주로 분류하였으나 2002년 버전 3.0에서는 VRA 핵심 범주를 17개 요소로 정의하였다(Visual Resources Association 2002).

핵심 범주의 각 요소들을 예제와 함께 간단한 정의를 제공하며, 다른 문화콘텐츠 메타데이터간의 상호운용성을 용이하게 하기 위해 매핑작업과 어휘통제를 위한 표준을 사용할 것을 권고하고 있다. VRA Core 3.0은 여전히 개발중에 있으며, CDWA와 MARC 및 다른 메타데이터 표준과의 매핑과 통제어휘 이용을 권고하는 안내 및 예제 일람표의 개발을 지속적으로 수행하고 있다.

3.4 CDWA

CDWA(Categories for the Description of Works of Art)는 예술정보의 제공과 이용을 위하여 예술역사가, 박물관 큐레이터 및 등록자, 시각자원전문가, 예술사서, 정보관리자 및 기술전문가 등의 커뮤니티 대표자들로 구성된 AITF(Art Information Task Force)에 의해 개발되었다.

CDWA의 개발 목적은 첫째, 학회와 연구자들간에 공유되는 예술 작품 대한 설명이 포함되도록 학술적 정보(scholarly information)를 정의하고 둘째, 예술 정보에 대한 접근을 더욱 용이하게 하는 적절한 공통 용어(common terminology)와 설명 요소를 규명하는 것이다. 셋째, 대학과 연구자들간의 정보 공유를 위한 표준화된 교환 포맷을 권고하는 것이다.

범주는 예술 정보를 제공하고 이용하도록 하며, 예술작품과 이미지를 설명하는데 사용되는 정보에 대한 구조를 제공하기 위해 개발되었다.

범주 작성의 목적은 예술 데이터베이스의 내용을 명확히 하기 위한 것이다. 예술작품의 설명 정보에 대한 유일하고 포괄적인 프레임워크를 제공함으로써, 예술정보를 공유하고자 하는 다양한 시스템간에 호환성을 강화시키도록 의도되었다. 범주에 의해 정의된 프레임워크는 공유와 교환에 적합한 정보를 나타낸다.

범주는 전 시대와 전 지역에 걸쳐 그림, 논문자료, 조각, 도예, 금속세공품, 가구, 디자인, 공연예술 등을 포함하여 동화

상과 이미지에 주안점을 둔다.

예술작품은 분명하게 정의할 수 있는 단위가 아니므로, 단체나 개별 분리 설명, 혹은 사례를 기반으로 결정해야 한다. 특히, 예술작품에 대한 정보는 복잡하다. 예술사적 자원에는 재고목록, 공중기록, 판매 청구서, 영수증, 편지, 학술논문, 다이어리, 계정 도서, 계약서, 생사기록 등 다양한 자료를 포괄하고 있다. 아울러 다른 예술 자료와 같이 지도, 인쇄물, 사진, 필름, 비디오, X-ray 등의 시각적 자료를 포함하기도 한다. 예술사적 연구자료의 다양성과 신뢰성의 변이는 예술 정보를 자동화하는 어려움으로써, 이들은 자동화된 환경에서 정보를 보전하는 것 또한 매우 중요하다.

2.0버전에서 범주는 작품의 본질적 정보와 비본질적 정보로 분류하였으나 3.0 버전에서는 객체, 구조 혹은 그룹(Object, Architecture, or Group)의 핵심 항목으로

