



이 저작물의 인용 표시 방법: 김현, 「유교문화 디지털 아카이브를 위한 디지털 큐레이션 모델 연구」, 세계유교문화박물관 디지털 아카이브 마스터플랜 연구, 2020. 3. 국학진흥원.

유교문화 디지털 아카이브를 위한 디지털 큐레이션 모델 연구

김 현¹⁾

1. ‘살아있는 이야기’를 전달하는 아카이브
 2. 디지털 큐레이션
 3. 백과사전적 아카이브
 4. 메타 아카이브
 5. 디지털 큐레이션을 위한 데이터 모델 설계
 6. 세계유교문화 박물관 디지털 아카이브의 디지털 큐레이션
- 《부록 1》 한국문화유산 엔사이브 온톨로지
《부록 2》 개별 문화유산의 시맨틱 데이터 및 확장 네트워크 예시

1. ‘살아있는 이야기’를 전달하는 아카이브

한국의 옛 모습을 보여주는 수많은 문화유산과 역사 기록은 이 나라가 오랫동안 유교의 가르침에 기반을 둔 유교 국가였음을 알게 한다. 그럼에도 불구하고 오늘날 한국인들 가운데 유교를 자신이 믿는 ‘종교’라고 생각하는 사람은 거의 없다고 해도 과언이 아니다. 신분 제도가 사라지고, 사회적 성취의 기준이 달라진 상황에서 유교는 ‘최고의 가르침’으로서의 지위를 잃을 수밖에 없었고, 외형적으로 유교를 신앙하는 모습은 더 이상 한국 사회에 남지 않게 되었다.

하지만, 오늘날에도 한국인들의 가정생활과 학교교육의 곳곳에서 이들이 전통적인 규범에 영향 받는 모습을 발견하기는 어렵지 않다. 가족, 친지, 사제, 직장 상사·동료 사이의 관계에서 요구되는 예의와 염치는 다분히 유교적 전통에 뿌리를 두는 것이다. 조상에 대한 제례는 한국의 많은 가정에서 여전히 중요한 의미를 갖는다. 바람직한 부모의 모습, 교사의 자세, 심지어는 정치인의 인격까지도 유교적 인간관이 평가의 기준이 되는 경우가 많다. 유교는 현대 한국 사회에서 종교로서보다는 문화적 전통이자 국민들 대다수가 의식적 또는 무의식적으로 따르는 삶의 규범으로 그 영향력을 유지하고 있다.²⁾

1) 한국학중앙연구원 한국학대학원 인문정보학 교수, 디지털인문학연구소장. 한국디지털 인문학협의회 회장

2) 김현, 「한국의 유교」, 『100가지 이야기로 배우는 한국의 역사와 문화』,
http://dh.aks.ac.kr/Korea100/wiki/index.php/한국의_유교

우리가 박물관이나 아카이브를 만드는 이유는 단순히 잊힌 역사의 조각을 흥밋거리로 보여주고자 함이 아닐 것이다. 역사는 오늘의 우리가 이 자리에 이 모습으로 있을 수 있게 한 배경과 이유를 설명한다. 유교문화박물관이나 유교문화 디지털 아카이브를 통해 사람들은 오늘의 한국 사회와 개개인의 삶의 깊은 곳에 여전히 살아있는 ‘유교적 전통’의 본질과 변화의 이야기를 찾을 수 있어야 한다. 박물관의 아카이브의 콘텐츠가 이미 죽은 것의 박제와 다름없다면 그러한 효과를 기대하기는 힘들다.

아카이브의 콘텐츠에 관해 필자가 주목하는 ‘죽은 것’과 ‘살아있는 것’의 차이는 ‘고립’과 ‘연결’이다. 한 편의 문서, 한 장의 사진, 한 채의 건물, 한 사람의 인물이 각각 고립된 정보로 우리에게 주어진다면, 그것은 단지 과거 사실의 조각, 조각일 뿐이다. 하지만 그 조각들이 서로에 대해 어떤 관계 속에서 만들어지고 존재하게 되었는지를 알게 하는 정보가 주어져서 그 개개의 사실들이 ‘이야기’로 엮어질 수 있다면, 그것은 바로 오늘날 우리가 이 사회를 살아가는 이야기와도 유사한 콘텐츠가 될 것이다. 그 이야기의 한 끝에서, (그것이 장소건, 인물건, 유물이건, 삶의 동질성에 대한 공감건....) 나의 현재의 관심사와 직결되는 연결고리가 발견된다면 그 과거의 이야기는 현재의 이야기의 연장선상에서 새로운 의미와 생명을 얻는 ‘살아있는 이야기’가 된다.



의성김씨 학봉종택의 한글 서간과 관련 지식 데이터

‘살아있는 이야기’를 추구하는 디지털 아카이브의 개념은 실물로 소장하고 있는 자료를 온라인상에서도 열람할 수 있게 한다는 종래의 정보화 개념과는 다른 이야기이다. 후자는 ‘아카이브 자료의 디지털화’(Digitalization of Archival Materials)라고 할 수 있겠지만, 이것을 바로 ‘디지털 아카이브’(Digital Archive)와 동일시할 수는 없다.

박물관의 유물 한 점이나 기록관의 문서 한 장은 그것이 독립적, 자기완결적으로 존재가치를 갖는 것이 아니라, 어떤 사건이나 사회 현상, 인물의 삶과의 관련 속에서, 또 다른 유물들과 함께 어울려 어떤 집합적 의미나 중요성을 갖는 경우가 대부분이다. 의미 있는 이야기를 하려면 다른 유물들에 대해서도 언급해야 하고, 유물이나 문서의 형태로 존재하지 않는 무형의 지식도 동원해야 한다. 디지털 아카이브가 나의 아카이브 소장물의 디지털 카탈로그에 머물게 되면, 이용자들이 그곳에서 발견할 이야기는 매우 협소하고 무미건조할 것이다.

실제로 박물관에서 이루어지는 기획 전시의 사례를 보도록 하자. 큐레이터가 특정 주제의 전시를 기획하면서, 이에 관한 풍부한 볼거리와 지식을 전달하려고 할 경우, 자기 박물관의 소장품만 가지고서는 그 목적을 이루기가 어렵다고 판단하고, 다른 박물관이 가지고 있는 유물들의 대여전시를 추진할 것이다. 아울러 전시 유물에 대한 설명만으로는 그 유물이 대변하는 역사나 문화의 문맥을 충분히 전달할 수 없기 때문에, 그 배경을 설명하는 텍스트와 지도, 사진을 보여 주는 안내판을 제작한다. 사실상 박물관의 관람자들은 스스로 박물관에 유물을 보러 간다고 생각할지 몰라도, 그들이 접하는 것은 큐레이터들이 박물관 유물을 소재로 엮어낸 ‘이야기’이고, 유물들은 그 ‘이야기’의 증거로서 기억되는 것이다. 소장품 위주의 상설전시라고 하더라도, 모든 전시는 큐레이션의 결과이고 그 큐레이션은 소장 유물뿐 아니라고, 그 유물의 역사적, 문화적 맥락을 이해하는 데 필요한 지식 정보를 동원해서 이루어지기 마련이다. 그런데, 유감스럽게도 이러한 기관의 디지털 아카이브라고 하는 곳에서는 그 큐레이션의 흔적을 찾아보기 어렵다. ‘살아있는 이야기’를 전달하는 디지털 아카이브란, 말하자면, ‘유물의 아카이브’가 아니라 ‘큐레이션의 아카이브’를 만들자는 이야기이다.

