# 디지타이제이션의 디지털 이미지와 메타데이터 표준화작업

## Imaging and Metadata Standards in Digitization : Practical Strategies

박 은 경(Eun G. Park)*

목 차

1. 서 론
2. 디지타이제이션
3. 디지털캡처와 이미지기술
3. 1 디지털 이미지
3. 2 메타데이터 기술
4. 주요 쟁점
4. 1 메타데이터 품질유지
4. 2 품질관리
5. 결 론

〈초 록〉

디지타이제이션은 문화유산기관의 소장품 보존에 탁월한 공헌을 해 왔으나, 디지타이제이션을 수행하는 상세한 세부기술전략을 표준화하여 지침으로 마련한 기관은 많지 않다. 본고는 디지타이제이션에서 중요한 두 단계인 이미지프로세스와 메타데이타기술(description)에 대한 실무가이드를 사례를 통해 설명한다. 그리고 디지타이제이션 프로젝트 진행에 관련된 중요한 쟁점들을 강조함으로 실무자들이 디지타이제이션을 실행하는데 필요한 전략을 제공한다.

주제어: 디지타이제이션, 이미지프로세스, 메타데이터 기술

<ABSTRACT>

Although digitization contributes to the conservation and preservation of cultural heritage resources, detailing the necessary techniques is a lack of standardization in existing digital repositories. The purpose of this paper is to present practical strategies to illustrate technical details of imaging processes and metadata description with examples. This paper also provides important management issues related to digitization management for project managers and practitioners.

Key words: digitization, imaging process, metadata description

* 맥길대학교 정보학대학원 기록관리담당주임 조교수(eun_park@lycos.com)

# 1. 서 론

디지타이제이션은 최근 전세계적으로 도서관, 박물관, 문화재보호기관 및 관련 문화유산 기관의 주요사업이 되고 있다. 그 기관들의 주요 업무기능인 수집, 데이터베이스, 온라인 목록, 이용자 서비스, 웹사이트, 디지털 참고봉사와 같은 본연의 업무를 컴퓨터시스템에 의해 수행하는 것은 물론, 최근에 많은 기관들이 디지타이제이션 프로젝트를 실시하여 소장목록에 넓은 기회들을 제공하고 있다.

'최근 10년 동안에 국내외적으로 디지타이제이션 프로젝트를 수행한 기관들은 급격히 증가하고 있다'고 보고되어 있다. 성공적인 사례로는 미국의 〈의회도서관〉, 〈국가기록청〉, 캐나다의 〈국립도서관〉, 〈국립기록보존소〉, 호주의 〈국립기록보존소〉 등 다수가 있다.

이들 프로젝트의 선례를 보면, 디지타이제이션은 문화유산물의 보존과 보전에 탁월한 기여를 하였음은 의심할 바가 없다. 지난 20여 년 동안 많은 실험과 시행착오를 통해, 방대한 양의 디지털 컨텐츠가 현재까지 구축되었고, 이 값진 경험들은 디지타이제이션 프로젝트를 수행하는데 중요한 사항들이 무엇인지 잘 보여주고 있다. 예를 들면, 디지타이제이션 업무흐름도(workflow), 전체 진행과정, 프로젝트운영, 디지타이징 툴, 기술적인 세부사항, 시스템 인프라, 이미징, 자료의 조직과 목록, 비용감가분석, 디지털 자료관리 및 보존, 자료관, 이용자 연구, 디지타이제이션 정책 등 여러 가지가 있다. 이런 주요사항들을 보면, 디지타이제이션은 단순히 스캔하여 디지털 이미지로 변환시키는 기술 작업이 아니라 프로젝트를 계획, 설계, 운영, 진행, 관리, 평가, 수정, 재평가하는 총체적인 프로젝트 경영관리임을 알 수 있다.

여러 가지 쟁점들이 있지만, 이 중에서 실무자와 이용자들에게는 무엇보다도 파일포맷, 구조, 내용기술(description) 등 디지털 자료를 다루는 상세한 기술적 내용들에 관심이 더욱 많을 것이다. 왜냐하면, 이런 기술적 세부사항이 이미지의 품질을 보장하고 장기간 보존할 수 있는 밑바탕이 되며, 나아가 프로젝트운영자로써의 위기관리와 의사결정을 위한 중요한 요소가 되기 때문이다.

디지털 자료들을 보관, 생산, 이용하는 자료관은 너무나 넘쳐난다. 그러나 필요한 세부노하우를 표준화하여 지침으로 마련한 자료관은 별로 없다. 선구적인 프로젝트들이 이런 기술전략을 몇 가지 제안하고 있으나 다른 기관에서 도입하여 적용할만한 표준화된 방법과 툴을 제시한 곳은 드물다.

이 논의에서 핵심은 과거의 실무경험에서 근거한 디지타이제이션 운영 관리 프로젝트를 표준화시켜 디지털 자료의 품질을 평가할 수 있는 평가기준표를 만듦으로써, 이를 다시 실무에 적용할 수 있는 어떤 표준화된 툴을 개발하자는 제안이다.

이 제안과 함께, 본고는 디지타이제이션 자료보존과 디지털 이미지품질 관리를 위한 실무에 응용할 수 있는 디지털 이미지와 메타데이터 결정의 실무전략을 예를 들어 설명하고, 디지타이제이션 프로젝트 진행에 관련된 중요한 쟁점들을 강조하여 실무자들이 디지타이제이션을 실행하는데 필요한 노하우를 제공하고자 한다.