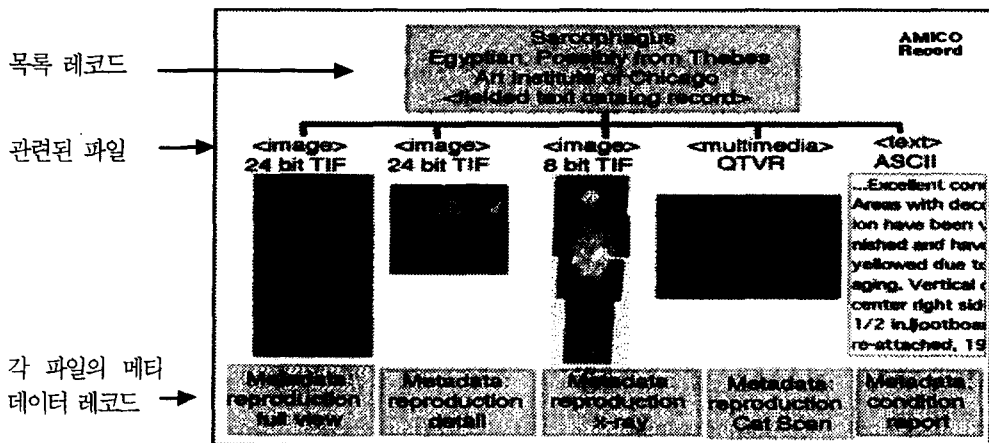
서 27개 요소와 전거/어휘통제(Authorities /Vocabulary Control)항목으로 4개 요소를 설정하고 있다.

3.5 AMICO의 Data Directory

AMICO(Art Museum Image Consortium)는 1997년에 구성되었다. 문화콘텐츠 자원을 협력적으로 생성하면서 박물관 멀티미디어의 교육적 이용이 가능하도록 협력하는 30개 이상의 박물관들로 구성된 비영리 조직이다.

AMICO의 데이터 명세서는 텍스트 레코드, 관련된 이미지 및 멀티미디어 파일과 데이터 사전(Data Directory)으로 구성된다. 데이터 사전은 CDWA를 기반으로 작성되는 목록레코드 필드와 더블린 코어를 기반으로 작성된 메타데이터 레코드 필드로 정의된다.

데이터 사전 1.3 버전에 의한 예술 작



<그림 3> AMICO의 정보조직 구성도

품의 정보조직은 <그림 3>과 같이 최소한의 모든 핵심 필드를 가지고 있는 목록 레코드와 최소 하나 이상의 이미지 파일과 관련된 멀티미디어 파일과 연계된 멀티미디어 파일 및 각각의 연계된 멀티미디어 파일에 대해 모든 핵심 필드를 기록한 메타데이터 레코드로 구성되어 있다.

3.6 CIMI의 Access Point

박물관 커뮤니티 정보전문가들에 의해 박물관 정보의 장기적 이용과 광범위한 접근을 위해서는 표준화가 필요하다는 인식이 1980년대를 지나면서 널리 확산되었다.

이에 CIMI(Consortium for Interchange of Museum Information)는 1990년 디지털 형태의 박물관 정보 보존을 지원하고 정보교환의 잠재력을 확장하는 표준을 개발하기 위한 연구 프로젝트를 수행하는 국제적인 박물관 조직으로 탄생하였다.

이론과 실용적 단계의 개발을 목적으로 CIDOC(The documentation Committee of the International Council of Museums)은 박물관 정보를 캡슐화하는 데이터 모델을 개발하였다. 이론적 단계에서는 박물관간 데이터 교환을 위한 표준들을 고찰하여 일반적인 정보 모델을 개발하고 실용적인 단계에서는 개발된 표준과 모델을 대상으로 프로토타입 개발을 목적으로 하였다(Light 1995).

CIMI의 주요 프로젝트로는 CHIO(Cultural Heritage Information Online)가 있다. 이 프로젝트는 탐색가능한 10,000개

이상의 객체에 대한 레코드와 민속예술 정보를 제공하고 있다. 대상으로는 이미지와 전통적인 박물관 데이터베이스 레코드와 전시목록의 전문을 포함하고 있다.

3.7 Object ID

Object ID는 문화 콘텐츠(객체)를 기술하는 국제 표준이다. 이것은 박물관 커뮤니티, 경찰 및 관세대행사, 예술품 무역 및 보험사, 예술 및 유물 감정사 등의 협력을 통하여 개발되었다.

Object ID 프로젝트는 J. Paul Getty Trust에 의해 1993년에 시작되었고 1997년 표준으로 정착되었다.

Getty Information Institute는 문화콘텐츠 객체의 식별을 위해 국제간 표준을 개발하기 위한 협력적 프로젝트를 시작했다. 이것은 FBI와 Scotland Yard와 인터폴 및 박물관, 문화유산, 예술무역 및 감정 관련 조직, 보험사 등을 포함하여 주요 법률대행사에 의해 장려되고 있다.