2. 디지털 큐레이션

영어의 ‘curation’은 ‘돌보다’(to watch over, attend)의 뜻을 가진 라틴어의 ‘curare’에서 온 말이다.³⁾ 오늘날 통용되는 좁은 의미에서의 큐레이션은 박물관, 갤러리 등에서 유물이나 작품을 선별하여 구입하는 것, 소장품을 정리, 관리, 대여하거나, 그것을 관람객이 볼 수 있도록 전시하는 일을 말한다. 이 경우, 그러한 일을 수행하는 전문직업인을 큐레이터(curator)라고 한다. 좀 더 넓게 해석하면 미술품뿐 아니라 다양한 유무형의 대상물을 일정한 원칙과 기준에 의해 선별하고 정리하여 체계를 갖춘 컬렉션으로 만드는 일을 모두 큐레이션이라고 할 수 있다. 어떤 기획 의도를 가지고 그에 부합하는 음원을 모아서 뮤직 앨범을 만든다거나, 어떤 주제에 관한 다양한 일화를 수집하여 이야기 자료집으로 꾸미는 일도 큐레이션으로 볼 수 있는 일이다. 이러한 식으로 생각하면 이 큐레이션이라는 일은 인문학 연구자가 한 편의 글을 쓰거나 일정 시간의 강연을 준비하는 일과도 크게 다르지 않다고 할 수 있다.⁴⁾

이성적인 이해나 감성적인 체험의 가치가 있는 대상을 발견했을 때, 그 발견을 홀로 즐기기보다는 다른 이들과 함께 공유하고 공감하는 것을 더 의미 있다고 생각하는 것은 인간의 모든 문화적 활동의 저변에 놓인 사고이다. 큐레이터의 큐레이션이나 인문학자들의 저술 활동은 모두 자신이 얻은 앎과 느낌을 공유하고 확산하고자 하는 노력이라는 점에서 유사하다. 아울러 그 활동들은 단편적인 사물이나 정보를 무차별적으로 나열하기보다는 그 요소들의 연결고리를 만들어 맥락이 있는 이야기를 만듦으로써 잘 이해하고 더 많이 공감하게 하려 하는 공통점이 있다. 다만 그 이야기의 구성 요소가 역사적인 유물이나 미술작품과 같은 유형의 실물 위주인 경우 그 이야기 구성의 행위를 큐레이션이라고 불렀고, 언어로 기술된 정보와 지식으로 이야기를 만드는 경우 ‘저술’이나 ‘강의’ 등의 이름으로 차별화 했을 뿐이다.

아날로그 시대에는 그 두 가지를 전혀 다른 일로 취급할 그만한 이유가 있었다. 박물관의 수장고와 전시실, 대학의 연구실과 강의실은 서로 다른 독립된 공간이었고, 그 각각의 공간에서 일하는 사람들은 서로 다른 일을 하는 전문가로 취급되었다. 지식의 성격과 내용은 유관한 것일지라도, 들여다보는 대상물이 다르면 연구방법이 달라지고, 그래서 그 일은 서로 다른 일이었다. 고객들도 그들의 분업 체계를 당연한 것으로 받아들였을 것이다. 유물을 보기 위해서는 박물관으로, 문서를 확인하기 위해서는 기록관으로, 그 증거들을 학술적으로 설명하는 지식을 얻기 위해서는 대학의 강

3) "curator", *Merriam-Webster*, <https://www.merriam-webster.com/dictionary/curator>

4) 학생들을 위한 강의를 준비하거나, 나의 관심사를 세상과 공유하기 위한 글을 쓸 때, 나는 오래 그것에 답을 다양한 화제를 떠올리고, 그 하나하나를 어떠한 순서로, 어떠한 논리로 엮어갈 것인가에 대해 고민한다. 이야기의 진행을 구체화하는 과정에서 먼저 생각했던 화제를 버리기도 하고 새로운 화제를 찾기도 한다. 추상적으로만 생각했을 뿐 구체적인 내용을 알지 못했던 화제에 대해서는 그것을 좀 더 정확하게 알기 위해 새로운 조사를 시작하기도 한다. 그 때문에 많은 시간을 쓰기도 하지만 때로는 그 과정에서 좀 더 흥미로운 새 이야깃거리를 발견하기도 한다. 화제와의 씨름을 계속하다 보면, 그 이야기의 전체적인 흐름이 내가 당초에 구상했던 것과 많이 달라져 가기도 한다. 나는 그 이야기가 나 자신과 나의 학생들, 독자들에게 좀 더 만족스러울 수 있도록 화제의 깊이와 정확도, 그 연결의 순서를 다듬는 일을 반복한다.

의실로 가야 했다.

하지만 그 모든 지적 감성적 경험의 세계에 ‘디지털’이라고 하는 새로운 변수가 등장하면서 상황은 달라지기 시작했다. 현대의 지식 고객들은 박물관의 유물이든, 아카이브의 문헌기록이든, 그것들을 가지고 역사를 탐구한 인문학자들의 논술이든, 이 모든 것들을 사이버 공간에서 디지털화된 콘텐츠로 접하는 데 점점 익숙해져 가고 있다. 이 세계에서는 박물관과 강의실이 분리된 공간일 필요가 없다. 유물의 형상과 색채를 보이는 이미지 자료와 그것의 제작기법이나 역사적 배경을 설명하는 연구논문이 다 같이 디지털 데이터로 존재하는 곳에서 두 세계를 넘나드는 행위는 한 번의 마우스 클릭이면 족하다.

디지털 인문학계에서 논의되는 바의 ‘현대 인문학의 관행에 대한 반성’ 가운데 하나는 우리가 그동안 아카이브와 대학의 연구실을 둘로 나누는 과오를 저질러 왔다는 것이다. 2009년 미국에서 발표된 디지털인문학 선언 2.0 은 전통적인 대학의 연구실, 도서관, 박물관이 모두 하나로 통섭하는 세계가 미래 인문학의 무대가 될 것으로 내다보고, 그곳에서 이루지는 융합적인 교육·연구 활동을 새로운 ‘큐레이션’으로 인식하였다. 이에 대한 이해를 돕기 위해 선언문의 일부를 인용한다.

확장된 학술 활동으로서의 큐레이션 curation as augmented scholarly practice

디지털 인문학자들은 **큐레이션**을 미래 인문학의 중심적이 특징으로 인식한다.
Digital Humanists recognize curation as a central feature of the future of the Humanities disciplines.

근대 대학이 큐레이션을 이차적 보조적 역할로 강등시키고 큐레이터를 박물관, 아카이브 및 도서관으로 추방하면서 학문에서 분리시켰던 반면, 디지털 인문학 혁명은 연구와 교육의 지형도를 근본적으로 재형성하도록 촉진한다.

Whereas the modern university segregated scholarship from curation, demoting the latter to a secondary, supportive role, and sending curators into exile within museums, archives, and libraries, the Digital Humanities revolution promotes a fundamental reshaping of the research and teaching landscape.

디지털 인문학은 **학자를 큐레이터로, 큐레이터를 학자로 재구성**하고, 이로 인한 가능성과 수요의 확장을 통해 학문 활동에 다시 활기를 북돋으며 박물관, 도서관 및 아카이브의 학술적 사명을 재개하려한다.

It recasts the scholar as curator and the curator as scholar, and, in so doing, sets out both to reinvigorate scholarly practice by means of an expanded set of possibilities and demands, and to renew the scholarly mission of museums, libraries, and archives.

대학 박물관은 적어도 대학 도서관만큼의 연구소 역할을 할 수 있어야 한다. 아카이브는 교육과 실천교육의 현장이 되어야 한다. 교실은 과거의 교육 자료와 더불어 자료 처리, 주석달기 및 배열 등의 과제가 학습과정에 통합된 실천적 참여의 장소가 되어야 한다.

A university museum worthy of its name must become at least as much a laboratory as, say, a university library. An archive must become a place of teaching and hands-on learning. The classroom must become a place of hands-on engagement with the material remains of the past where the tasks of processing, annotating, and sequencing are integral to process of learning.

큐레이션은 또한 건강한 겸손함을 가지고 있다. 그것은 모든 것에 통달하겠다는 가당찮은 고집을 부리지 않으며, 실체를 접할 수 있고 변종도 있을 수 있는 국지적 지식을 포용한다. 또한 큐레이션은 추상적 이론을 멀리하며, 형상화, 대상화할 수 있는 것들을 모아놓는 것을 선호한다.

Curation also has a healthy modesty: it does not insist on an ever more impossible mastery of the all; it embraces the tactility and mutability of local knowledge, and eschews disembodied Theory in favor of the nitty-gritty of imagescapes and objecthood.

큐레이션은 단어, 이미지, 소리뿐만 아니라 물체를 통해 논의하는 것이다.

Curation means making arguments through objects as well as words, images, and sounds.

이것은 비평적·서술적 과업을 공간 속에서 수행한다는 의미를 내포한다. 역사가 들은 그러한 과업을 언어로만 수행해 왔다. 그 일을 물리적이거나 가상적인 공간 안에서 하는 것은 본질적으로 다른 일이다.