## 2. 디지타이제이션

디지타이제이션은, 기술적으로 설명하면, 아날로그로 된 자료의 내용을 0과 1로 된 이진수로 표현하여 컴퓨터가 읽을 수 있도록 전환하는 간단한 것이다(Hughes, 2004, p.4).

디지털데이터는 포맷이 다양하여 여러 가지 매체에 담아서 디스플레이 될 수 있다. CD, DVD 등 다양한 디지털 방식의 매체들이 그 예이다.[1)]

Institute of Museum and Library Services Workshop(2000)의 보고서에 따르면, 디지타이제이션은 기관과 자료에 여러 가지 장점을 가져다준다고 설명한다. 즉 디지타이제이션은 우선, 소장자료에 접근성을 높여서 다양한 이용자들을 확보하여 이용자의 수가 기존방식보다 월등하게 증가하였다. 자료보호의 측면에서는 손상되고 상태가 나쁜 원본 자료들의 마모나 상태악화를 막아 원본을 보호해 주고, 이용자에게는 디지털로 된 이미지대용물이 이용자의 이용을 촉진시켜주었다.

Hughes는, 디지타이제이션이 자료를 소장한 기관에 주는 이점으로, 기관간의 협력을 통해 기존자료의 부족한 갭을 보완해준다고 설명한다(Hughes, 2004, pp.19-29). 이를 통해 필요한 기술을 보완하고 교육과 연구의 기회를 제공함으로 기관의 위상을 끌어올리는 계기가 된다고 한다. 또한 Saffady는 특히 자료에 대한 온라인 원거리접근을 통해 소장품관리나 실무업무에 중요한 이점으로 작용함을 강조한다(2002, p. 13).

이러한 디지타이제이션을 프로젝트경영관리의 관점에서 설명하면, 디지타이제이션은 한마디로 원본자료와 거기에서 추출된 기술을 새로 만들어진 디지털대용물에 서로 연결시키는 작업이라고 본다(Hughes, 2004). 원본자료는 목록상에 기록되고, 포장이 되거나 제본이 되어 있다.

디지타이징 하기 전에 원본자료는 제본을 뜯고, 스테이플 철침을 제거하고, 깨끗이 해야 한다. 디지털화될 각 자료는 목록에 첨가되고 고유번호가 부가되고 디지털파일의 파일명과 서로 상응해야한다.

원본자료를 스캔하여 디지털이미지를 만들어낸다. 이것은 디지털 캡쳐라고 부른다. 스캔은 원본으로 직접할 수도 있고 원본상태가 적절하지 않을때는 다른 자료를 이용하기도 한다. 원본의 상태가 아주 열악할 때는 스캔 작업이 원본에 치명적인 손상을 줄수도 있기 때문이다. 처음에는 원본에 대한 내용기술을 가능한 한 상세하게 하고, 그 중에서 기술내용을 선별적으로 골라 원본자료와 디지털이미지를 연결시킬 메타데이터를 결정한다.

메타데이터는 이용자가 검색을 통해 디지털 이미지에 접근하도록 접근점 역할을 함으로, 메타데이터가 없다면, 원본자료와 디지털 이미지 모두, 심지어 디지타이제이션 작업 자체도 무의미하게 될 것이다. 이런 중요한 이미지 기술작업이 메타데이터의 기술정도에 대한 합의없이 실행되어지고 있는 안타까운 실정이다.

Besser는 그의 연구에서 메타데이터표준의 부적절한 사용과 일관성이 없는 자료기술, 자

1) 디지털 형태의 저장 매체에 대해서는 Saffady, William. 2002. Managing Electronic Records(MER). 3rd ed. Lexana, KS: ARMA International, 2002. pp.19-52. 참조.

료접근의 기술적 장애로 인하여 상호 운용성(interoperability), 장기보존성(longevity), 그리고 접근성(accessibility)이 보장되지 못하고 있다고 지적한다(1997). 디지털이미지가 메타데이터와 연계되지 못하고 있기도 하고, 특히, 메타데이터 기술시 디지털이미지의 내용뿐만 아니라 이미지의 물리적인 세부사항들, 예를 들면, 파일포맷, 파일크기, 압축 등의 개별 아이템으로의 파일세부사항도 함께 자세히 기술하여함을 강조한다. 그리고 낱개 이미지뿐 아니라 소장품전체에 내용기술의 일관성을 유지하는 것이 소장품과 시스템간의 상호 운용성을 보장하는 가장 중요한 핵심이다.

최근 2005년 3월에 캐나다에서 아주 흥미로운 자료조사가 보고 되었다. 캐나다 〈퀘벡주 박물관협회(Sociètè des Musèes Quèbècois)〉는 지난 2년 동안 캐나다 문화재정보네트워크기관(Canadian Heritage Information Network (CHIN)에 가입한 회원 박물관소장품 데이터베이스 전체를 조사하여, 자료 기술 시에 가장 빈번하게 사용되어지는 메타데이터 필드를 조사하고, 그 중요도 우선순위를 보고하였다. 이 결과를 바탕으로 CHIN에서는 캐나다전역의 박물관소장품을 저장하는 하나의 통합데이터베이스를 구축할 계획이다. 다음 〈표 1〉은 상위13개의 메타데이터와 우선순위를 잘 보여주고 있다.

가장 빈번하게 쓰이는 메타데이터 필드의 순서가 그림, 조각과 같은 예술품이나 문화재 유물과 같은 공간을 차지하는 입체형임을 고려할 때, 첫째는 작품명과 작품타입이고, 카테고리와 하위카테고리, 작가, 날짜, 출처 등의 순서인 것이 타당해 보인다. 여기까지의 상위순위는 어떤 종류의 소장품이라도 유사할 것이다.

