

V. 디지털 아카이브

아카이브의 실물을 관리하는 입장에서 그것과 관련한 인문지식을 포용하고자 하는 시도는 기존의 정보과학에서 강도 높게 추구되어 왔다. 인문정보학은 그 반대 방향에서의 접근을 일차적인 소임으로 삼는다. 즉, 인문학 연구의 결과물이 디지털 네트워크상에서 아카이브학의 자원으로 재이용될 수 있도록 그것을 구조적으로 체계화하고 맥락화 하려는 것이다. 인문지식이 학계와 아카이브 두 세계에서 분업의 결과물로 나뉘기보다는 디지털 협업의 플랫폼 위에서 하나로 소통되게 하자는 것이 이 분야에서 인문정보학이 추구하는 목표이다.¹⁾

V-1. 디지털 아카이브란?

인문지식이 디지털 아카이브(Digital Archives)의 세계에 다가가기 위해서는 그 대상에 대한 이해가 선행되어야 한다. 디지털 아카이브란 무엇인가? 일각에서는 음원, 영상 등 디지털 신호로 만들어진 데이터를 보존하는 것을 디지털 아카이빙(Digital Archiving)이라고 이름하기도 하고,²⁾ 또 다른 영역에서는 아카이브 기능을 수행하는 곳(Institutional Repository)에서 소장품의 디지털 사본을 제공하는 경우, 이를 디지털 아카이브라고 부르기도 한다.³⁾

- 1) 아카이브에 소장된 기록물이나 박물관에서 전시하는 문화유산의 역사적, 문화적 가치를 탐구하고 드러내는 일은 대학의 연구자들이 '인문학'이라는 이름으로 추구해 오던 것과 다르지 않다. 이론 연구에 치중하는 인문학자들도 그 이론을 입증하는 증거를 아카이브나 박물관의 실물에서 찾았고, 수장고에 쌓인 유물들을 분류하고 그것의 역사적·문화적 맥락을 좇아 전시실을 꾸미는 큐레이터들도 자신의 직무에 필요한 지식을 인문학적 연구 성과에서 얻을 수 있었다. 이 양자 사이의 소통이 디지털 네트워크상에서 더욱 혁신적으로 이루어질 수 있다는 점이 인문정보학의 관심사이다.
- 2) 국내에서 '디지털 아카이브'라는 말이 쓰이기 시작한 것은 1990년대 후반, TV 방송사들을 중심으로 '영상물 데이터베이스' 구축에 대한 논의가 시작되면서부터이다. (1999년 문화관광부의 '디지털 아카이브 시스템 구축 사업'[디지털 방송 시대에 대비하여 각 방송사의 영상물을 디지털 신호로 데이터베이스화하는 사업] 등) 이 때, 디지털 신호로 이루어진 방송 영상 데이터와 그에 관한 메타데이터를 저장하는 데이터베이스를 '디지털 아카이브'라고 지칭하였는데, 이는 아카이빙 대상 자원이 디지털 형태임을 의미하는 것이다. 최근에도, 아날로그 매체(필름, 테이프, 음반)에 담겨 있던 영상물 또는 음원의 보존과 활용을 위해 그것을 디지털 신호로 전환하고 데이터베이스에 수록하는 경우, 이를 '디지털 아카이브'로 지칭하는 사례를 볼 수 있다. (2013년 11월 대한민국 국방부에서 구축·발표한 「6·25전쟁 영상 디지털 아카이브」 등)
- 3) 디지털 도서관, 디지털 박물관, 디지털 기록관 등을 표방하는 국내의 공공분야 정보시스템의 입장에서 '디지털 아카이브'의 개념을 정의한다면 그것은, '1차 자료에 관한 정보를 디지털 형태로 정리, 축적한 데이터베이스'라고 할 수 있다.
우리나라에서는 전자정부 구현의 정책적 추동에 힘입어 상당수의 공공기관이 기관 고유 업무의 전자정보화를 이루어냈고, 또 그곳에서 축적된 콘텐츠를 대국민 서비스 자원으로 활용하고 있다. 특히, 도서관, 박물관, 기록관 등 아카이브 기능을 갖고 있는 공공기관의 경우, 소장하고 있는 자료에 관한 정보를 수록한 데이터베이스를 구축하고 그 콘텐츠를 온라인 네트워크 상에서 국민들에게 제공하는 것

전통적으로 도서관⁴⁾, 박물관⁵⁾, 기록관⁶⁾은 각기 차별화된 기능을 수행하는 조직으로 인식되어 왔다. 도서관은 주로 출판물을 소장하며, 박물관은 역사성이 있는 유물을, 기록관은 사건의 증거가 되는 기록물의 원본을 보존한다. 고객에게 제공하는 서비스의 형태 또한 도서관은 대출, 박물관은 전시, 기록관은 열람의 방법을 위주로 하였다. 그러나 현대의 도서관, 박물관, 기록관은 그 기능이 확대되면서 상호간의 경계가 점차 모호해져 가고 있으며, 그 융합적인 기능을 지칭하는 새로운 용어⁷⁾까지 등장하였다. 그 기능을 디지털 세계로 옮겨 놓은 곳에서는 융합의 성격이 더욱 두드러진다.

현실 세계의 사물과 달리 디지털 콘텐츠는 어느 한 공간에 귀속될 필요가 없다. 의미적으로 관련이 있다면 그 원시 자료가 유물인지 책인지 기록물 원본인지에 구애됨이 없이 하나의 스토리를 갖는 지식 정보로 연결 지을 수 있다. 인문정보학이 봉사하고자 하는 디지털 아카이브는 전통적인 기록관 개념의 연장선상에 있다기보다 디지털 도서관, 디지털 박물관과 통용되는 융합적인 개념이다.