It implies a spatialization of the sort of critical and narrative tasks that, while not unfamiliar to historians, are fundamentally different when carried out in space—physical, virtual, or both—rather than in language alone.

이는 연구대상이 되는 것들을 실제로 수집하고, 조합하고, 선별하고, 구조화하고, 해석하는 일에 종사하게 되는 것을 의미한다. 이 모든 일들은 우리가 전통적인 서술적 학문과 동등한 것으로 간주하는 큐레이션의 과업이다.

It means becoming engaged in collecting, assembling, sifting, structuring, and interpreting corpora. All of which is to say that we consider curation on a par with traditional narrative scholarship.

큐레이션은 자체적으로 명시적 언어, 기술 도구 및 복잡성을 가지고 있는 수단이다. 이 수단은 이 순간에도 변화하고 확장되어 가고 있으며, 그와 함께 가상현실에서 구현되는 갤러리, 학습 환경, 다양한 세계들이 학술의 영역에서도 중요한

특징으로 부각되고 있다.

It is a medium with its own distinctive language, skill sets, and complexities; a medium currently in a phase of transformation and expansion as virtual galleries, learning environments, and worlds become important features of the scholarly landscape.

.....

큐레이션은 강화된 학술연구 활동이다. 그것은 또한 가르치는 것과 배우는 것도 강도높게 증진시킨다.

Curation is an augmented scholarly practice that also powerfully augments teaching and learning.

큐레이션의 방법으로, 미래 세대의 인문학자들은 문화와 역사연구의 가장 기본적인 원자재를 가지고 연구를 시작할 것이며, 실제 실험실과도 같은 환경에서 전문적인 연구자들의 지도 하에 지식을 수집하고 생산하는 일에 직접 참여하게 될 것이다. 큐레이션 작업이 학술연구 “성과”의 범위에 추가됨으로써 인문학 연구 세계는 광대하게 풍요로워진다.

It summons future generations of humanists to set to work right from the start with the very stuff of culture and history: to become directly engaged in the gathering and production of knowledge under the guidance of expert researchers in a true laboratory-like setting. The universe of Humanities research is vastly enriched by the addition of curatorial work to the range of recognized and supported "outputs" for scholarship.

큐레이션은 학문의 선결조건을 새롭게 만들어 낸다. 그것은 자기들만의 고유한 전문 용어의 경계에서 벗어나서 보다 유연한 공공의 영역으로 나아가라는 것이다. 그 열린 영역에서는, 전통적인 형태의 학문이 학자의 전문적 지도하에 아카이브의 광범위한 참여자들을 위해 다양한 목적으로 활용될 수 있다.

Curation creates the preconditions for modes of scholarship that step outside the boundaries of one's own expert language into a more fluid public realm, where traditional forms of scholarship can be multipurposed for the large-scale participatory generation of archival repositories under the expert guidance of a scholar.⁵⁾

박물관의 큐레이터가 전시를 위해 유물 하나하나와 관련이 있는 다양한 지식 정보를 찾아내고, 의미의 연결고리를 찾아 그 유물의 문화적 가치를 설명하는 이야기를 만들 듯이, 미래 인문학의 연구자와 학생들은 디지털 세계에서 유의미한 지식의 조

5) The Digital Humanities Manifesto 2.0,
<http://manifesto.humanities.ucla.edu/2009/05/29/the-digital-humanities-manifesto-20/>

각을 탐색하고 그것들 사이의 관계성을 탐구하는 방법으로 새로운 사실을 발견하고 대상 세계에 대한 이해를 넓혀갈 것이다. 그곳에서 탐구해야 할 대상이 아카이브의 유물이나 책 속의 지식이냐를 가려서 어느 한 쪽만 어느 한 쪽에만 집중하고, 다른 쪽은 소홀히 해야 할 이유가 전혀 없다.

전통적인 인문학계가 아카이브와 큐레이션을 소홀히 했던 것처럼 전통적인 아카이브도 인문학적 지식의 확장에 적극적이지 않았다. 소장하고 있는 물품이나 기록물에 대한 해석을 유보한 채, 원천 자료를 그대로 정리·보존하여 제공하는, ‘연구 지원’ 역할을 수행하고, 그 자료에 대한 해석이나 의미부여는 대학의 연구실에서 ‘학술 연구’의 행위로 수행한다는 입장에 머물렀던 것이다. 이러한 사고의 연장선상에서 현대의 디지털 아카이브는 ‘해석되지 않은 원천 자료의 디지털 사본’의 형태에 고집스럽게 머물고 있는 듯한 양상이다. 문제는 아카이브 자료는 비록 이러한 고전적 형태로나마, 디지털 기술에 힘입어 이용자 접근성이 나날이 향상되고 있는데, 이 아카이브에서 배제된 문맥 이해와 해석의 데이터는 어디에서도 찾아보기가 어렵다는 것이다.

종래 아날로그 세계에서 행하던 독립적, 자기완결적 큐레이션을 디지털 미디어 상에서 하는 일을 가지고 ‘디지털 큐레이션’이라고 하는 이들도 있다. 그것은 큐레이션을 여전히 박물관 안에서의 일에 한정하는 좁은 사고에 기인하는 발상이다. 디지털 인문학의 입장에서 이해하는 ‘디지털 큐레이션’은 바로 아카이브와 인문학연구실의 경계를 넘어서는 융합적인 지식의 탐구이다. 디지털 환경에서 그와 같은 지식 탐구를 실천하게 되면, 그 탐구의 과정과 결과가 고스란히 디지털 콘텐츠로 남을 수 있게 되고, 이것은 또한 아카이브 컬렉션의 문화적 문맥을 드러내고 역사적 가치를 입증하는 온라인 전시 콘텐츠로서 활용될 수 있게 된다. 그것은 다시 개방적인 클라우드 환경에서 새로운 큐레이션을 유도할 것이고 유의미한 지식의 확장을 촉진할 것이다. 이것이 내가 생각하는 디지털 아카이브의 미래 모델이다.

3. 백과사전적 아카이브: 백과사전과 아카이브의 융합

필자는 2009년 한국학중앙연구원 한국학대학원에 인문정보학 전공 과정을 개설한 직후부터 아카이브의 컬렉션의 디지털 사본과 그 유물들의 사회적·문화적 문맥을 알려 주는 백과사전적 지식 정보를 긴밀하게 연계시키는 정보 시스템의 구체적인 프레임워크를 개발하는 일과 이 시스템을 구현하고 운영할 디지털 큐레이터를 양성하는 것을 이 전공의 중심적인 교육·연구 과제로 다루어 왔다. ‘백과사전적 아카이브(Encyclopedic Archives)’, ‘약칭 엔사이브(Encyves)’는 이 융합적 시스템에 부여한 이름이다.⁶⁾ 세계유교문화 디지털 아카이브의 프레임워크 개발에 참조가 될 수 있도록, 필자가 구상해 온 백과사전적 아카이브의 개념과 그것을 추구하는 이유, 그것이 지향하는 목표를 소개하겠다.

1) 디지털 아카이브와 인문지식의 네트워크

디지털 아카이브의 기능은 실물 자료의 디지털 사본을 온라인상에서 제공하는 것에 머물지 않는다. 유관한 정보들이 의미의 연결 고리를 그물처럼 맺고 있는 지식 네트워크상에서 인물, 사건, 개념, 문헌 등 다른 유형의 지식과 결합하여 한 시대의 역사와 문화를 보여주는 문맥을 형성하게 하는 것이 현대에 추구되는 디지털 아카이브의