이를 바탕으로 나온 최종 데이터베이스의 표준화된 메타데이터 필드의 결과가 사뭇 궁금해진다.

이조사의 결과에 동감하면서, 디지털화된 자료를 기술할 때 어떤 필드를 반드시 포함시켜야할지 박물관의 메타데이터 표준은 어떤 필드

〈표 1〉 빈번하게 사용되는 메타데이터와 우선순위*

| 필드 | 표준화의 용이성(easiness) | 우선순위 | 선정이유 |
|---|---|---|---|
| 작품명/타입 | 어려움 | 1 | 연구에 중요 |
| 카테고리/하위카테고리 | 중간 | 2 | 연구에 중요 |
| 작가/예술가 | 상대적으로 쉬움 | 3 | 연구에 중요 |
| 받은 날짜/보낸 날짜 | 쉬움 | 4 | 연구에 중요<br>표준화에 용이 |
| 출처/국가 | 상대적으로 쉬움 | 5 | 연구에 중요 |
| 주제/이미지 | 중간 | 6 | 연구에 중요 |
| 기간 | 중간 | 7 | 연구에 중요 |
| 매체 | 상대적으로 쉬움 | 8 | 표준화에 상대적으로 용이 |
| 지원자 | 상대적으로 쉬움 | 9 | 표준화에 상대적으로 용이 |
| 재료 | 중간 | 10 | 연구에 중요 |
| 문화 | 중간 | 11 | 연구에 중요 |
| 테크닉 | 어려움 | 12 | 연구에 중요 |
| 학파/스타일 | 중간 | 13 | 연구에 중요 |

* 원문보고서의 표를 축약, 수정한 것임.

로 이루어질지 실무자들에게 큰 보탬이 될 것이라 기대된다.

## 3. 디지털캡쳐와 이미지기술의 실무

디지털자료의 특성을 기술·기록하는 것은 그 자료를 생산, 이용, 관리, 기록, 보존하는데 아주 중요한 작업이다(Beamsley, 1999). 보존관리할 자료의 특징에 맞추어 그것에 알맞은 메타데이터를 정하여 자료의 지식적인(intellectual) 내용과 물리적인(physical) 형태를 기술하는 것은 실무자에게는 가장 중요한 작업이 된다. 현재 사용되고 있는 메타데이터표준들을 살펴보면, 흥미롭게도, 도서관 커뮤니티는 메타데이터표준이 너무나 많아서 적절한 선택이 힘든 반면에, 박물관 커뮤니티는 소장품 기술을 위한 적절한 메타데이터표준이 부족하여 선택이 어렵다. 범람이든 고갈이든 문화유산기관이 디지타이제이션 자료 기술을 위한 가장 적절한 메타데이터표준을 찾는 것은 여전히 어려운 실정이다.

이 장에서는 이미지표준과 메타데이터표준을 마련하기 위한 현재 커뮤니티에서 가장 널리 이용되는 기본적인 사항을 소개하겠다.

### 3.1 디지털 이미지

좋은 디지털 이미지를 만들려면 디지털 기술에 관련된 여러 가지 기술적인 요소들을 이해해야 한다.

이미지의 질은 여러 가지 요소에 의해 복합적으로 그 결과가 결정되기 때문이다(Hughes, 2004; Stielow, 2003; Koelling, 2004; Lee, 2001). 예를 들면, 디지털 파일의 해상도, 칼라 비트심도, 압축, 파일 크기, 파일포맷 등이다. 이 중에서 중요한 요소들을 설명하겠다.

픽셀(Pixel)은 컴퓨터 화면에 표시될 수 있는 화상의 최소요소로 화소(畵素)라고 하며, 2차원 매트릭스에 비트맵으로 표시하여 1비트맵 속에 색깔이나 음영이 몇 개나 있는지 그 정도를 나타내준다.

해상도(resolution)는 이미지를 디스플레이할 때 사용하는 픽셀의 수를 말하며 이미지파일의 상세정도를 표시해준다.

프린트할 때 색의 선명도를 말하는 단위는 1인치 내 점의 수를 dpi(dots per inch)로 표시하고, 컴퓨터 모니터 화면상의 선명도를 말할 때 1인치 내 픽셀의 수를 ppi(pixels per inch)로 표시하여 이미지파일의 해상도를 나타낸다. Dpi나 ppi는 숫자가 클수록 더 해상도가 높아 이미지가 선명하고 정확하다. 해상도는 자료의 크기와, 상세정도, 자료의 목적에 따라 바뀌질 수 있다.

비트심도(bit-depth)는 디지털화 되었을 때의 색깔의 선명도를 말하며, 1bit는 한 픽셀내의 색깔의 선명도를 2진수의 제곱 승으로 표시한다. 즉, 1bit 이미지는 한 픽셀 내에 음영 없이 검은색과 흰색 2(=21)가지로만 이루어진 색깔이다. 8bit 이미지는 한 픽셀 내에 256가지 음영으로 이루어진 색깔이다(28=256).

압축(compression)은 한 공간에 데이터를 더 많이 넣어 파일크기를 줄일 수 있도록 압축정도를 말하는 것으로, 압축을 해도 이미지의 품질을 잃지 않는 로스리스(lossless) 압축과 파일크기가 많이 줄면서 이미지의 질을 떨어뜨리는 로시(lossy)압축이 있다.