은 그 기관이 기본적으로 갖추어야 할 기능의 하나로 인식하게 되었다. 이와 같은 디지털 콘텐츠 서비스는 원시 자료 또는 기관의 성격에 따라 디지털 도서관, 디지털 박물관, 디지털 기록관 등으로 구분하여 이름 짓기도 하나 '아카이빙 자료'의 디지털 서비스라는 점에서 모두 유사한 성격의 '디지털 아카이브'라고 볼 수 있다.

- 4) 도서관(library): 도서관은 체계화된 정보 자료를 고객이 열람하거나 대출할 수 있는 곳이다. 이용자들이 직접적으로 자료를 열람(Physical Access)하거나 디지털화된 자료를 열람(Digital Access)할 수 있고, 열람 장소는 빌딩 또는 방(room), 또는 가상공간(Virtual Space), 또는 두 장소 모두가 될 수 있다. 도서관에 축적된 자료로는 도서, 정기간행물, 신문, 수고본, 영화, 지도, 유인물, 문서, 마이크로 필름, CD, 카세트테이프, 비디오테이프, DVD, 블루레이 디스크, 전자책, 오디오북, 데이터베이스, 그리고 기타 자료들이 포함된다. (Library, Wikipedia. <http://en.wikipedia.org/wiki/Library>)
- 5) 박물관(museum): 인류와 그들의 환경이 남긴 유무형의 유산을 교육, 연구, 오락의 목적으로 획득, 보존, 연구, 소통, 전시하는 곳이다. 사회봉사 및 발전을 위해 비영리적, 영속적, 개방적으로 존재한다. (Museum Definition, The World Museum Community. <http://icom.museum/the-vision/museum-definition/>)
- 6) 기록관(archives): 편지, 보고서, 메모(notes), 사진, 또는 그 밖의 원시 자료로부터 일차적인 사실(facts), 데이터, 증거 등을 얻을 수 있는 장소이다. (What's an Archives, The U.S. National Archives and Records Administration. <http://www.archives.gov/about/info/whats-an-archives.html>)
- 7) Larchiveum: library, archive, museum의 합성어. 2008년 미국 텍사스 대학의 Megan Winget 교수가 복합 수집기관(Multidisciplinary Collecting Institution)의 새로운 모델로 제안. 우리나라에서는 경상북도 문화콘텐츠진흥원이 2013년 12월 「라키비움(Larchiveum)」이라고 명명한 복합 문화공간을 개관하였으며, 안동시에서 건립중인 '세계유교문화박물관'도 라키비움 모델을 표방하고 있다.

V-2. 개방과 공유: 디지털 아카이브의 새로운 패러다임

디지털 데이터의 생산과 유통에 관한 국제사회의 선진적인 논의는 이른바 ‘오픈 아카이브’(Open Archives)⁸⁾, 또는 ‘링크드 오픈 데이터’(Linked Open Data)⁹⁾라고 하는, 개방과 공유의 세계를 지향하는 데 초점을 모으고 있다. 오픈 아카이브는 아카이브 기능을 갖는 기관(Institutional Repository)들이 표준적인 메타데이터를 생산하고 공유함으로써 이용자들로 하여금 장소와 조직의 경계에 구애됨이 없이 필요한 데이터에 자유롭게 접근할 수 있게 하는 것이며, 링크드 오픈 데이터(LOD)는 디지털 자원들이 서로에 대해 갖는 ‘의미’를 명시적으로 기술함으로써 인터넷상에 유관한 자원들로 엮인 광대한 의미망이 형성될 수 있도록 하려는 제안이다.

인문지식이 디지털 아카이브의 실물 정보와 소통하게 하려는 인문정보학의 노력은 바꿔 말해, 디지털 아카이브 영역에서 추구되는 새로운 패러다임을 인문학 지식의 생산과 유통에도 적용시키려는 것이라고 할 수 있다. 이를 위해서는 그 패러다임에 대한 이해가 선행되어야 한다. 이 점에서 OAI(Open Archives Initiative), 국제박물관협의회(International Council of Museums, ICOM), 유로피아나(Europeana) 등이 제시하는 학술과 문화, 예술 영역의 지식정보 유통 체계는 매우 유용한 참조 모델이자 학습 과제이다. 디지털 아카이브로 연장되는 인문정보를 전문적으로 연구하기 위해서는 그와 같은 선행 모델을 깊이 있게 들여다보아야 하겠지만, 인문정보학의 연구 주제를 안내하는 이 글에서는 각 기술의 용도와 특성만 간략히 소개하기로 한다.

1) OAI의 기술 표준: 웹 콘텐츠의 상호 운용성을 위한 표준

-
- 8) 실물 자원은 소장처라는 공간 안에서만 이용할 수 있는 데 반해, 디지털 자원은 온라인 네트워크 상에서 누구나 그것에 접근하고, 다양한 형태로 재활용(reuse)할 수 있다. 이러한 이유에서 ‘디지털 아카이브’는 단순히 디지털 복제물의 생산과 서비스 차원을 넘어서서 그것의 광범위한 공유와 재활용의 방법을 모색하게 되었는데, 이것을 가능케 하는 새로운 개념의 디지털 아카이브 네트워크를 ‘Open Archives’라고 한다. OAI(Open Archives Initiative)는 Library & Information Science 분야의 전문가 그룹으로서 Open Archives를 위한 기본적인 기술 표준을 제시하고 있다. (OAI 지원 기구: Andrew W. Mellon Foundation, Coalition for Networked Information, the Digital Library Federation, National Science Foundation)
- 9) 월드와이드웹 상에 존재하는 각각의 데이터에 유일한 식별자(Uniform Resource Identifier, URI)를 부여하고, 이를 이용하여 데이터 상호간의 연결이 가능하도록 하는 네트워크 모델을 ‘링크드 데이터’(Linked Data)라고 하며, 이것이 개방적으로 운용되는 것을 ‘링크드 오픈 데이터’(Linked Open Data, LOD)라고 한다. W3C(World Wide Web Consortium)가 이를 구현하기 위한 기술 표준을 선도하고 있다.