6) 필자가 이 시스템을 처음 구상하게 된 것은 한국학중앙연구의 지식정보센터의 책임자로서 이 기관의 정보 콘텐츠 편찬·서비스 업무를 총괄할 때였다. (2007~2010) 그 때 부서에서 하던 일의 성격은 크게 두 가지로 나눌 수 있었다. 하나는 한국민족문화대백과사전, 한국향토문화전자대전과 같은 디지털 백과사전을 편찬하는 일이었고, 다른 하나는 장서각이 소장하고 있는 고전적의 디지털 아카이브를 만드는 일이었다. 둘 다 디지털 콘텐츠를 제작하고 데이터베이스를 구축하는 일로서 기술적으로 유사한 성격의 일이었지만, 각각의 일은 별도의 조직에 의해 수행되었고, 그 두 팀의 업무는 어떠한 연결고리도 없이 독립적으로 수행되었다. 일부터 그렇게 한 것은 아니지만, 책자 형태의 저작물을 만들어 내던 시절부터 그 두 가지 일은 별도의 조직에서 다른 연구진들에 의해 수행되어 온 일이었기 때문에 디지털 사업을 시작할 때 예전의 체계를 그대로 답습한 것이었다. 사실상 두 사업의 총괄책임자였던 나 이외는 그 조직의 어느 누구도 다른 팀의 일에 대해 잘 알지 못했고, 또 알려고도 하지 않는 듯했다. 나 역시 처음에는 그 두 가지 일이 기술적 프레임워크 면에서 유사할 뿐 별개의 일이라고 생각했다. 그러나 시간이 흐르면서 콘텐츠의 내용에 대한 나의 이해가 깊어질수록 그 두 가지 일은 나의 눈에 별개의 일로 보이지 않게 되었고, 오히려 그 두 영역 사이에 소통과 연계가 없는 것이 더 이상한 일로 생각되었다. 말하자면, 어느 동일한 문헌자료 한 권을 두고 볼 때, 그 책에 대한 해제와 관련 인물에 대한 정보가 민족문화대백과사전에 있고, 그 책의 원문 이미지 데이터는 장서각 디지털 아카이브에 있는데, 그 관련 있는 두 가지 콘텐츠를 서로 연결되지 못하고 있었던 것이다. 한 기관의 한 부서 내에서 이런 문제가 인식되었으면, 당연히 개선해야 된다고 생각하고 두 가지 콘텐츠를 하나의 시스템으로 엮는 문제를 부서원들과 의논하기 시작했다. 그 일은 논의의 시작부터 난관에 봉착했다. 그 때, 내가 대면한 문제점은 이런 것이었다. 첫째, 각 팀의 기존 실무자들은 다른 팀의 일에 무관심했을 뿐 아니라, 상대방의 일이 자기 일에 영향을 주면 기존의 안정화된 업무 체계가 무너질 것을 과도하게 염려했다. 둘째, 양쪽의 데이터를 문맥이 통하게 연계시켜서 종합적인 지식 콘텐츠를 만들려면 기존에 해 오던 사전편찬이나 영인물 간행의 수준을 넘어서서 훨씬 더 치밀하고 깊이 있는 조사연구가 필요한데, 그 일을 위한 전문 인력의 충원이나 조직 개편이 용이하지 않았다. 셋째, 디지털 백과사전 편찬과 디지털 아카이브 구축 사이 연계 시스템에 대한 구체적인 설계도나 참조 모델이 없었다. 나의 머릿속에만 있는 구상을 가지고, 디지털 세계에 대한 이해가 부족한 부서원들의 이해와 동의를 얻는 것은 가능하지 않았다. 이 일을 경험 후, 나는 백과사전적 아카이브 구축을 위한 실험의 무대를 대학원의 인문정보학 랩으로 옮겨서, 구체적인 데이터 모델 개발과 함께 그 일을 수행할 수 있는 인력의 양성에 주력했다.

이상적인 모습이다. 하지만 이것은 아카이브 데이터베이스의 메타데이터만 확장한다고 되는 일이 아니다. 한 개의 벚루가 500여 년의 왕실 역사와 관련이 있다고 해서, 500년 동안 일어난 모든 일을 그 벚루의 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.

LIDO⁷⁾나 유로피아나⁸⁾가 추구하는 아카이브 실물에 관한 지식의 확장은 모두 시맨틱 웹(Semantic Web)⁹⁾ 또는 링크드 데이터(Linked Data)¹⁰⁾ 개념의 디지털 환경에서 이루어질 수 있는 일이다. 학계의 인문학 연구자들이 생산하는 인문지식도 같은 환경에서 같은 방식으로 만들어진다면, 그 두 세계의 지식은 언제든지 필요할 때 의미 있고 흥미로운 문맥을 형성하며 소통하게 된다.¹¹⁾

7) LIDO(Lightweight Information Describing Objects)는 국제박물관협의회(ICOM, International Council of Museums)에서 제정한, 문화유산 정보의 기술을 위한 XML 스키마이다. 2010년 국제박물관협의회 국제문서화위원회(ICOM-CIDOC, International Committee for Documentation) 컨퍼런스에서 LIDO v1.0이 공표되었다. LIDO는 박물관의 유물 등 문화유산에 관한 정보를 명시적으로 전달하기 위한 목적으로 고안되었다. 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 설명 정보(Descriptive Information)를 담을 수 있는 틀을 제공한다. LIDO의 가장 중요한 부분은 CIDOC CRM에서 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 CRM: Conceptual Reference Model. 국제박물관협의회 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)

8) 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. 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 연관성을 가지고 서비스될 수 있도록 하였으며, 이를 기반으로 유럽 전역의 문화유산 디지털 콘텐츠를 포함하는 거대한 지식망을 구현해 가고 있다.

9) 시맨틱 웹이란 컴퓨터가 인식할 수 있도록 명확한 의미를 부가한 데이터의 웹이다. 인터넷상에 존재하는 여러 가지 자원들이 각각 별개의 것으로 보이기보다, 서로에 대해 어떠한 의미적 관계를 갖는지를 알게 하는 것을 지향한다. '시맨틱 웹'은 어떤 기술이나 표준을 특정하는 말이기보다, 미래의 '월드와이드웹'(World Wide Web)에 대한 기대와 이상을 표현하는 용어이다. 하지만 그 이상을 실현하기 위해서는 기술을 개발하고 표준을 정하는 노력이 있어야 한다. 시맨틱 웹은 월드와이드웹의 창시자 팀 버너스리가 1998년 제안한 이후, 월드와이드웹 컨소시엄(World Wide Web Consortium: W3C)을 중심으로 그것의 구현을 위한 기술의 확장과 표준화 작업이 진행 중이다.

10) 시맨틱 웹이 데이터의 연결망으로서의 미래의 웹에 대한 포괄적인 구상이라고 한다면, LOD는 HTTP URI(Uniform Resource Identifier)와 RDF(Resource Describing Framework) 등의 데이터 기술 형식에 의존하여, 일정한 영역에서 데이터의 연결을 구현하는 구체적인 방법의 제안이라고 할 수 있다. 팀 버너스리는 시맨틱 웹에 다가가기 위한 실천 전략으로 'Linked Data'와 'Linked Open Data'의 개념을 소개하였다. Tim Berners-Lee, "Linked Data", W3C Design Issues (2006, 2009), <https://www.w3.org/DesignIssues/LinkedData.html>

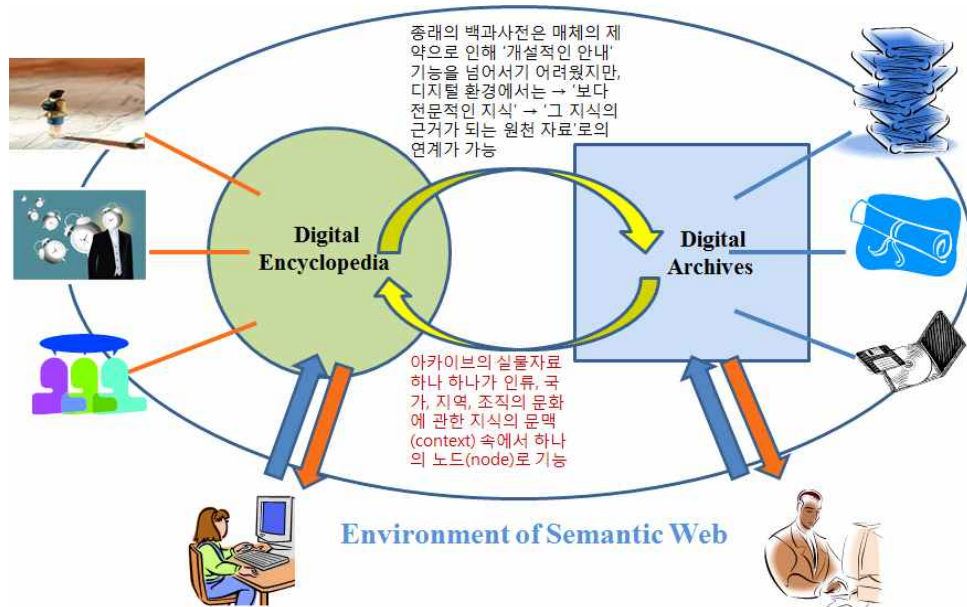
11) 김현, '디지털 아카이브와 인문지식 네트워크', 『디지털인문학 입문』 (제2쇄본, 1028) p. 206

2) 디지털 백과사전과 디지털 아카이브

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 인해 ‘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 발전시켜 가야 한다. 과거에는 가치 있는 실물 자료를 수집하고 보존하는 것이 아카이브의 주된 임무였다. 오늘날에는 그 실물 자료의 ‘활용성’을 증대시키는 일이 중요한 과제로 부상하였다.¹²⁾ 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.