**3가지 파일 만들기: 매스터(master), 액세스(access), 섬네일(thumbnail) 파일**

디지타이제이션 프로젝트를 수행할 때, 디지털 이미지는 1개가 아니라 보통 3개에서 4개의 파일을 만들어야 한다. 즉, 매스터파일, 액세스파일(1-2종류), 섬네일파일이 그것이다.

매스터파일은 원본을 최대한 잘 보존하기 위한 최고의 선명도와 원본에 가장 가까운 파일크기를 가진 최고의 디지털 대용물로 만들고, 다른 파일들과 같은 서버나 네트워크상에 저장하지 않고, 안전한 개별서버나 오프라인 데이터베이스에 보관한다. 매스터파일은 파일사이즈를 전혀 줄이지 않고, 또한 압축도 하지 않는 것이 원칙이다.

액세스파일은 이용자가 웹상에서 검색하여 원본대용물로 접근할 수 있도록 스크린 크기로 만든다. 기관에 따라서는 액세스파일을 풀스크린 크기와 50%-70%크기로 2가지 종류를 만들어 서비스하기도 한다. 어느 정도의 선명도를 유지하면서 이용자가 파일을 빨리 로딩할 수 있도록 원래의 파일보다 파일의 규모를 반정도로 줄이는 경향이다.

섬네일파일은 웹상에서 이용자가 검색할 때 인덱스 기능을 할 수 있도록 만든 파일로, 내용을 인식할 수 있는 정도의 길이로 즉, 보통 가로세로 1인치 정도로 만들고, 검색시 파일을 찾아주는 인덱스기능만 감당하면 되기 때문에 파일의 byte크기는 최대한 작도록 만든다. 액세스와 섬네일파일은 선명도보다는 파일의 크기를 줄이는 것이 중요하므로 액세스파일은 압축을 적당량하고 섬네일은 가능한 많이 압축한다.

**파일포맷**

다양한 종류의 파일포맷이 업계에 존재하고, 표준이 기술의 변화에 따라 변하고 있다. 대표적인 이미지파일포맷은 주로 TIFF(Tagged Image File Format), JPEG(Joint Photographic Experts Group), GIF(Graphics Interchange Format), PNG(Portable Network Graphics)가 쓰이고 있다.

TIFF파일은 고품질의 디지털이미지를 캡처할 수 있는 장점이 있어, 특히 아카이브즈에 장기보존용으로 만들어지는 매스터파일에 좋은 포맷이다(Kenney and Rieger, 2000).

JPEG파일은 네트워크상에서, 웹이나 소프트웨어에서 디스플레이용으로 좋은 포맷이다.

JPEG와 GIF는 로시압축(lossy compression)을 하여 파일크기를 줄이고 이미지를 검색하여 볼 수 있는 포맷으로 널리 쓰인다. 액세스파일과 섬네일파일은 보통 JPEG이나 GIF포맷으로 저장하고, PNG포맷은 파일이 크거나 멀티미디어에 많이 이용된다.

디지털파일의 표준포맷에 대한 결정은 전문가로부터 기술적인 컨설팅을 받아 디지타이제이션 프로젝트가 시작되는 초기단계부터 결정하는 것이 파일관리에 좋다.[2)]

JIDI Feasibility Study는 디지털파일의 적당한 해상도비교에 대한 설명과 원본자료의 질을 유지하기위한 기본적인 지침을 담고 있다(Tanner and Robinson, 1998). 또 코넬대학교 도서관의 디지털이미지 만들기 튜토리얼을 보면 이미지의 선명도를 서로 비교하여 설명하고 있다. 디지털이미지의 품질을 유지하는 여러

2) 파일포맷 비교 지침은 National Institute of Standards and Technology의 Digital Media File Type: Survey of Common Formats <http://www.itl.nist.gov/div895/isis/filetypes.html>을 보라.

〈표 2〉 디지털파일 가이드

| | 파일포맷 | 해상도 및 픽셀 | 파일크기 |
|---|---|---|---|
| 매스터파일 | 압축안함<br>TIFF포맷 | 원본에 가장 근접한 최고해상도 유지<br>3000-5000 픽셀 | 원본크기대로 |
| 액세스파일 | JPEG(8 bit, 혹은 24 bit)<br>GIF(1-8 bit, bitonal, indexed gray, indexed color)<br>PDF | 800 × 600 ppi<br>150-250 ppi<br>120 dpi | 100kb-500kb |
| 섬네일파일 | JPEG(1 bit gray 혹은 8 bit color) | 150-200 pixels 72 ppi | 50kb보다 작게 |

연구에서 나온 결과들을 집계, 종합하여 위와 같이 〈표 2〉를 제안한다.

자료를 스캔하여 디지털파일로 바꿀 때, 여러 가지 기술적요소가 관련하여 그 질을 좌우할 수 있다.

스캐너와 프린트기의 조건에 따라, 스캐너가 아무리 좋아도 디스플레이하는 모니터의 품질에 따라서도 다르게 디스플레이 될 수 있으며, 이용자 입장에서 검색할 때, 모니터상태와 조건에 따라서 다양하게 다르게 보여 질 수도 있다. 그러나 프로젝트 실무담당자는 항상 자료제공자의 입장에서 이미지의 품질을 유지하는 방안을 찾도록 노력해야한다.

디지털이미지의 품질과 이미지 생산가격 사이에는 불가분의 관계가 있다.