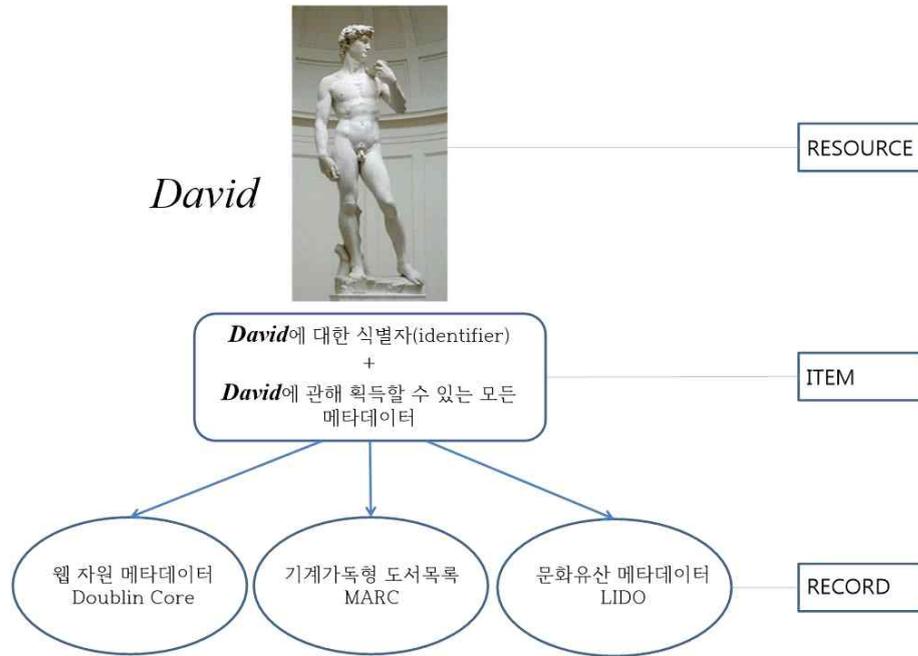
정보과학 분야의 전문가 그룹이 이끄는 OAI(Open Archives Initiative)는 디지털 콘텐츠의 효율적인 보급을 촉진하는 기술 표준을 개발하고 있다. OAI 활동은 원래 전자문서(e-print) 아카이브에 대한 접근을 용이하게 함으로써 학술정보의 소통을 증진시키려는 목적에서 출발하였는데, 현재는 다양한 디지털 자원의 개방적인 이용 환경을 조성하는 방향으로 연구 영역을 확대해 가고 있다. 미국의 앤드류 멜론 재단(Andrew W. Mellon Foundation), CNI(Coalition for Networked Information), DLF(Digital Library Federation), NSF(National Science Foundation) 등이 이 OAI 활동을 재정적으로 지원하며, 성과 확산을 돕고 있다.

- OAI-PMH: 메타데이터 교환을 위한 상호 운용성 확보 표준

OAI-PMH(Open Archives Initiative Protocol for Metadata Harvesting)는 여러 곳의 다른 아카이브로부터 그들이 가지고 있는 디지털 자원의 메타데이터를 수합하는 규약이다. 데이터를 제공하는 기관이 이 규약에 따라 자기 데이터를 구조화된 형태로 노출시키면, 그 데이터를 받으려는 측에서 역시 이 규약에 따라 데이터를 요청하여 획득할 수 있도록 하고 있다.

- OAI-ORE: 웹 자원의 집성체(aggregation)를 위한 기술 표준

OAI-ORE(Open Archives Initiative Object Reuse and Exchange)는 집성체 형태의 웹 자원을 기술하고 교환하기 위한 표준이다. 웹 자원은 텍스트, 이미지, 데이터, 비디오 등 분산되어 존재하는 멀티미디어 자원이 하나로 묶여진 형태(집성체: aggregation)로 만들어질 수 있다. 누군가가 이렇게 집성체 형태로 만들어진 웹 자원을 재이용(reuse)하고자 한다면 그것의 내용에 대한 충분한 정보가 있어야 할 것이다. OAI-ORE는 웹 자원의 재이용과 교환을 가능하게 하기 위해 그와 같은 집성체를 표준적인 방법으로 기술하는 틀을 제공한다.



OAI-PMH를 통해 취합할 수 있는 메타데이터 형식의 예시¹⁰⁾

2) LIDO: 문화유산 정보 스키마

LIDO(Lightweight Information Describing Objects)는 국제박물관협의회(ICOM, International Council of Museums)에서 제정한,¹¹⁾ 문화유산 정보의 기술을 위한 XML 스키마이다. 박물관의 유물 등 문화유산에 관한 정보를 명시적으로 전달하기 위한 목적으로 고안되었다.

오늘날 박물관들은 자체적으로 문화유산 정보를 제공하는 데 머물지 않고, 그것을 다양한 수준의 포털 - 주제별, 지역별, 국가적, 국제적 - 에 제공함으로써 종합적인 지식정보 네트워크를 이루어낼 필요성에 당면하고 있다. 그런데 개별 박물관들은 대부분 독자적인 유물관리 시스템을 운영하기 때문에 여기에 담긴 데이터가 그대로 온라인 네트워크 상에서 통용되기 어려운 실정이다.