12) 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: 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 첫 번째 과제는 디지털 지식 네트워크상에서 운용할 수 있는, 새로운 형태의 '데이터 기반 인문지식 백과사전'을 편찬하는 것이다.¹³⁾

3) 백과사전적 아카이브 구축 사례: 한국문화 엔사이브

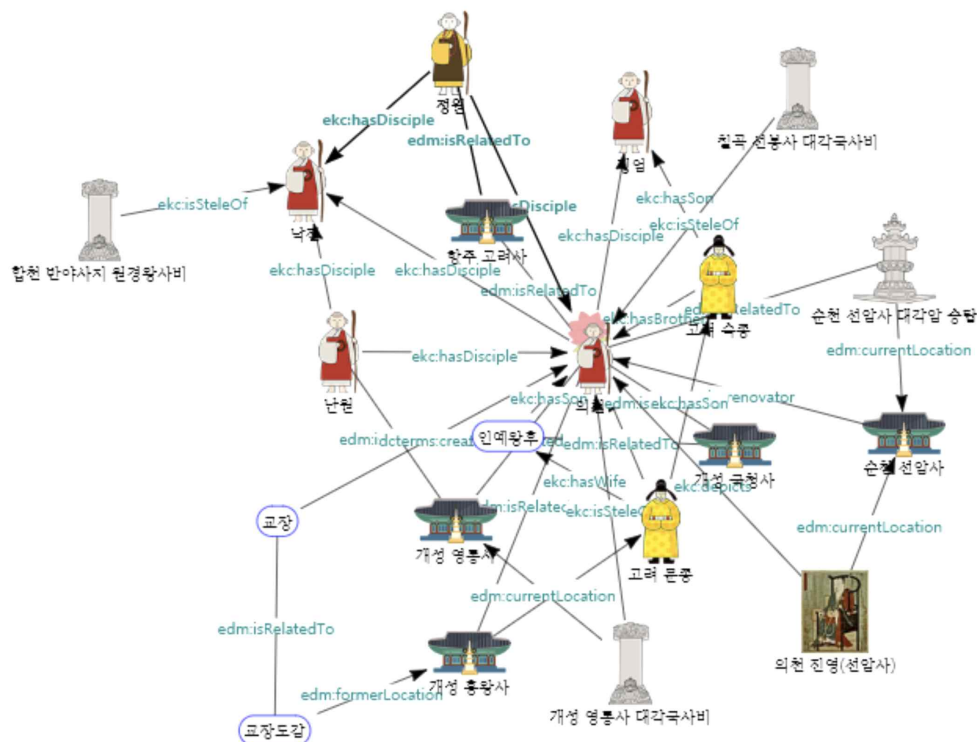
'한국문화 엔사이브(Encyves of Korean Culture)'는 데이터 기반 인문지식 백과사전의 모델을 제시하려는 목적으로 가지고 편찬된 주제 중심 디지털 백과사전이다.¹⁴⁾ 한국의 대표적인 기록문화 유산 중 '한글고문헌', '궁중기록화', '민족기록화', '불교 사찰의 승탑 비문', '옛 인물의 초상화' 등 다섯 가지 영역을 연구 대상으로 선정

13) 김현, '백과사전적 아카이브', 같은 책 pp. 208~210.

14) 이 디지털 백과사전은 한국학중앙연구원이 지원한 '한국 기록유산의 디지털 스토리텔링 자원 개발'(2016~2017) 및 '문화유산 속의 인물에 관한 시각적 스토리텔링 자원 개발'(2017) 연구의 일환으로, 한국학중앙연구원 인문정보학교실 연구진들에 의해 편찬되었다. 오늘날 디지털 미디어 상에서 서비스되는 백과사전류의 콘텐츠가 많이 있지만, 그것은 대부분 '책'이라는 아날로그 미디어에 적합한 형태로 편찬된 콘텐츠를 디지털로 서비스하는 것'에 지나지 않는다. 이것은 디지털 시대의 지식수요에 대한 적절한 대응이 아니라는 인식에서 새로운 디지털 백과사전 모델 연구에 착수하였고, '한국문화 엔사이브'를 시범적인 연구 편찬의 성과로 제시하였다.

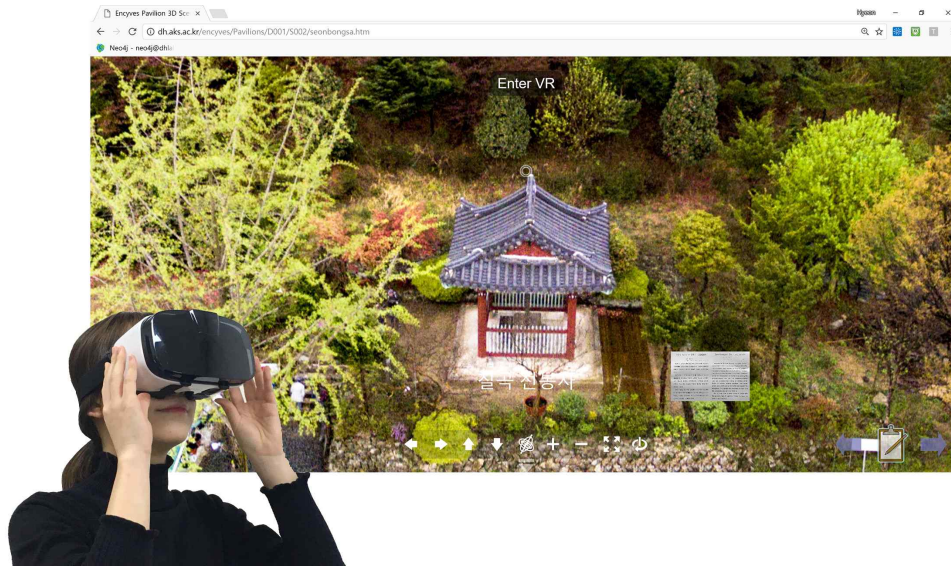
<http://dh.aks.ac.kr/Encyves/wiki>

하고, 그 범주에 속하는 대표적인 기록물과 관련 지식 정보를 디지털 콘텐츠로 제작하여, 다양한 수준의 한국학 교육에 활용할 수 있게 하였다.



『한국문화 엔사이브』의 기사별 네트워크 그래프 예시¹⁵⁾

15) <http://dh.aks.ac.kr/Encyves/wiki/index.php/>의천



『한국문화 엔사이브』의 기사별 가상현실 예시¹⁶⁾

이 백과사전 콘텐츠에는 데이터 네트워크, 전자지도, 연자연표 가상현실 등 다양한 소프트웨어 기술이 적용되었지만, 이용자가 1차로 접하는 기본 플랫폼은 독자가 곧 저자가 될 수 있는 개방형 공동 데이터 편집기 ‘위키’ 소프트웨어이다. 위키를 기본 플랫폼으로 채택한 것은 향후 일반 독자들도 이 백과사전의 편찬자로 참여할 수 있는 길을 열어 주기 위해서이다. ³¹⁾

4) 데이터 기반 인문지식 백과사전의 요건

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

아날로그의 세계에서 백과사전과 아카이브는 서로 독립적으로 존재했었다. 하지만 이제 그것이 모두 디지털 세계에 있고, 데이터의 형태로 존재한다면, 그 두 가지는 더 이상 별개의 것이어야 할 필요가 없다. 한 번의 클릭으로 아카이브의 실물에 다가갈 수 있는 데이터 시대의 백과사전은 기존의 백과사전과 무엇이 다른가? 데이터 기반 인문 지식 백과사전의 특징을 다음과 같이 정리해 볼 수 있다.

① 데이터 네트워크 형태의 백과사전

다양한 주제의 정보가 독립적인 ‘항목’으로 나열되어 있는 것이 종래의 백과사전이

16) http://dh.aks.ac.kr/Encyves/wiki/index.php/칠곡_선봉사

31) 김현, ‘한국문화 엔사이브’, 같은 책 pp. 39~41.