Object ID 프로젝트는 문화콘텐츠의 설명적 요소의 표준 제정을 통하여 전세계적으로 적용을 고무하는 다양한 커뮤니티와 예술품 도난을 제거하기 위해 노력하고 있다. 1999년 Object ID 프로젝트는 CoPAT(Council for the Prevention of Art Theft)으로 이전되었다. CoPAT은 예술, 유물 및 건축 분야에 범죄를 방지하는 임무를 지닌 단체로서 1992년에 설립되었다. 이들의 임무는 예술, 유물과 건축 분야에서 도난을 예방하는 것이다.

Object ID 범주에 요소가 포함되려면 문화유산, 법 집행, 예술품 무역, 평가, 보험, 세관 등과 관련된 6개의 커뮤니티를 통한 국제설문조사의 결과에서 모든 분야에서 75%이상의 승인을 받았을 경우에 가능하다. 현재 사진(photographs)을 비롯하여 11개의 요소가 핵심 범주로 구성되어 있다. 아울러 부가요소로 권고되고 있는 요소는 객체 ID번호(Object ID Number)를 비롯하여 5개 요소가 있고, 아직 합의를 이루지 못하고 사용되고 있는 요소로는 객체 상태(Condition of Object)를 비롯하여 7개의 요소가 있다.

4 문화콘텐츠 메타 데이터 포맷의 비교 분석

4.1 포맷의 선정 원칙

문화콘텐츠 관련 메타데이터에 대한 단일화된 하나의 국제 표준은 없다. 최근 메타데이터는 디지털콘텐츠 객체에 대하여 Dublin Core와 같이 비교적 단순한 포맷에서 CDWA와 같이 상당히 복잡하고 풍부한 포맷에 이르기까지 복잡성과 풍부성 수준이 다른 다양한 메타데이터 포맷이 표준으로 사용되고 있다.

본 연구에서는 문화콘텐츠를 대상으로 하는 메타데이터 포맷 중에서 그 범위와 영향을 기반으로 6개의 포맷을 선정하였다. 포맷 선정의 기준으로는 첫째, 메타데이터 포맷의 범위이다. 특정 영역의 포맷

보다는 문화콘텐츠 객체에 대한 일반적 메타데이터 포맷에 중점을 두었다. 둘째, 특정 커뮤니티에 의해 개발된 포맷보다는 광범위하게 이행되고 실험된 포맷을 선정했다. 이러한 기준에 의거하여 Dublin Core, VRA의 Core Category, EAD, CDWA, CIMI Access Point, Object ID 등 6개의 메타데이터 포맷을 문화콘텐츠 메타데이터 비교 포맷으로 선정하였다.

본 연구의 초점은 일관성, 구문론이나 레코드 구조가 아니라 이들 메타데이터 포맷의 요소의 선정에 관한 것이다. 각 포맷의 요소 비교는 최신 정보와 문헌적 조사를 기반으로 수행하였다.

4.2 유형별 요소 비교

각 포맷의 요소에 대한 비교분석 방법으로는 Gilliland-Swetland(1998)의 메타데이터 유형을 기반으로 관리, 설명, 보존, 기술, 이용으로 범주화하여 비교표를 작성하였다. 이 비교표는 Getty에서 제공하는 메타데이터 표준 참조테이블(Metadata Standards Crosswalk 2001)과 해당 홈페이지에서 제공하는 최신의 정보를 기반으로 작성하였다. 기본적으로 하나의 요소는 하나의 유형으로만 분류하였다.

때문에 포맷과 같이 기술과 보존의 유형에 해당하는 요소인 경우에도 좀 더 비중이 있다고 판단되는 하나의 요소인 보존으로 분류하였다.