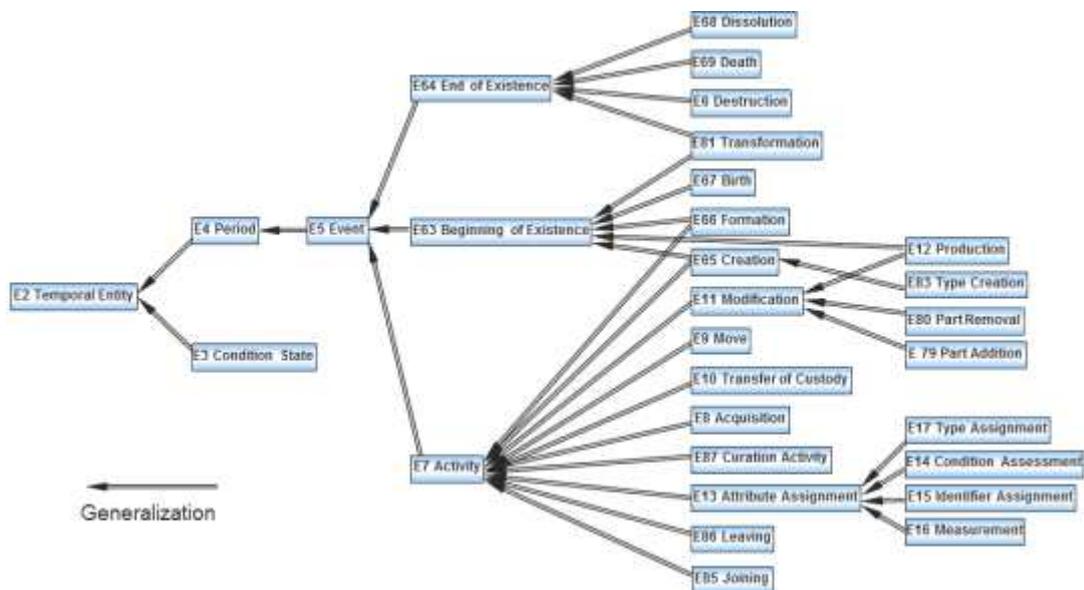
LIDO의 제정 목적은 여러 박물관이 표준적인 형식을 좇아 문화유산 데이터를 생산하게 함으로써 종합적인 포털 서비스와 데이터 교환을 용이하게 하자는 것이다. LIDO는 문화, 기술, 자연과학 등 모든 영역의 문화유산에 대한 설

10) Main Technical Ideas of OAI-PMH, Open Archives Forum, <http://www.oaforum.org/tutorial/english/page3.htm>

11) 2010년 국제박물관협의회(ICOM-CIDOC, International Committee for Documentation) 컨퍼런스에서 LIDO v1.0이 발표되었다.

명 정보(Descriptive Information)를 담을 수 있는 틀을 제공한다.

LIDO의 가장 중요한 부분은 CIDOC CRM¹²⁾에서 채용한 ‘event’ 개념이다. 예를 들어, 어떤 대상의 창작, 수집, 사용 등이 모두 ‘event’로 서술되는데, 이 ‘event’ 요소는 그 안에서 다시 언제(date), 어디서(place), 누가(actor) 했는지를 기술할 수 있게 하고 있다.



CIDOC CRM의 Event 클래스 구조¹³⁾

3) 유로피아나

유로피아나(Europeana)는 유럽 전역의 3,000여 박물관, 도서관, 기록관이 보유하고 있는 문화적 자료(도서, 그림, 영화, 그 밖의 박물관 소장 유물 및 기록물 등)의 디지털 콘텐츠를 서비스하는 가상의 디지털 박물관이다. 유로피아나의 설립은 유럽연합(EU) 차원의 문화 정책에 의한 것이다. 2008년 브뤼셀에서 열린 유로피아나 출범 각료회의는 유로피아나의 설립 취지를 다음과 공표하였다.

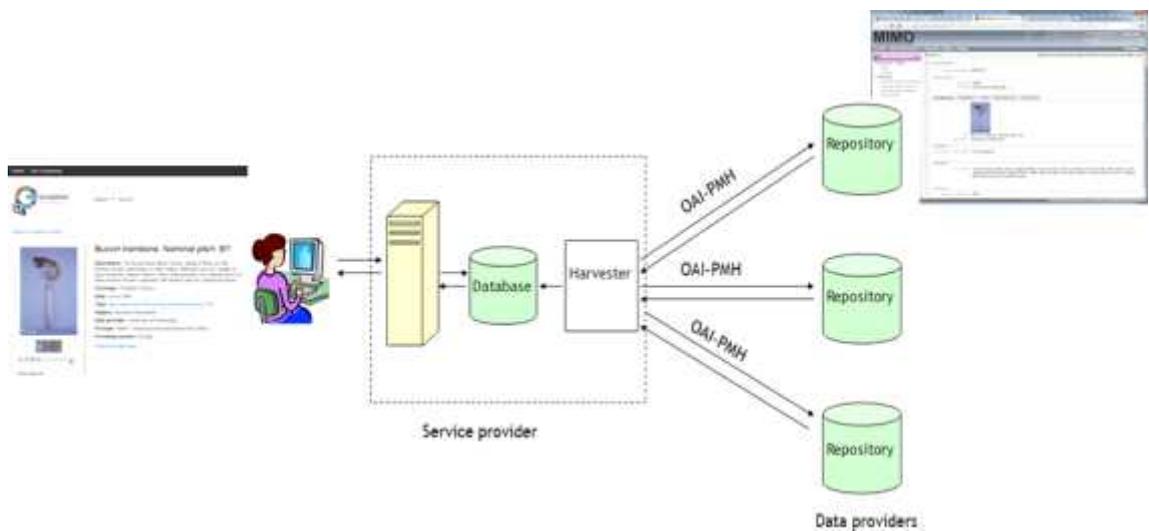
12) CIDOC CRM: Conceptual Reference Model. 국제박물관협회의 국제문서화위원회(International Committee for Documentation of the International Council of Museums)에서 제안한, 문화적인 개념의 데이터화를 위한 온톨로지. 86개의 클래스(class)와 137개의 속성(property)으로 이루어져 있다. 2006년 ISO에 의해 국제표준으로 채택되었다. (ISO 21127:2006)