큰 파일은 선명도를 높여 고질의 이미지를 이용자에게 제공하지만, 낱개 이미지 한 개에 1Kb를 높일 경우, 전체 소장품을 완성하는데 필요한 부가적인 다른 스텝들의 업무량, 업무시간, 저장서버, 이용자의 로딩시간을 모두 고려한다면 비용과 시간은 급진적으로 늘어날 수 있다. 또, 디지타이제이션 전체 프로젝트의 흐름도에도 영향을 줄 수가 있다. 항상 적당한 선에서 타협하여 이미지의 선명도와 파일크기를 유지하되, 또 동시에 비용을 고려하는 절충된 선을 선택하는 것이 가장 중요하다.

다시 말해서, 디지타이제이션 프로젝트는 비용을 떠나서는 생각할 수 없으므로 의사결정과정에서 실무자는 항상 디지털이미지의 품질을 유지하고, 또 그에 따른 비용을 계산하여야 한다. 그리고 동시에 디지타이제이션의 목적이 기관과 이용자의 이용도 증대에 있음을 잊지 않도록 하고, 이용자들의 대부분은 보통의 모니터와 프린터를 가지고 소장품을 검색하고, 이용자그룹의 장비수준도 다양할 수 있음을 기억하여 이들을 넓게 만족할 수 있는 품질을 유지하도록 해야한다(NOF-Digitise, 2002).

### 3.2 메타데이터 기술

메타데이터는 간단히 '정보에 관한 정보', '데이터에 관한 데이터'라고 정의할 수 있다.

디지털환경에서 메타데이터는 디지털자료를 기술하여 이용자가 검색하고, 관리자가 운영관리를 용이하게 하도록, 주로 아키비스트나 사서에 의해 원 자료에 부가되는 정보이다(Hughes, 2004).

메타데이터는 자료의 원래 지식적인 부분에 관한 내용들, 즉, 작가나 주제, 날짜, 카테고리, 학파, 출처, 키워드, 이용자, 편집, 출판사, 저작권, 보유기관 등을 기술한다. 또 디지털파일의 물리적인 세부사항인 파일포맷, 파일크기, 포

맷, 해상도, 색상, 압축, 구조, 소프트웨어, 애니메이션, 음성파일 등도 기술한다.

잘 만들어진 메타데이터는 그 자료를 장기간 이용되고 보존되기 위한 필수요소가 된다. 메타데이터는 일반적으로 이용목적에 따라 임의적으로 기술(記述)적(descriptive) 메타데이터, 업무관리적 메타데이터, 기술적(technical) 메타데이터, 보존용 메타데이터와 활용 메타데이터로 나누기도 한다(Gilliland-Swetland, 1998).

자료에 메타데이터를 부가하는 것은 어떤 의미에서는 디지털 파일자체를 만드는 작업보다도 디지털파일을 사용할 수 있도록 가치를 부여하는 더 중요한 작업이다. 그러므로 메타데이터 결정은 프로젝트의 초기단계에서부터 전문가팀을 구성하여 시간과 노력을 아끼지 않고 처음부터 메타데이터표준을 작성하는 것이 바람직하다. 왜냐하면, 디지타이제이션 프로젝트의 성공여부는 메타데이터에 달려있다고 해도 과언이 아니며, 장기간 프로젝트가 진행될 때 메타데이터기술담당직원의 인건비로 가장 많은 비용이 들어가고, 실수를 하여 수정을 하는 실제 시행착오가 가장 빈번한 부분이기도 하기 때문이다.

이런 메타데이터에 대한 연구는 활발히 진행되어, 세계적인 연구팀에 의해서 여러 가지 메타데이터 표준이 제안되었다.

개별 표준에 대한 상세한 설명은 본고에서 하지 않겠으나, 간략하게 몇 예를 소개하면, 우선 널리 알려진, Dublin Core는 가장 기본적인 15개의 공통분모요소로 구성되어 여러 문화유산기관에서 많이 채용하고 있다. Dublin Core는 웹상에서 자료검색의 목적으로 개발되었고, 널리 사용됨으로 어떤 종류의 디지털자료에도 이용될 수 있다.

METS(Metadata and Encoding Transmission Standard)는 디지털라이브러리에서 종류에 상관없이 디지털 자료를 인코딩하기 위한 목적으로 XML텍스트로 기술하는 표준이다.

〈표 3〉은 XML로 기술된 METS문서의 예이다.

〈표 3〉 METS 텍스트 예[3)]

```
〈?xml version="1.0" standalone="no" ?〉
〈mets:mets xmlns:mets="http://www.loc.gov/METS/"
      xmlns:mods="http://www.loc.gov/mods/"
      xmlns:xlink="http://www.w3.org/TR/xlink"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xmlns="http://www.loc.gov/METS/"
      xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd
      http://www.loc.gov/mods/ http://www.loc.gov/standards/mods/mods.xsd"
      OBJID="ark:/13030/kt9s2012hz" TYPE="cartographic" LABEL="San Francisco 15-minute Quadrangle"〉
      〈mets:metsHdr CREATEDATE="2001-10-23T00:00:00" LASTMODDATE="2003-04-18T07:00:00"〉
〈mets:agent ROLE="CREATOR"〉
〈mets:name〉Rick Beaubien〈/mets:name〉
〈/mets:agent〉
〈/mets:metsHdr〉
〈mets:dmdSec ID="DM1"〉
〈mets:dmdSec ID="DM2"〉
〈/mets:dmdSec〉
    :
```

3) METS 공식 사이트에서 참조. 〈http://www.loc.gov/standards/mets/mets-examples.html〉

전형적인 기술메타데이터의 예로는 MARC (Machine Readable Cataloging)와 EAD(Encoded Archival Description)를 들 수 있다.