〈표 2〉에서와 같이 문화콘텐츠 메타데이터 포맷들은 주로 설명요소를 중심으로

〈표 2〉 메타데이터 포맷의 유형별 요소

메타데이터 포맷	요소 수	유형별 요소의 수와 비율				
		관리	설명	보존	기술	이용
Dublin Core	15	2(14%)	12(80%)	1(6%)		
EAD	60	5(8%)	44(73%)	10(17%)	1(2%)	
VRA 3.0	17	2(12%)	12(70%)	3(18%)		
CDWA	31	4(13%)	20(65%)	5(16%)		2(6%)
CIMI	60	8(13%)	40(67%)	6(10%)	5(8%)	1(2%)
Object ID	11	1(9%)	8(73%)	2(18%)		

구성되어 있음을 알 수 있다. 특히 하드웨어와 소프트웨어 문서화를 비롯한 시스템 반응시간 추적 및 승인과 보안데이터 등의 시스템운용과 작동에 관련된 메타데이터 요소는 거의 포함하고 있지 않음을 알 수 있다.

Burnett, Na, and Park(1999)은 메타데이터 접근방법을 문헌정보학적 서지통정 접근방법과 컴퓨터공학적 데이터관리접근방법으로 나누어 비교하였다. 이 연구결과에서 문헌정보학적 접근방법의 대표적인 포맷인 MARC나 더블린 코어는 자원의 본질적 요소에 치중하면서 관리요소는 거의 포함하고 있지 않음을 지적하였다. 반면에 컴퓨터공학적 접근방법인 데이터 관리접근방법의 대표적인 포맷인 TEI Header, IAFA Templates, Semantic Header는 본질적 요소이외에도 부가적 요소로서 시스템 운영, 관리를 위한 관리요소를 상당부분 선정하고 있음을 지적하였다.

그런데 문화콘텐츠 메타데이터 포맷들은 대부분 도서관, 박물관, 미술관을 중심

으로 문헌정보학적 서지통정 접근방법으로 접근하고 있어 대부분의 요소가 자원객체의 설명과 보존 요소에 치중되어 있는 것으로 판단된다.

분산 네트워크 환경에서 문화콘텐츠가 보다 효율적으로 관리되고 서비스되기 위해서 관리요소는 반드시 필요하다. 따라서 향후 관리 및 기술 유형의 요소가 추가되어야 할 것이며, 이용과 관련한 요소도 추가되어야 할 요소로 판단된다.

4.3 포맷간 요소의 비교

문화콘텐츠 메타데이터 포맷의 요소는 Object ID나 더블린 코어와 같이 단순한 형식에서 CIMI와 같이 풍부한 포맷에 이르기까지 선정하고 있는 요소의 수가 다양하다. 또한 각 포맷이 핵심요소로 선정하고 있는 요소의 명칭도 다양하다.

본 연구의 비교 포맷으로는 문화콘텐츠 분야에서 일반적으로 표준으로 널리 사용되고 있는 더블린 코어, VRA 3.0, CDWA, CIMI, Object ID를 대상으로 〈표 3〉

〈표 3〉 포맷간 일반성이 높은 요소 비교

유형 포맷	Dublin Core	VRA 3.0	CDWA	CIMI	Object ID
설 명	Title	Title	Title or Names	objectTitle bibliographicTitle objectName	Title
	Creator	Creator	Creation Creator Identification	creatorGeneral creatorRole creatorName creatorDateOfBirth creatorDateOfDeath creatorNationalityCultureRace creatorInfo	Maker
	Subject	Subject	Subject Matter Subject Identification Classification	Subject contentGeneral content	Subject
	Description	Description	Orientation/Arrangement Inscriptions/Marks Conservation/Treat. Hstry Descriptive Note Physical Description	physicalDescription inscription/Mark	Description Inscriptions & Marks Distinguishing Features
	Date	Date		dateOfBirth dateOfDeath dateOfOrigin dateCollected	Date or Period
	Type	Type Record Type	Object/Work-Type	typeSpecimen	Type of Object
	Relation	Relation	Related Works Related Visual Doc. Related Textual Ref.	relatedObjects relatedObjects-rendition relatedTextualReferences	
		Style/Period	Styles/Periods/Groups /Movements	stylePeriod periodName	
	Identifier	ID Number		Object ID	
관 리	Rights	Rights	Copyright/Restrictions	copyrightRestriction provenance	
		Location	Current Location Place/Lication Iden.	repositoryName repositoryPlace	
보 존	Format	Measurements	Measurements	quantity dimensions	Measurements
		Technique Materials	Materials & Techniques	processTechnique materialMedium	Materials & Techniques

은 세 개 이상의 포맷에서 공통적으로 선정하고 있는 요소를 중심으로 비교표를 작성하였다. 그밖에 두 개 이상 포맷에서 선정하고 있는 요소와 특정 포맷에서만

단독으로 선정하고 있는 요소 비교는 〈표 4〉에서 정리하였다. 단, EAD 포맷은 주로 아카이브의 기록보존을 위한 메타데이터 요소들로 이루어져 비교 대상에서 제

의하였다.