13) CIDOC CRM Class Hierarchy, International Council of Museums, http://www.cidoc-crm.org/cidoc_graphical_representation/crm_class_hierarchy.htm

“디지털화와 온라인 접근은 문화 및 학문적 유산을 복원하고 새로운 콘텐츠의 개발을 촉진하며, 최신 온라인 서비스를 배양하는 중요한 채널이다. 이 두 요소는 정보 접근성을 민주화하고 정보사회와 지식기반 경제를 발전시키는 데 기여한다.” (2008. 11. 유로피아나 출범 유럽 각료회의. 브뤼셀)

유로피아나에서 서비스되는 모든 디지털 자료는 한 곳의 중앙 시스템에 저장되어 있는 것이 아니고, 유럽 각처에 있는 기관에서 네트워크를 통해 제공된다. 그 기관(도서관, 박물관, 기록관, 시청각 자료관 등)들은 각각 다른 표준에 따라 그들의 자료를 생산하며, 그 자료에 대한 접근 방법 역시 다양하다. 무엇을 어떻게 디지털화할지는 원천 자료의 소장 기관이 결정하며, 유로피아나는 디지털화된 콘텐츠의 메타데이터만을 제공받는다.

하지만 유로피아나는 이 메타데이터의 형식을 설계함에 있어, 다양한 디지털 데이터가 의미적 연관을 가지고 서비스될 수 있도록 하였으며, 이를 기반으로 유럽 전역의 문화유산 디지털 콘텐츠를 포함하는 거대한 지식망을 구현하는데 성공하였다.



유로피아나와 데이터 제공 기관 사이의 메타데이터 수·발신 체계 (OAI-PMH 기반)

유로피아나의 데이터 모델 EDM은 다양한 문화유산 소장 기관이 유로피아나에 제공하는 데이터를 구조적으로 표현하는 방법이다. EDM은 시맨틱 웹 (Semantic Web)과 링크드 데이터(Linked Data)의 확산에 기여하고자 하는 취지에서 그것의 핵심 원칙을 준수하고 있으며, 이미 제정된 시맨틱 웹 개념의 여러 가지 표준의 기반¹⁴⁾ 위에서 만들어졌다.

유로피아나 문화유산 콘텐츠를 담은 데이터 구조는 3종의 핵심 클래스(Core Class)와¹⁵⁾ 4가지 유형의 문맥 정보 클래스(Contextual Classes)로¹⁶⁾ 구성된 다. 유로피아나가 단순히 아카이브 소장 물품의 정보 서비스에 머물지 않고, 그것들 사이의 지식망을 형성할 수 있게 하는 것은, 문화유산 콘텐츠 속에서 문맥 정보를 식별해 내고, 그것을 매개로 유관 지식의 연계를 이루어낼 수 있게 하기 때문이다.

V-3. 디지털 아카이브와 인문지식 네트워크

시맨틱 웹 시대의 디지털 아카이브는 실물 자료의 디지털 사본을 온라인상에서 제공하는 기능에 머물지 않는다. 유관한 정보들이 의미의 연결 고리를 그물처럼 맺고 있는 지식 네트워크 상에서 인물, 사건, 개념, 문헌 등 다른 유형의 지식과 결합하여 한 시대의 역사와 문화를 보여주는 문맥을 형성하게 하는 것이 현대에 추구되는 디지털 아카이브의 이상적인 모습이다. 하지만 이것은 아카이브 데이터베이스의 메타데이터만 확장한다고 되는 일이 아니다. 한 개의 베틀루가 500여 년의 왕실 역사와 관련이 있다고 해서, 500년 동안 일어난 모든 일을 그 베틀루의 속성으로 기록할 수는 없다. 500년 동안 일어났던 많은 사건과 그 시대를 살아온 인물들, 그에 관한 다양한 기록과 이야기들이 언제든지 필요할 때 그 실물의 정보와 결합될 수 있도록 하면 되는 것이다. 이것이 가능하기 위해서는 아카이브의 실물이 인문학 지식에 다가가는 길과 함께 인문학 지식의 연결 고리가 아카이브의 실물에 도달할 수 있는 장치를 마련해야 한다.

LIDO나 유로피아나가 추구하는 아카이브 실물에 관한 지식의 확장은 모두 링크드 데이터(Linked Data) 또는 시맨틱 웹(Semantic Web) 개념의 디지털 환경에서 이루어질 수 있는 일이다. 학계의 인문학 연구자들이 생산하는 인문 지식도 같은 환경에서 같은 방식으로 만들어진다면, 그 두 세계의 지식은 언제든지 필요할 때 의미 있고 흥미로운 문맥을 형성하며 소통하게 된다.