었다고 한다면, 데이터 기반 백과사전은 항목과 항목, 또는 개별 항목 속의 지식 요소들이 서로 어떠한 의미의 연관이 있는지 그 ‘관계성’을 보이는 데 주력한다. 데이터 기반 백과사전은 ‘가나다’ 순으로 정리된 목록 형태가 아니라, 방대한 규모의 데이터 네트워크의 형태로 편찬된다. 그 네트워크의 노드 하나하나가 백과사전의 항목이고, 노드와 노드 사이를 연결하는 연결선은 항목과 항목 사이의 다양한 연관관계를 대변한다.

② ‘지식’과 ‘자료’를 넘나드는 융합 콘텐츠

데이터 기반 백과사전 속에서는 이론화된 지식과 그 지식의 증거 자료가 단절 없이 연결될 수 있다. 아날로그 세계에서는 ‘연구실’과 ‘아카이브’가 서로 분리된 영역에 속했지만, 한 번의 클릭으로 그 두 세계를 넘나들 수 있는 디지털 환경에서는 ‘지식’과 ‘자료’가 훨씬 가깝게 묶일 수 있기 때문이다. 데이터 기반 백과사전은 서로 유관한 ‘지식’과 ‘자료’를 함께 탐구할 수 있는 길을 제공한다.

③ 가상현실에서 지식을 탐구하는 시각적 체험 공간

새로운 백과사전의 데이터는 텍스트에 한정되지 않는 멀티미디어 콘텐츠이다. 역사와 문화의 자취를 담은 지리적 공간을 3차원 가상현실의 속에서 체험하면서, 그 구석구석에서 호기심에 답하는 백과사전적 지식을 열어 볼 수 있는 유저 인터페이스를 제공한다.

④ 집단 지성의 기여로 확장되는 개방적 데이터

데이터 기반 백과사전의 지식 관계 네트워크는 학자들이 생산한 전문 지식뿐 아니라, 다양한 층위의 ‘유용한 지식’에 접근할 수 있는 길을 열어준다. ‘위키피디아’처럼 독자가 직접 백과사전 기사(네트워크의 한 노드)를 수정, 증보할 수 있을 뿐 아니라, 노드 사이의 링크 데이터를 추가하여 지식 관계 네트워크를 확장해 갈 수 있다. 이 방법으로 사회과학이나 산업 분야의 빅데이터에 대한 연결고리가 만들어질 수 있으며, 4차산업혁명을 이끄는 현대 사회의 빅데이터 속에서 고전 인문지식이 유효한 역할을 할 수 있게 한다.³²⁾

32) 김현, ‘데이터 기반 인문지식 백과사전: 기존의 백과사전과 무엇이 다른가?’,
http://dh.aks.ac.kr/Encyves/wiki/index.php/데이터_기반_인문지식_백과사전

4. 메타 아카이브

메타 아카이브(Meta-Archive)의 어휘 그대로의 뜻은 ‘아카이브들의 아카이브’라는 풀이가 될 것이다. 어느 특정한 범주 안에서, 여러 곳의 실물 아카이브들이 각각 어떠한 유물을 소장하고 있으며, 각 유물에 대해서 어떤 정보를 제공하는지 알게 하는 데이터베이스가 있다면, 이것도 일종의 메타 아카이브로 간주할 수 있다. 우리나라의 경우, 전국의 국공립 박물관의 소장 유물의 목록을 통합하여 서비스하고 있는 e-뮤지엄도 그것에 해당한다고 할 수 있다. 하지만, 필자가 지금 이 지면을 통해 소개하고자 하는 ‘메타 아카이브’는 그러한 수준의 ‘목록 데이터 통합 검색 서비스’와는 다른 것이다. 필자가 ‘메타 아카이브’의 이름으로 세계유교문화 디지털 아카이브에 제안하려는 일은 개별 아카이브나 독립적인 저작물의 데이터 사이에서 소통의 큐레이션을 하는 것이다.³³⁾

아날로그 시대의 실물 아카이브는 ‘물리적인 사물’의 아카이브였기 때문에 그 실물을 활용하는 활동도 위치한 장소의 공간적 범위 안에서 수행되었다. 어느 박물관이 소장하고 있는 그림을 보려면 당연히 그곳을 찾아야 했고, 사람들은 그것을 당연하게 여겼다. 조선시대의 대표적인 화가 겸재 정선(鄭敼, 1676~ 1759)의 그림을 보고자 할 때, 가장 널리 알려진 금강전도(1734)나 인왕제색도를 보기 위해서는 삼성미술관 리움으로 가야 하고, 그가 만년에 금강산 여러 곳의 풍경을 그린 해악전신첩(1747)을 보려면 간송박물관으로, 젊은 시절의 금강산 그림 풍악도첩(1711)을 보려면 국립중앙박물관으로 가고, 2006년 독일에서 돌아온 그의 화첩을 보려면 왜관의 성베테딕토 수도원을 방문해야 한다. 실물 작품을 보고자 할 때는 그 작품을 소장하고 있는 박물관을 하나하나 찾아가는 것이 당연하다. 하지만 그 유명한 정선의 그림을 실물로 본 사람이 지금 이 글을 읽는 독자 가운데 얼마나 되겠는가? 예전에는 책에서 보았을 테고, 지금 그 그림을 보거나 그 그림에 대해서 알고 싶다면 인터넷을 검색할 것이다.

정선의 그림에 관한 포털의 검색 결과는 무수히 많지만 앞서 언급한 그 그림들을 한꺼번에 보여주는 곳은 어디에도 없다. 그것을 모아서 일목요연하게 정리하는 것은 사용자가 스스로 손품을 팔아가며 해야 할 일이다. 뿐만 아니다. 정선이 어떠한 사람인지, 그가 그린 금강산의 명소들은 어느 곳이며 또 지금 실제의 모습은 어떠한지, 그의 절친이었고 큰 영향을 주었다는 시인 이병연(李秉淵)은 어떤 시를 남겼는지, 정선이 현령을 지냈다는 양천현은 오늘날 어느 곳인지, 화가 정선에 관해 조금이라도 깊이 알아보고자 할 때, 나의 궁금증에 답할 수 있는 정보를 찾는 일 또한 쉽지가 않다. 인터넷상의 다양한 정보들을 종합적으로 제공하는 포털 사이트가 있고, 또 정부 재원으로 만들어진 갖가지 통합검색 시스템들이 존재하는데, 왜 그것들은 나의 호기심이 구하는 지식을 바로 제공해 주지 못하는 것일까?

33) 그러한 의미에서 보면, ‘메타 아카이브’의 개념은 앞에서 언급한 ‘백과사전적 아카이브’와 동일하다고 할 수 있다. 다만, 아카이브 데이터의 원천 소스가 어느 특정 아카이브에 한정되지 않고, 일정한 주제로 묶을 수 있는 다수의 아카이브에서 취득된다고 하는 점이 강조되는 것이다.

우리가 오늘날의 인터넷 상에서 검색을 통해 찾을 수 있는 것은 단편적인 지식의 조각들일 뿐이다. 그 조각들 사이의 어떤 의미의 연결고리가 있어서 어떤 이야기로 엮여질지에 대한 정보는 사실상 존재하지 않는다. 다시 말해 인터넷 상의 데이터는 ‘큐레이션’이 부재(不在)한 원시 데이터(raw data)일 뿐이다.³⁴⁾

우리는 오랫동안 인터넷이 많은 정보를 한꺼번에 찾을 수 있게 해 준다는 사실에 놀라고, 고마와해 왔지만, 그 모여진 정보를 좀 더 가치 있고 유용하게 하는 일에는 거의 무관심했다고 해도 과언이 아니다. 좀 더 정확하게 말하면, 그 일은 정보 기술자나 포털 사업체가 할 일이지, 인문학자와 큐레이터가 관심 가질 일이 아니었다. 곳곳에 뿔뿔이 흩어져 있는 우리나라 대표적인 전통화가의 그림을 체계적으로 묶어서 볼 수 있게 하는 일은 미술사의 지식을 가진 연구자나 큐레이터의 일이 아니고, 포털 사업체의 프로그래머의 일이라고 생각해 온 것이다.