모든 포맷에서 공통적으로 핵심 요소로 선정하고 있는 요소는 <표 3>에서 보는 바와 같이 표제(Title), 제작자(Creator), 주제(Subject), 설명(Description), 유형(Type), 형식(Format) 등 6개 요소에 불과하다. 또한 <표 4>와 같이 각 포맷에서 단독으로 선정하고 있는 요소는 더블린 코어의 '출판사(Publisher)'나 '범위(Coverage)'를 비롯하여 상당수가 된다. 아울러 각 포맷에서 선정하고 있는 요소간 명칭 표현이 상당히 상이함을 알 수 있다.

인터넷자원의 메타데이터 포맷으로 개발된 더블린 코어와 국제간 예술, 유물 및 건축 분야에 도난을 예방하기 위해 개발된 Object ID 메타데이터 포맷은 핵심적인 간단한 요소만을 기본요소로 선정하

고 있다. 따라서 전자는 주로 인터넷 디지털콘텐츠의 식별과 검색에 필요한 요소들이 기본요소로 선정되었으며, 후자는 문화예술품의 국제적 도난을 예방하기 위한 목적으로 표준이 제정되었으므로, 문화예술품의 식별요소들이 기본요소로 선정되어 있다. 두 포맷의 요소명칭에 있어서도 더블린 코어의 Creator와 Object ID의 Maker와 같이 각 분야에서 일반적으로 통용되는 용어에 기반하고 있다.

VRA 3.0 포맷은 시각자료들의 식별요소를 중심으로 핵심요소만을 선정하고 있으나 Title, Creator, Date, Measurements, Material, ID Number, Relation, Location 등의 요소는 한정어(Qualifiers)로 하위요소들을 제시함으로써 핵심요소들을 세분화하고 있다.

<표 4> 포맷간 특정성이 높은 요소 비교 * 는 단독 선정요소임

유형 포맷	Dublin Core	VRA 3.0	CDWA	CIMI	Object ID
설명	Source	Source		resource	
	Language			objectLanguage	
관리			Context	contextHistorical contextArchaeological	
	Contributor			association associationGeneral association-name	
보존 이용			Ownership/Cldt. Hstry	owner	
			Condition/Exam Hstry	condition	
기타			Exhibition/Loan Hstry	Exhibition History	
	*Publisher *Coverage	*Culture	*Object/Work *State *Edition *Facture *Cataloging History *Generic Concept Iden. *Critical Responses	*nationalityCultureRace *creditLine *wallTextLabel *protectionStatus *fieldCollector *administrativeEvent *administrator의 다수	*Photographs

CDWA 포맷은 예술정보를 위한 메타데이터 요소와 전거 및 어휘통제를 위한 요소를 포함하고 있다. 목록정보에서 이미지와 멀티미디어 메타데이터 정보에 이르는 다양한 요소들을 기본요소로 선정하고 있다. 각 기본요소들은 하위요소를 템플릿 형식으로 제공하고 있어, 다양하고 풍부한 메타데이터를 생성할 수 있도록 제시하고 있다.

CIMI 포맷은 박물관 소장품들의 식별과 검색에 필요한 요소들을 풍부하게 선정하고 있는 포맷이다. 이미지나 멀티미디어 메타데이터와 관련한 요소가 빈약하나 전시정보를 포함하여 풍부한 관리요소를 선정하고 있다.

이상에서 살펴본 바와 같이 명확하게 메타데이터 범주에 포함되는 요소의 결정은 메타데이터 포맷의 근원적 관계를 잘 반영하고 있다. 아울러 메타데이터 작성 수준이 다양하게 차이가 있음에도 불구하고 자원의 확인과 식별을 위한 요소는 모든 메타데이터 포맷의 핵심적인 요소로서 이들은 모두 일반적 관계라는 것을 명확하게 보여 주고 있다.