14) RDF(S), OAI-ORE, SKOS, DC(Dublin Core) 등

15) 3종의 EDM 핵심 클래스(Core Classes):

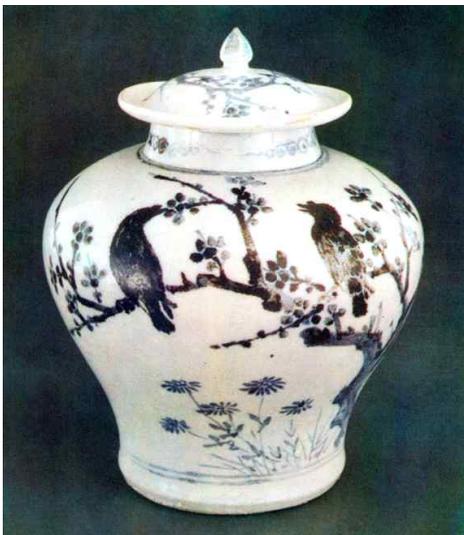
- . EDM:ProvidedCHO - 문화유산 그 자체에 대한 정보
- . EDM:WebResource - 문화유산에 관한 디지털 웹 자원의 정보
- . ORE:Aggregation - 문화유산 정보와 디지털 웹 자원의 집합체

16) 4종의 EDM 문맥 정보 클래스(Contextual Classes):

- . EDM:Agent - 개인, 기관, 단체 등 행위의 주체 LIDO의 actor에 상응하는 요소
- . EDM:Place - 장소
- . EDM:TimeSpan - 시간, 사건
- . SKOS:Concept - 개념, 용어

‘인문지식’이라고 추상적으로 표현하면 그 범위가 너무 넓어서 이러한 소통의 노력이 다소 허황하게 보일 수 있다. 그래서 그 범위를 ‘백과사전의 콘텐츠를 이루는 인문지식’ 정도로 축소해 놓고 보기로 한다. 한국의 역사, 문화, 사회에 관한 다양한 정보를 수록하고 있는 백과사전과 박물관이 소장하고 있는 우리나라 각 시대의 유물들의 관계를 생각해 보자. 국립중앙박물관의 도자 전시실에는 ‘청화(靑華)’라는 기법으로 만들어진 도자기가 있고, 『한국민족문화대백과사전』에는 ‘청화백자’라는 항목의 기사가 있다. 또한 그 기사의 말미에는 「청화백자의 안료에 대하여」나 「조선 전기 청화백자의 변천」과 같은 참고문헌이 소개되어 있다. 각각의 정보 사이에 연결 통로가 마련되면, 박물관의 전시 카드에 쓰인 서너 줄짜리 안내 정보가 백과사전의 개론적 지식으로, 필요한 경우 전문적인 학술연구 정보로까지 확장되어 갈 수 있게 된다.

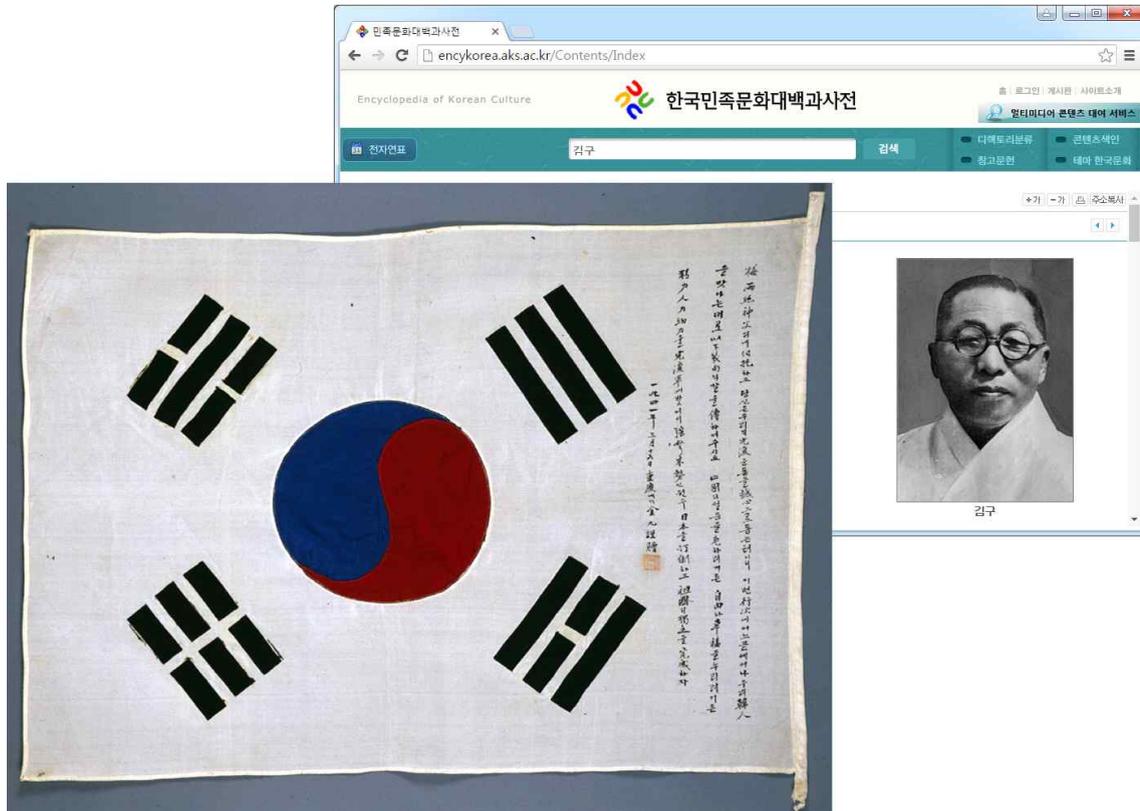
반대 방향으로 가는 연결에 대해서도 생각해 보자. 『한국민족문화대백과사전』에는 백범 김구 선생의 생애와 업적을 기술한 항목이 있다. 이 기사를 읽는 사람이 독립기념관이 소장하고 있는 ‘김구 서명문 태극기’를 보고, 거기에 쓰인 김구의 친필 문장¹⁷⁾을 읽는다면 열정적인 독립운동가 김구의 이미지를 더욱 선명하게 갖게 될 것이다.