기관과 조직, 전공분야의 벽에 의해 고립되고 흩어진 데이터 사이에서 소통의 큐레이션을 하는 것은 구글이나 네이버가 해 줄 수 있는 일이 아니다. 특정한 주제 분야에서 깊이 있는 지식을 가지고 있는, 그래서 이 박물관의 유물과 저 도서관의 기록물이 서로 어떤 연관이 있는지 찾아내고 설명할 수 있는 전문 큐레이터의 역할이 필요한 일이다. 우리 사회에서 ‘디지털 아카이브’의 올바른 위상이 정립되기 위해서는, 개별 아카이브 안에서만 큐레이터의 역할이 있고, 인터넷 상에서 여러 아카이브를 집성하는 일은 단지 기술적인 작업일 뿐이라는 사고가 하루라도 빨리 불식되어야 한다.

메타 아카이브는 기본적으로 이 세상에 서로 유관한 여러 곳의 실물 아카이브가 존재한다는 전제 위에서 성립하는 개념이다. 그 유관성은 관점에 따라 여러 가지로 상정될 수 있다. 주제, 지역, 시대와 사건, 인물, 작품의 유형 등..... 지금 우리가 설계하고자 하는 ‘세계 유교 문화 박물관 디지털 아카이브’를 이 관점에서 생각해 보자.

실물의 디지털 사본을 실물 수집하듯이 모아들이고 그것으로 박물관 전시실을 흉내 낸 가상의 전시실을 꾸미는 것은 진정성 없는 실물 아카이브의 디지털 아류가 되기 쉽다. 이 디지털 아카이브는 기존이 실물 아카이브가 하지 못한 일을 해야 한다. 기

34) 20년 전 대한민국 정부가 이른바, “자식정보자원관리사업(2000~2009)” 등의 대형 국책 사업으로 공공 데이터의 디지털화 사업을 전개했을 때만해도, 그 사업의 기획자들은 아날로그 데이터를 디지털로 전환하는 것만으로도 다양한 데이터의 연계가 가능해지고, 새로운 지식과 이야기의 생산이 폭발적으로 증대될 것으로 기대했었다. 말하자면, 승정원일기와 민족문화대백과사전 국공립박물관의 소장품목록이 다 같이 인터넷 상에 존재하면 그 안에 있는 유관한 정보는 서로 연결이 되어서 커다란 지식의 네트워크를 만들어 줄 것으로 생각한 것이다. 그런 일은 일어나지 않았다.

당시 이 사업의 주관 부처였던 정보통신부나 실행기관인 한국전산원(현 한국정보화진흥원), 그리고 그들에게 자문하였던 정보기술 전문가들은 ‘지식의 큐레이션’에 대한 이해가 터무니없이 부족했다. 자식정보자원관리사업으로 생산되는 모든 데이터에 확실적인 메타데이터 몇 항목을 부가하게 함으로써 데이터의 유용한 지식의 연계가 이루어질 수 있을 것으로 기대했던 것이다.

이 사업은 막연히 기대했던 ‘지식의 연계’는 성공적으로 이루어내지 못했다 할지라도, 공공 영역에서 생산된 다양한 아날로그 텍스트를 대규모의 원시 데이터(raw data)로 디지털화 하는 성과는 거두었고, 그것은 우리나라가 인터넷 강국으로 도약하는 데에 작지 않은 기여를 했다. 그러나 이때 성과를 보지 못한 ‘지식 정보 데이터의 연계’에 관해서는 정부도 학계도 별다른 노력도, 심화된 논의조차도 없이 오늘에 이르고 있다. 현대 사회를 ‘초고도 연결 사회’ (Hyper-connectivity society), ‘4차 산업혁명 시대의 사회’라고 하면서.....

관과 조직, 그리고 세분화된 분과학문의 벽을 넘어서는 소통의 큐레이션이 그것이다.

한국에는 유교박물관이라 할 만한 것이 전국 곳곳에 있다. 국공립박물관이나 인가된 사립박물관뿐 아니라, 오래 동안 대를 이어온 가문의 종택이나 유교 관련 인물의 유적지에 건립된 기념관, 유물 전시관.... 그밖에도 무수히 많은 개인, 기관, 단체들이 보존할 가치가 있는 유교 문화의 유물을 가지고 있다. 하지만 이런 식으로 흩어져 있는 유물들은 그것의 가치를 드러내는 큐레이션의 손길이 닿기 어렵거니와 일반인들에게 그 존재를 알리기도 쉽지가 않다. 메타 아카이브의 디지털 콜렉션은 이처럼 실물로 수집할 수 없는, 하지만 이야기로 꾸미고 세상에 알릴 필요가 있는 것들을 포함한다.

역사의 흔적을 자기 몸에 담고 있는 실물 유산은 여러 곳에 흩어져 있는 것이 당연하다. 인위적으로 그것을 한 곳에 모으려는 시도는 자칫 그 유물이 갖는 장소성의 가치를 훼손시킬 수도 있다. 흩어져 있는 유물들을 체계적으로 모아서 그 분야의 전체를 볼 수 있게 하는 망라성은 유물 수집의 한계가 명백한 실물 아카이브가 아니라, 그 한계에 구애될 필요가 없는 메타 아카이브가 지향할 일이다. 실물 아카이브와 메타 아카이브는 각각 다른 존재 이유가 있고, 그 두 가지는 불가결한 상보의 관계를 갖는다. 실물 아카이브는 사라져서는 안 될 실물을 발굴하고 보존하는 데 역점을 두는 곳이고, 메타 아카이브는 그것에 대해 더 많은 사람들이 알고, 그 앎을 부분적인 것에서 전체적인 것으로, 얇은 지식에서 문화에 대한 깊이 있는 성찰로 심화시키는 곳이다.

5. 디지털 큐레이션을 위한 데이터 모델 설계

백과사전적 아카이브 또는 메타 아카이브에 대한 논의가 이론에 머물거나 추상적인 계획에 그치지 않기 위해서는 그것을 실제로 구현할 수 있게 하는 구체적인 데이터 모델이 마련되어야 한다. 필자와 함께 한국학중앙연구원에서 앞에서 소개한 ‘한국 문화 엔사이브’ 개발연구를 비롯해서, ‘전통문화 융복합 지원을 위한 지능형 검색 플랫폼 구축’³⁵⁾, ‘조선왕실 태실 조사 자료 아카이브’³⁶⁾, 문화재 해설문 데이터베이스 구축 연구³⁷⁾ 등 전통문화유산을 대상으로 하는 일련의 디지털 인문학 연구 프로젝트를 수행하면서 백과사전적 아카이브 또는 메타 아카이브의 구현에 적용할 수 있는 플랫폼을 만들고, 실제적인 시스템 구현을 통해 그것을 검증하고 개선하는 노력을 기울여 왔다. 세계유교문화박물관 디지털 아카이브 설계에 참고가 되도록, 융합적 디지털 아카이브를 위한 기본적인 데이터 모델 (엔사이브 데이터 모델, Encyves Data Model) 을 소개한다.

1) 온톨로지

엔사이브 데이터 모델은 기본적으로 온톨로지(ontology) 개념에 입각하여 만들어진 네트워크 데이터베이스 모델이다. 정보 요소 하나하나가 서로에 대해 어떤 의미의 관계가 있는지를 보여 주기 때문에 ‘의미 기반 데이터베이스’(Semantics-based Database, Semantic Database)라고도 한다.