4.4 최소 수준의 데이터 요소

본 연구는 세계적으로 널리 사용되고 있는 문화콘텐츠 메타데이터 포맷으로 Dublin Core, VRA, CDWA, CIMI, Object ID의 데이터 요소를 중심으로 비교, 분석하였다. 이를 통하여 문화콘텐츠 통합을 위한 최소 수준의 데이터 요소를

선정한 일반적인 기준은 다음과 같다.

첫째, 텍스트, 이미지, 사운드, 그래픽, 애니메이션, 동영상 등 형태 구별없이 모든 자료를 포괄할 수 있는 데이터 요소를 선정하였다. 이는 개별 시스템간 상이한 메타데이터 요소 중 최소한의 공통된 필수 데이터 요소를 제시하여 시스템간 상호운용성을 확보하기 위한 것이다.

둘째, 문화콘텐츠 메타데이터 포맷에서 일반성이 높은 데이터 요소를 최소 수준의 데이터 요소로 선정하였다. 이는 기존 데이터 요소를 최대한 수용하여 시스템간 연동이나 정보공유가 가능하도록 하기 위함이다.

셋째, 문화콘텐츠 통합을 위한 최소 수준의 데이터 요소는 국내외 문화콘텐츠의 현황 및 표준화 동향을 지속적으로 반영할 수 있어야 한다. 이를 위하여 기존 데이터 요소를 최대한 수용함과 동시에 변화하는 환경에 적절한 요소를 수용할 수 있는 여지가 있어야 한다.

이상의 일반적인 기준을 근거로 문화콘텐츠 메타데이터 표준으로 널리 이용되고 있는 포맷간 요소를 비교, 분석하여 문화콘텐츠 통합을 위한 최소 수준의 메타데이터 요소를 선정하였다.

최소 수준의 메타데이터 요소는 <표 5>에서와 같이 Gilliland-Swetland(1998)가 기능적 측면에서 분류한 관리, 설명, 보존, 기술, 이용 유형으로 범주화하였다.

관리범주는 자원의 관리와 운영을 위한 데이터 요소로서 권리정보, 위치정보, 마

크 및 관리정보 등을 데이터 요소로 선정하였다. 이 중 관리정보는 CIMI 포맷에서만 선정하고 있는 요소이나 시스템의 효율적인 운용과 관리에 필요한 요소로 판단되어 추가하였다.

설명범주는 자원의 확인과 식별을 위한 데이터 요소로서 식별기호, 표제, 제작자, 유형, 주제, 연도, 설명, 스타일, 관련정보 등을 데이터 요소로 선정하였다. 식별기호를 제외한 모든 요소는 4개 이상의 포

맷에서 선정된 요소들로 구성하였다. 식별기호는 Dublin Core, VRA, CIMI 포맷에서 선정하고 있으나, 자원의 유일한 식별을 위해 반드시 필요한 필수요소라고 판단되어 추가하였다.

보존범주는 자원의 보존 및 관리를 위한 데이터 요소로서 형식, 재료, 상태 등을 데이터 요소로 선정하였다. 형식과 재료는 4개 이상의 포맷에서 선정하고 있는 요소이나 상태요소는 문화콘텐츠의 특성