The screenshot shows the website interface for the Korea National Encyclopedia. The article title is '청화백자'. The content includes a definition, a detailed history of the technique from the 14th to 15th centuries, and a list of related terms like '회화청' and '회화백자'.

백자 청화매조죽문 유개항아리. 국보 제170호. 국립중앙박물관 소장.

17) “.... 亡國의 서러움을 免하려거든, 自由와 幸福을 누리려거든, 精力, 人力, 物力을 光復軍에게 貸하여서 强弩末勢인 日本을 타도하고 祖國의 光復을 完成하자”



김구 서명문 태극기. 등록문화재 388호. 독립기념관 소장

V-4. 백과사전적 아카이브

각기 다른 곳에서 만들어진, 그렇지만 의미적으로 유관한 지식들을 함께 검색해서 찾아내는 것은 오늘날의 인터넷상에서 가능하다. 하지만 이것은 그 정보들 사이에 의미적 연결 고리(Semantic Link)가 있어서가 아니라, 텍스트 안에 같은 어휘가 쓰이고 있어서 발견되는 것이다. 인문정보학은 이러한 지식의 연계가 훨씬 더 정확하고 체계적으로 이루어지는 방법을 개발하고, 그것을 새로운 지식의 편찬에 적용하는 노력을 기울인다.

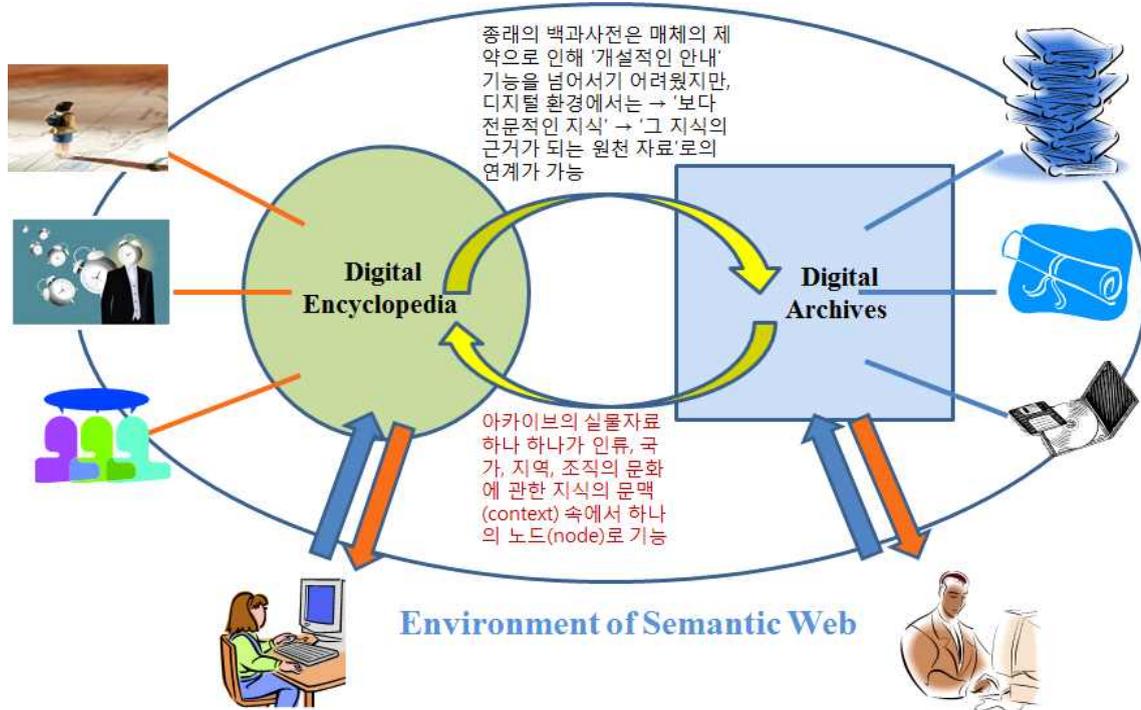
‘백과사전적 아카이브’(Encyclopedic Archives, Encyves)는 인문지식의 ‘일천 자료’이자 그 지식의 진실성을 입증하는 ‘증거’인 ‘실물 자료’(기록물, 유물 등) 데이터가 광대한 인문지식 네트워크의 노드(node)로 존재하는 세계이다.

아날로그 시대에 백과사전은 대중들이 분야별 전문지식의 세계로 들어가는 관문의 역할을 담당했다. 디지털 시대에는 인터넷상에 구현된 월드와이드웹이 종래의 ‘백과사전’의 역할을 대신한다. 종래의 백과사전은 매체의 제약으로 인

해 ‘개설적인 안내’ 기능을 넘어서기 어려웠지만, 디지털 환경에서는 → ‘보다 전문적인 지식’ → ‘그 지식의 근거가 되는 원천 자료’로의 연계가 가능하다. 시맨틱 웹 기술을 기반으로, 유관한 지식의 조각들이 다양한 방향으로 지식의 문맥을 형성할 수 있게 하는 것이 디지털 시대 백과사전의 새로운 모습이다. 이러한 디지털 백과사전은 그 네트워크의 연장선상에서 자연스럽게 모든 유관한 영역의 디지털 아카이브를 포섭한다.