EAD는 SGML로 시작하여 요즘은 XML로 텍스트를 작성하는 아카이브즈를 위한 전자 finding aids로써, 아카이브즈 기록물이나 수서본의 인코딩을 위해 개발되었다.

TEI(Text Encoding Initiative)는 다양한 포맷의 문서를 인코딩하기 위한 텍스트포맷용으로 개발되어 인코딩표준으로 다른 메타데이터 표준들의 인코딩에 이용되고 있다.

The VRA(Visual Resources Association) Core Categories Version 3.0는 주로 시각자료의 기술을 위해 개발, 이용된다. 이 외에도 MODS, CDWI, MPEG, RDF이 널리 알려져 있고, 이들 외에도 분야별로 여러 가지가 있다.

이런 메타데이터는 위에서 설명한 3가지 종류의 파일과 함께 디스플레이되는데, 상세함과 적용도가 아주 다양하여 실무담당자가 기관에 따라서 상세함 정도를 결정하여야 한다. 예를 들면, 섬네일파일은 인덱스로 사용되는 작은 이미지와 파일의 간단한 표제정도만 표기해주는 것이 좋다. 기관에 따라 다르지만 보통, 제목, 작가, 날자, 저작권 등이 주로 쓰인다. 액세스파일은 좀더 상세한 내용이 함께 기술된다. 원본에 대한 최대한 상세한 기술을 액세스파일과 보여주기도 하고, 또 경제적인 면을 고려하여 선별된 몇 필드만 골라서 디스플레이된다.

〈표 4〉 EAD 예[4)]

[Invitation to 84th anniversary celebration of mexican independence, 1894]
Creator: Hernandez, Frank X.
Date: 1894
Physical Description: Invitation 11 × 17 cm.
Notes: At top of front page of invitation: Mexican flag, with green at left, eagle in middle, and red at right, "Oct. 16, 1894," written below flag.
Subjects: Diversity; Ephemera; Invitations; Mexican Americans; Organizations--California--San Jose; Santa Clara County(Calif.)--
History;
Place of Origin: ark:/13030/kt809nc9gn
Language: Spanish
Identifier(s): casjhsj 1997-220-5
Source: local, tgm Ⅰ-Ⅱ
From: Silicon Valley History Online History San José Silicon Valley History Online collection
Originally Published By: History San Jose
Contributor(s): History San Jose
Repository: History San Jose Research Library
Suggested Citation: [Identification of Item]. Available from the Online Archive of California; http://ark.cdlib.org/ark:/13030/kt809nc9gn.
Terms and Conditions of Use: History San José can only claim physical ownership of the collection. Users are responsible for satisflying any claims of the copyright holder. Permission to copy or publish any portion of History SanJosé collection must be given by History San José.

4) EAD 공식 사이트에서 참조. 〈http://ark.cdlib.org/ark:/13030/kt809nc9〉

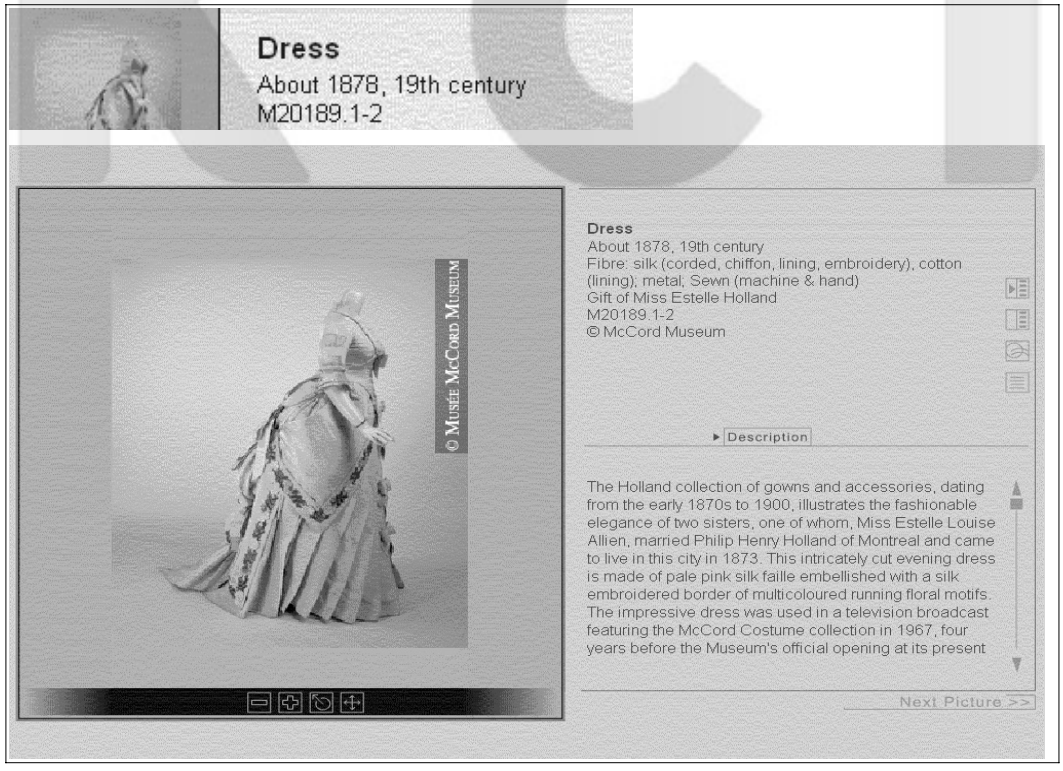
메타데이터 단계에서 가장 중요한 결정이 바로 이 액세스파일과 함께 디스플레이 될 메타데이터의 상세함 정도를 결정하는 부분이다. 기술적으로 타당한, 기관의 여건에 맞는, 담당할 인력이 있는지, 재정적인 지원이 있는지 등의 질문과 함께 적절한 타협선에서 결정하여야 할 것이다. 이미지의 품질과 그에 상응하는 적절한 상세기술을 제공하여 이용자의 요구를 만족시킬 수 있고, 또 동시에 기관의 프로젝트 운영의 경제적, 전체적인 고려하여 결정하여야 한다.