<표 5> 최소 수준의 메타데이터 요소

범주	데이터 요소	하위 요소(값)	수	필수/선택
관리	권리정보(rights)	copyright, restriction	n	선택
	마크(Mark)	inscriptions, mark	1	선택
	위치정보(location)	currentLocation, repositoryPlace	n	선택
	관리정보(administration)	ownership, contributor, administrativeInfo	n	선택
설명	식별기호(identifier)	ISAN, V-ISAN, ISBN, ISMN, ISSN, ISRC, ISTC, ISWC, PII, SICI, URNs, PURLs, BICI, DOI	n	필수
	표제(title)	objectTitle, bibliographicTitle	n	필수
	제작자(creator)	creatorGeneral, creatorRole, creatorName, creatorDateOfBirth, creatorDateOfDeath, creatorNationality	n	필수
	유형(type)	objectType, workType, RecordType	n	필수
	주제(subject)	DDC, KDC, LC, LCSH 등	n	필수
	연도(date)	dateOfBirth, dateOfDeath, dateOfOrigin	n	선택
	설명(description)	descriptionNote, physicalDescription	n	필수
	스타일(style/period)	styles, periods, groups, movements	n	선택
보존	관련정보(relation)	relatedIdentity, relatedType, relatedObjects, relatedWorks	n	선택
	형식(format)	documentation, measurements, version	n	필수
	재료(materials)	processTechniques, materialMedium	n	선택
	상태(condition)	protectionStatus, protectionDate	n	선택
기술	시스템 기능(system functions)	hardware, software	n	선택
	승인정보(Authentication)	id, password, security data	n	선택
이용	전시정보(exhibition)	exhibition history, loan history	n	선택

상 보존을 위해 필요한 요소로 판단되어 선택요소로 추가하였다.

기술범주는 문화콘텐츠 시스템의 운용 및 작동을 위한 데이터 요소로 시스템 기능, 승인정보 등을 데이터 요소로 선정하였다. 이들 요소는 문화콘텐츠 포맷에서 선정되어 있지는 않지만, 필요한 요소로 판단되어 임의로 추가한 요소이다.

마지막으로 이용범주에는 문화콘텐츠의 이용 수준과 유형에 관한 데이터 요소로 전시정보, 이용정보 등을 데이터 요소로 선정하였다. 전시정보는 CDWA와 CIMI 포맷에서 선정하고 있는 요소이다.

5 결론 및 제언

문화콘텐츠는 일반 콘텐츠와 달리 박물관, 미술관, 도서관 등 관련기관에 따라 다양한 메타데이터 포맷을 표준으로 적용하고 있다. 또한 포맷들은 특정 목적을 위해 작성됨으로써 각 포맷간 메타데이터 요소가 상당히 상이하다. 따라서 예술품 도난 예방을 위한 메타데이터 포맷으로 개발된 Object ID와 같이 단순한 포맷에서 박물관 소장품 정리를 위해 개발된 CIMI와 같이 풍부한 포맷에 이르기까지 다양한 요소의 포맷들이 있다.

문화콘텐츠 분야에서 널리 사용되고 있는 메타데이터 포맷인 Dublin Core, VRA, CDWA, CIMI, Object ID 포맷을 비교 분석한 결과를 요약하여 정리하면

다음과 같다.

첫째, 문화콘텐츠 메타데이터 포맷은 대부분 문헌정보학적 서지통정 접근방법으로 접근하고 있어 대부분의 요소가 자원 객체의 설명과 보존을 위한 요소에 치중되어 있고, 기술 및 이용 요소는 상대적으로 상당히 빈약하였다.

둘째, 문화콘텐츠의 메타데이터 포맷에서 공통적으로 선정하고 있는 요소는 Title, Creator, Subject, Description, Type, Format 등 6개의 요소에 불과하였다. 이들은 모두 문화콘텐츠의 식별과 검색을 지원하는 요소라는 점에서 공통점을 지니고 있으나, 요소의 명칭이나 표현이 서로 상이하여 문화콘텐츠 시스템간 연동이나 정보공유에 어려움이 있다.

셋째, 네트워크를 통한 문화콘텐츠의 효율적인 유통과 상호운용성을 확보하기 위한 필수적인 요소로 식별자 요소가 중요하나 더블린 코어와 VRA, CIMI에서만 기본요소로 선정하고 있었다.

넷째, 문화콘텐츠 메타데이터 요소의 비교, 분석 결과를 토대로 문화콘텐츠 통합을 위한 최소 수준의 메타데이터 요소를 선정하였다. 최소 수준의 데이터 요소 중 필수요소는 각 포맷에서 모두 선정하고 있는 요소로 구성하였다. 아울러 각 포맷에서 모두 선정하지는 않았지만, 필요한 요소로 판단되는 것은 선택요소로 채택하였다.

이에 본 연구의 결과를 토대로 제언하면 다음과 같다.