도서관, 박물관 기록관 등 아카이브 기능을 수행하는 곳에서도 그들의 디지털 아카이브가 미래의 백과사전적 아카이브의 일부가 될 수 있도록 발전시켜야 한다. 과거에는 가치 있는 실물 자료를 수집하고 보존하는 것이 아카이브 주된 임무였다. 오늘날에는 그 실물 자료의 ‘활용성’을 증대시키는 일이 중요한 과제로 부상하였다.¹⁸⁾ 아카이브의 실물 자료가 독립적으로 존재하기보다 세상 사람들의 다양한 관심사에 긴밀하게 연계되어 있음을 밝히는 노력이 필요하다. 그 노력이란 아카이브의 실물 자료 하나하나가 인류, 국가, 지역, 조직의 문화에 관한 지식의 문맥(context) 속에서 하나의 노드(node)로 기능할 수 있도록 하는 것이다.

18) 도서관(library), 박물관(museum), 기록관(archives)과 같은 조직들은 실물을 소장하고 관리하는 일을 하기 때문에, 일찍부터 그 일을 보조하는 수단으로 정보기술의 활용을 모색해 왔다. 초기에는 자료의 수집, 정리, 전시, 대출 등의 업무 프로세스를 자동화하는 것을 위주로 하였지만, 정보통신 기술 환경의 급진적인 진화와 함께 이곳의 정보과학도 변화의 국면을 맞이하게 되었다. 소장하고 있는 실물의 운용뿐 아니라 그 실물의 의미와 가치를 알리는 ‘지식’을 함께 다루어야 하는 과제가 대두된 것이다.



백과사전적 아카이브(Encyclopedic Archives, Encyves):
디지털 백과사전과 디지털 아카이브의 융합 모델

백과사전적 아카이브는 어느 한 곳의 박물관이나 기록관을 그런 형태로 만들자는 제안이 아니라, 백과사전적 지식정보 데이터베이스를 중개자로 삼아 수많은 디지털 아카이브의 실물 정보들이 의미적 관계를 좇아 서로 연결될 수 있게 하자는 것이다. 그 네트워크의 연장은 아주 전문적인 학술 지식에 도달하기도 하고, 매우 대중적인 오락거리 상식과 접촉할 수도 있다.

백과사전적 아카이브의 구현은 학계와 아카이브계 양쪽에서 추구되어야 할 과제이지만, 그 실현의 관건이 되는 첫번째 과제는 디지털 지식 네트워크 상에서 운용할 수 있는, 새로운 형태의 '디지털 인문지식 백과사전'을 편찬하는 것이다.

오늘날 세계의 어느 유명 출판사도 더 이상 종이로 된 백과사전을 간행하지 않는다. 과거에 백과사전이 누렸던 지위를 이제는 인터넷이 차지하고 있다. 그렇다고 해서, 지난날 인문학자들에게 가장 큰 규모의 지식 사업으로 여겨졌던 백과사전 편찬이 디지털 세계에서 사라지게 되는 것은 아니다. 아카이브의 실물 정보가 그 활용성 증대를 위해 확장된 지식으로 나아가려 하고 있고, 학계의 전문적인 학술연구가 대중의 지적 호기심에 닿을 수 있는 실마리를 찾으려 하고 있다. 한층 넓어진 지식 유통의 세계에서는 수요와 공급이 적정하게 만날 수 있도록 하는 지식의 매개자 역할이 더욱 중요하다. 과거의 백과사전

콘텐츠를 그대로 디지털 미디어에 담아내는 방식으로는 그러한 역할을 기대할 수 없다. 디지털 시대의 백과사전은 시맨틱 웹의 철학을 좇아 새로운 체제로 편성되어야 한다.

디지털 백과사전 편찬이라는 새로운 과제를 두고 디지털 아카이브의 영역에 주목하는 이유는 두 가지로 이야기할 수 있다.

첫째, 백과사전 편찬에 응용할 수 있는 인문지식 온톨로지의 선도적 모델을 그곳에서 찾을 수 있다. 유로피아나, LIDO, CIDOC-CRM 등에서 조직적으로 체계화 하려는 문화유산 관련 지식은 학계에서 탐구하는 인문지식과 다른 별개의 것이 아니다. 학계의 학술연구는 주로 개인 연구자 차원에서 이루어졌기 때문에 지식의 공익적 확산을 위한 노력이 상대적으로 미흡했던 반면, 아카이브 분야의 세계적인 기관들은 공익적 봉사가 그들의 존재 이유이기 때문에, 그리고 그것의 촉진을 위한 자원 조달이 가능했기 때문에 선도적인 모델의 제시가 가능했다고 할 수 있다.

둘째, 디지털 정보로 편성된 인문지식 백과사전이 만들어졌을 때, 기술적인 면에서 그것과 바로 소통할 수 있는 지식의 수요처가 디지털 아카이브의 세계이다. 미래의 디지털 백과사전은 학술 연구, 학교 교육, 사회 교육, 지식문화 산업 등 인문학 지식을 필요로 하는 모든 영역에서 지식의 중재자 역할을 하게 될 것이다. 그 중에서도 지식의 중재 기능이 자동적으로, 다시 말해 사람의 개입을 최소화하고 디지털 정보시스템 안에서 지능적으로 이루어질 것을 기대할 수 있는 영역이 디지털 아카이브이다. 아카이브의 실물 정보는 인문지식과의 연계를 통해 그 큰 가치를 드러낼 것이고, 그렇기 때문에 디지털 세계에서 지식 확장을 위한 노력이 그곳에서도 지속적으로 경주될 것이기 때문이다.