캐나다 몬트리올에 소재한 〈McCord Museum〉에서 만들어진 액세스파일과 섬네일 파일의 메타데이터기술예가 〈표 5〉에 잘 나타나 있다.

디지털라이브러리 환경에서 사용되는 메타데이터를 전체적으로 이해할 수 있는 좋은 예로는, 「Harvard Library Digitization Initiative의 안내매뉴얼」로 문화유산기관에서 다양하게 이용할 수 있는 실무지침안내가 잘 나와 있다.

Digital Library Federation은 효과적인 디지털라이브러리 소장품관리를 위한 실무 전략 가이드를 종합적으로 제공하고 있다. CHIN은 앞서도 말했듯이 캐나다의 국립기관으로 좋은 지침을 많이 제공하는데, 메타데이터, 메타데이터어휘, 분류, 데이터포맷, 데이터교환 등에 관해 설명하고, 특히 박물관분야에서 이용할 수 있는 메타데이터표준, 목록, 소장품관리 등의

〈표 5〉 엑세스파일, 섬네일 파일의 메타데이터 기술[5)]




5) McCord 박물관 사이트에서 참조. 〈http://www.mccord-museum.qc.ca/en/collection〉

안내지침과 메타데이터표준들을 모아서 보여주고 있다. 영국의 Arts and Humanities Data Service는 인문대학과 시각예술분야에서 이용할 수 있는 실무전략지침을 제공하고 있다.

## 4. 주요 쟁점

### 4.1 메타데이터 품질유지

현재 존재하는 메타데이터표준은 수적으로는 너무나 넘쳐나지만, 그중에서 개별기관에서 필요한 가장 적절한, 아니면 최소한 그럴만한 하나를 선택하자면 찾기가 힘들다는 것이 일반적으로 실무자들이 메타데이터에 대한 경험이다. 선택의 폭이 많아서 선택이 더 어려워지기도 하고, 한기관의 기대치에 완벽하게 맞는 맞춤형 표준은 없기 때문이다. 그래서 적절한 메타데이터를 만들고자 분야별로 계속 더 표준을 만들어내어 메타데이터스킴 범람을 계속하고 있다.

실무담당자들에게 메타데이터에 관련하여 유의해야할 점은 한 자료를 기술할 때 전체 소장품 내에서 무엇보다도 기술의 일관성을 유지하는 점이다. 이 점을 항상 염두에 두어야 한다. 왜냐하면, 메타데이터기술을 실제로 해보면 어휘선정에서도 일관성이 힘들고, 현재 나와 있는 메타데이터 표준들 간에 필드명을 정의하는데도 일관성이 부족하고, 필드의 구조를 또한 일관되게 하지 못하고 있음을 많이 보게 된다. 그래서 한 자료그룹에 대해서는 같은 아키비스트가 기술을 하도록 하고 기술의 원칙을 프로젝트 처음부터 정하여 일관성을 유지하도록 한다.

연구자들의 관점에서 보면, 이 문제들은 오랫동안 논의되어 온 메타데이터의 연구주제이고 여전히 해결하지 못한 과제이다. 또 주제어를 선택할때 기존에 널리 쓰이는 어휘집을 선택하는 것이 중요하다.[6)]

### 4.2 품질관리

디지털파일의 품질은 여러 가지 요소에 의해 영향 받는다. 예를 들면, 원본자료, 스캐닝기기, 스캐닝기술세부사항, 이미지편집, 메타데이터, 그리고 디스플레이 설비 등이다.

디지털파일의 품질은 이미지품질과 비용간의 타협의 결정을 내리는 것이라고 볼 수 있다. 디지타이제이션으로 인해 발생하는 파일저장, 부가적인 비용과 시간에 대해서 면밀하게 검토하여 신중히 선택하여야 한다.

또, 각 단계에는 예상했던 것보다 항상 더 요구사항이 많고 비용이 많이 든다. 디지타이제이션 프로젝트 수행시 품질관리는 전체 프로젝트흐름도 과정에서, 이미지 캡쳐과정에서, 메타데이터 기술과정에서 일관성 있게 종합적으로 고려하여야 한다.

## 5. 결 론

본고에서는 '디지털이미지 만들기'와 '메타

6) 대표적인 분류스킴이나 시소러스로 미국 의회 주제분류표, 'The Art and Architecture Thesaurus', Getty Information Institute의 'Thesaurus of Geographic Names', 국제박물관협회(International Council of Museums)의 'International Committee for documentation standards for museum documentation'이 있다.

데이터 기술과정'에 관한 기본적인 지침을 예와 함께 설명하고, 디지타이제이션 프로젝트를 진행할 때 보장해야 할 품질관리에 관하여 설명하였다.