첫째, 문화콘텐츠의 통합을 위한 메타 데이터 포맷은 시스템간 상호운용성을 확보할 수 있는 최소한의 공통된 요소를 포함하여야 한다. 또한 이들 최소 수준의 요소는 국내외 표준화 동향이나 변화하는 환경에 적응하여 적절한 요소를 수용할 수 있도록 설계되어야 한다.

둘째, 현재와 같이 분산 네트워크 환경에서 문화콘텐츠의 효율적인 유통과 서비스를 위해서는 자원의 식별과 검색을 위한 설명 및 보존 요소이외에 시스템의 효율적인 운용 및 관리 요소, 서비스 이용에 대한 수준과 유형에 대한 요소의 추가와 확대가 검토되어야 할 것이다.

참 고 문 헌

문헌정보처리연구회 편. 1998. 메타데이터 형식과 구조. 서울 : 문헌정보처리연구회.

황동열 외. 2003. 『우리문화 원형의 디지털콘텐츠화 종합계획 수립 : 2003 - 2007』. 서울 : 한국문화콘텐츠진흥원.

심상민. 2002. 『콘텐츠 비즈니스의 새 흐름과 대응전략』. 서울 : 삼성경제연구소.

한국. 문화관광부. 2002. 2002년도 문화산업 진흥연차보고서.

Gill, Tony. "The Value of Digital Cultural Content." Position Statement of the Cultural Content Forum.

2002, UKOLN. [cited 2003, 5. 11].<<http://www.ukoln.ac.uk/interop-focus/ccs/positions/creation.html>>.

Visual Resources Association Data Standards Committee. 2002. VRA Core Categories, Version 3.0. [cited 2003, 5. 11]. <<http://php.indiana.edu/%7Efryp/vracore3.htm>>.

Visual Resources Association, 2002. Information on the VRA Core. [cited 2003, 5. 11]. <<http://php.indiana.edu/%7Efryp/coreinfopage.htm>>.

Weibel, Stuart and Innella, Renato. 1997. "The 4th Dublin core Metadata Workshop Report". D-Lib Magazine, June [cited 2003, 5. 11]. <<http://www.dlib.org/dlib/june97/metadata/06weibel.html>>.

Zang, Marcia Lei. 1999. "Metadata element for object description and representation : a case report from a digitized historical fashion project". Journal of the American Society for Information Science, 50(13) : 1193-1208.

Stephenson, Christie. 1999. "Recent development in cultural heritage image database : directions for user-centered design", Library

- Trends, 48(2) : 410-437.
- Basser, Howard and Jennifer Trant. 1995. "Introduction to image : issues in constructing an image database", [cited 2003. 5. 11]. <<http://www.getty.edu/gri/standard/introimages>>.
- Busch, Joseph A. 1995. "SGML for culture heritage information". [cited 2003. 5. 11]. <http://www.cimi.org/old_site/CIMI?SGML_for_CHI.htm>.
- Trant, J. 1995. "Framing the picture : standards for imaging systems". [cited 2003. 5. 11]. <<http://www.archimuse.com/papers/jt.ichim/ichim.4.image.doc.html>>.
- Gilliland-Swetland, Anne J. 1998. "Defining metadata". [cited 2003. 5. 11]. <<http://www.getty.edu/gri/standard/intrometadata>>.
- Castellano, Vanessa. 1997. "Multimedia access to the world cultural heritage". [cited 2003. 5. 11]. <<http://contemporaryartonline.com/articles/mus/mus2.html>>.
- Burnett, Kathleen, Kwong Bor Ng and Soyeon Park. 1999. "A comparison of the two traditions of metadata development". *Journal of the American Society for Information Science*. 50(13) : 1209-1217.
- Joseph A. Busch, "SGML for Cultural Heritage Information", [cited 2003. 5. 11]. <http://www.cimi.org/old_site/CIMI/SGML_for_CHI.html>.
- Encoded Archival Description Tag Library. 2002. Encoded Archival Description Working Group of the Society of American Archivists and the Network Development and MARC Standards Office of the Library of Congress. [cited 2003. 5. 11]. <http://www.loc.gov/ead/tglib/appendix_d.html>.