디지타이제이션은 다단계의 복잡한 과정으로 이루어지며, 기술적인 분야만아니라, 종합적으로 프로젝트운영의 관점에서 접근해야 성공적으로 끝낼 수 있다. 각 단계마다 실무운영자는 의사결정을 계속 해야 하고, 기관은 소중한 자료들을 장기보존할 수 있는 소장품 관리운영지침을 마련해야한다. 전체 프로젝트에서 품질관리를 하는 것은 일관성을 유지시켜나가고 기관책임자와 협조체제를 유지하고 대화함으로 가능하다.

본고를 쓸 때, 처음에 생겼던 두 가지 질문을 이 글을 맺으면서 실무담당자에게 해답을 찾기 위해 되물으려고 한다. '자료를 장기간 유지하고 보존하기 위해서, 자료의 어떤 특징을 알아야하는가?', 그리고 '자료의 개념적인 내용을 기록하고 자료의 진본성을 보장하기 위해 어떤 메타데이터가 필요한가?' 하는 질문이다. 이것은 실무자들이 직접 해답을 찾기 바란다.

마지막으로, 이 분야에서 앞으로 더 개발해야 할 연구과제는 표준화된 디지타이제이션을 수행하기 위한 효과적인 툴을 개발하고, 특정 커뮤니티를 위한 표준프로그램을 설계하고, 이것이 나아가 다른 커뮤니티에서도 널리 쓰이는 툴을 개발하는 것이다.

디지타이제이션 프로젝트가 진행하는 각 기관마다 성공적으로 수행되어 소장품들을 더욱 잘 보존하고 뛰어난 이용자서비스를 제공함으로 실제적인 열매를 많이 맺기 바란다.

## 참 고 문 헌


Beamsley, Teresa G. Securing digital image assets in museums and libraries:A risk management approach. Library Trends, 48, 2(1999), 358-78.

Besser, Howard. The Changing Role of Photographic Collections with the Advent of digitization.(The Wired Museum, Washington: American Association of Museums, 1997). [cited 2005.8.15] 〈http://www.gseis.ucla.edu/~howard/Papers/garmil-eastman.html〉 August 15, 2005.

Canadian Heritage Information Network, 〈http://www.chin.gc.ca/English/Standards/introduction.html〉

Cornell University Library. Research Department. Moving Theory into Practice: a Digital Imaging Tutorial. 〈http://www.library.cornell.edu/preservation/tutorial/contents.html〉

Digital Library Federation(DLF), 〈http://www.diglib.org/standards.htm〉

Dublin Core Metadata Element Set, 〈http://dublincore.org〉

Encoded Archival Description(EAD), 〈http://www.loc.gov/ead〉

Gilliland-Swetland, Anne. "Setting the Stage." From Introduction to Metadata: Pathways to Digital Information. Baca, Murtha, ed. Los Angeles, CA: Getty Research Institute, 1998: 1-12.

Harvard Library Digitization Initiative, http://hul.harvard.edu/ldi/html/metadata.html.

Hughes, Lorna M. Digitizing Collections: Strategies Issues for the Information Manager. Bodmin, Cornwall: Facet Publishing, 2004.

Kenney, Anne R. and Rieger, Oya.(ed). Moving Theory into Practice: Digital Imaging for Libraries and Archives. Research Libraries Group, 2000. [cited 2005.8.15] 〈http://www.rlg.org/preserv/mtip2000.html〉 August 15, 2005.

Koelling, Jill Marie. Digital Imaging: A Practical Approach. Walnut Creek, CA: AltaMira Press, 2004.

Lee, Stuart D. Digital Imaging : A Practical Handbook. New York, NY: Neal-Schuman Publishers, Inc, 2001.

MARC(Machine Readable Cataloging), 〈http://lcweb.loc.gov/marc〉

METS(Metadata and Encoding Transmission Standard), 〈http://www.loc.gov/standards/mets〉

NOF-Digitise Technical Advisory Service. The Digitisation Process. [cited 2005. 8.15] 〈http://www.ukoln.ac.uk/nof/support/help/papers/digitisation_process〉 August 15, 2005.

Société des Musées Québécois. Artefacts Canada Standardization: Analysis and Recommendations. 2005. 〈http://www.chin.gc.ca/English/Pdf/Standards/ACStandardization.doc〉

Stielow, Frederick. Building Digital Archives, Descriptions, and Displays. New York, NY: Neal-Schuman Publishers, Inc,. 2003.

The Institute of Museum and Library Services. Report of the Workshop on Opportunities for Research on the Creation, Management, Preservation and Use of Digital Content. September 2003. [cited 2005.8.15]. 〈http://www.imls.gov/pubs/pdf/digitalopp.pdf〉 August 15, 2005.

Saffady, William. Managing Electronic Records (MER). 3rd ed. Lexana, KS: ARMA International, 2002.

Tanner, Simon and Robinson, Brian. JISC Image Digitisation Initiative: Feasibility Study(Final Report). The Higher Education Digitisation Service. 1998. 〈http://heds.herts.ac.uk/resources/papers/jidi_fs.pdf〉 〈www.tei-c.org/Guidelines2/index.html〉

Visual Arts Data Service: Catherine Grout, Phill Purdy, Janine Rymer. Creating Digital Resources for the Visual Arts: Standards and Good Practice. [cited 2005.8.15]. 〈http://vads.ahds.ac.uk/guides/creating_guide/contents.html〉 August 15, 2004.

VRA(Visual Resources Association) Core Categories Version 3.0. 〈http://www.vraweb.org/vracore3.htm〉

KCI