‘온톨로지’란 정보화의 대상이 되는 세계를 전자적으로 표현할 수 있도록 구성된 데이터 기술 체계이다.³⁸⁾ 원래 온톨로지라는 말은 철학에서 ‘존재론’이라고 번역되는 용어로서 ‘존재에 대한 이해를 추구하는 학문’의 의미를 갖는 말이었다. 그러한 용어가 정보과학 분야에서 중요한 개념으로 등장하게 된 것은 인간이 세계를 이해하는 틀과 컴퓨터가 정보화 대상(콘텐츠)을 이해하는 틀 사이에 유사성이 있다고 보았기 때문이다. 그 틀은 바로 대상을 구성하는 요소들에 대응하는 개념들과 그 개념들 간의 연관 관계이다.³⁹⁾

35) 문화체육관광부의 2017년 문화기술 연구개발 지원 사업 (2017~2019)

36) ‘고양 서삼릉 내 묘역 및 태실 역사성 회복 연구’ (2019, 문화재청 연구사업)의 일환으로 구축

37) 문화재청의 2019년 ‘문화재 안내판 정비사업’ (2019~2023)

38) 정보기술 분야에서 말하는 ‘온톨로지(ontology)’에 대한 가장 일반적인 정의는 그루버(Gruber, Thomas. 1959~)가 말한 ‘명시적 명세화의 방법에 의한 개념화’(explicit specification of a conceptualization)이다. (Gruber, ‘A Translation Approach to Portable Ontology Specifications’, *Knowledge Systems Laboratory Technical Report KSL 92-71*, Stanford University, 1992) 여기서 ‘개념화’(conceptualization)라는 것은 정보화하고자 하는 대상 세계를 일정한 체계 속에서 파악하는 것, 예를 들면 그 세계에 무엇이 있고, 그것은 어떤 속성을 품고 있으며, 그것들 사이의 관계는 무엇인가 하는 일정한 질문의 틀 속에서 대상 세계를 이해하는 방식이라고 할 수 있다. ‘명세화’(specification)란 대상 세계에 존재하는 개체, 속성, 관계 등을 일목요연한 목록으로 정리하는 것, 그리고 ‘명시적’(explicit)이라는 그 정리된 목록을 사람뿐 아니라 ‘컴퓨터가 읽을 수 있도록’(machine readable) 한다는 것이다.

39) 김현, 「한국 고전적 전산화의 발전 방향 - 고전 문집 지식 정보 시스템 개발 전략 -」, 『민족문화』

넓은 의미에서는 모든 정보화의 틀이 다 온톨로지일 수 있겠지만, 대상 자원을 ‘클래스’(class)로 범주화하고, 각각의 클래스에 속하는 개체(individuals)들이 공통의 ‘속성’(attribute)을 갖도록 하고, 그 개체들이 다른 개체들과 맺는 ‘관계’(relation)를 명시적으로 기술하는 것이 가장 일반적인 온톨로지 설계 방법이라고 할 수 있다.⁴⁰⁾

엔사이브 데이터 모델의 표준적인 구성은 3 개의 기본 데이터 섹션과 3개의 부가 데이터 섹션으로 이루어진다.

구분	데이터 모델	성격
기본 데이터 섹션	① 사실과 문맥 데이터 (Factual Data)	데이터 네트워크
	② 카테고리 데이터 (Category Data)	분류 목록
	③ 스토리라인 데이터 (Storyline Data)	스토리텔링
부가 데이터 섹션	① 공간 데이터 (Spatial Data)	전자지도
	② 시간 데이터 (Temporal Data)	전자연표
	③ 대역어 데이터 (Translation Data)	다국어 정보 서비스

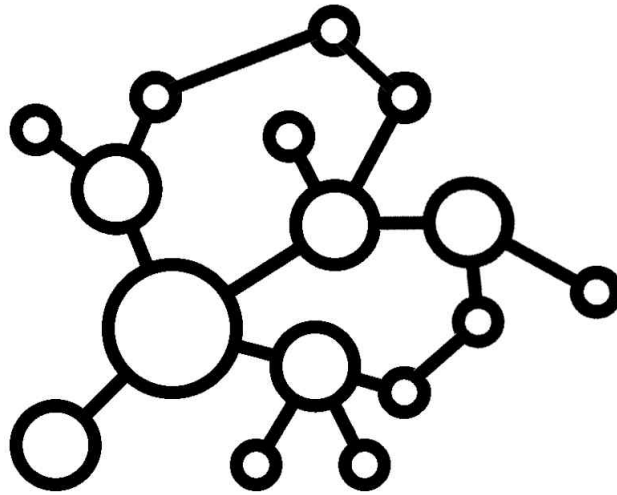
다음은 이 6개 데이터 모델의 온톨로지와 데이터 예시이다. 데이터 모델 별로 클래스의 종류와 각각의 클래스 속하는 개체(Individual Object)들 사이의 관계를 제시한다.

2) 기본 데이터 섹션

① 사실과 문맥 데이터 (Factual Data)

28 (2005)
40) 온톨로지 설계 용어

온톨로지 구성 요소	용도	Web Ontology Language (OWL)
Class, 클래스	공통의 속성을 가진 개체들을 묶는 범주 a group of individuals that belong together because they share some properties.	owl:Class
Individual, 개체	클래스에 속하는 개체 Instances of classes	owl:NamedIndividual
Relation, 관계	(같거나 다른 클래스에 속하는) 개체들 사이의 관계 relationships between pairs of individuals	owl:ObjectProperty
Attribute, 속성	개체가 속성으로 갖는 데이터 값 relationships from individuals to data values	owl:DatatypeProperty
Relation Attribute, 관계 속성	관계 정보에 부수되는 속성 attributes related to relations	N/A in OWL
Domain, 정의역	특정 Object Property 또는 Datatype Property의 주어가 될 수 있는 클래스를 한정 A domain of a property which limits the individuals to which the property can be applied	rdfs:domain
Range, 치역	특정 Object Property의 목적어가 될 수 있는 클래스를 한정 The range of a property limits the individuals that the property may have as its value	rdfs:range

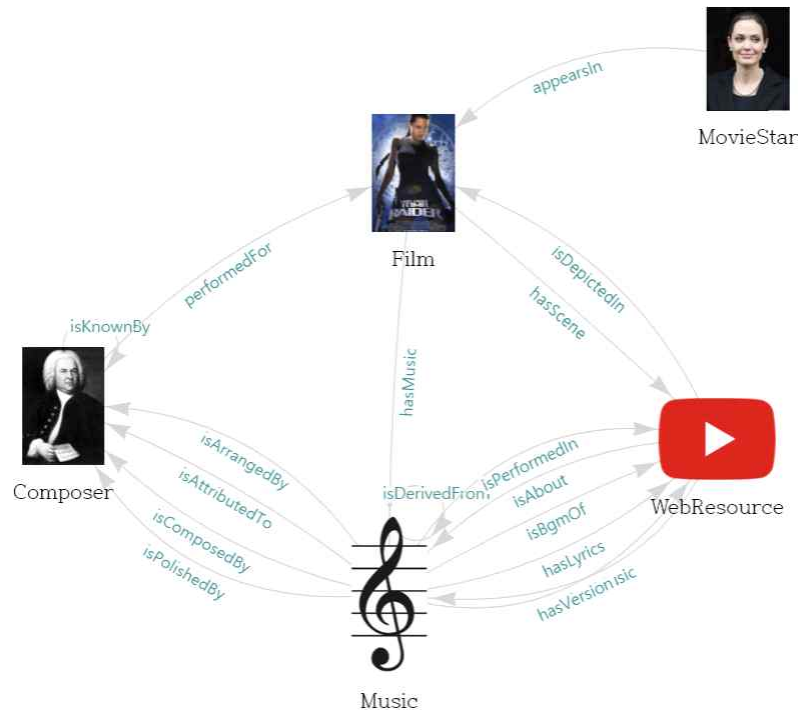


‘사실과 문맥 데이터’ 개념도

‘사실과 문맥 데이터’ 섹션은 백과사전적 아카이브의 모든 지식 요소를 노드와 링크로 구성한 시맨틱 네트워크 데이터 전체를 말한다. ‘사실과 문맥 데이터’의 세부 온톨로지는 아카이빙 및 큐레이션의 대상, 목적 등에 따라 다르게 설계될 수 있으며, 기존의 온톨로지도 계속 확장시켜 갈 수 있다.

문화유산에 관한 ‘사실과 문맥 데이터’ 온톨로지의 예시는 《부록 1》에 수록하였다. 여기서는 ‘사실과 문맥 데이터’의 성격을 쉽게 이해할 수 있도록 단순화한 예시를 소개하겠다. 아래의 그림은 영화 속에서 연주된 클래식 음악의 컬렉션을 중심으로 그 음악과 영화에 대한 이야기를 보이기 위한 ‘사실과 문맥 데이터’의 온톨로지 예시이다. ‘음악 작품’, ‘음악가’, ‘영화’, ‘영화배우’, ‘웹 자원’ 등 5개의 클래스를 설계하고, 각 클래스 속의 정보 노드들 사이의 다양한 관계를 정의하였다.

※ 온톨로지 설계 예시 큐레이션 주제: 속의 클래식 음악



- Class: Music, Composer, Film, MovieStar, WebResource
- Relation:

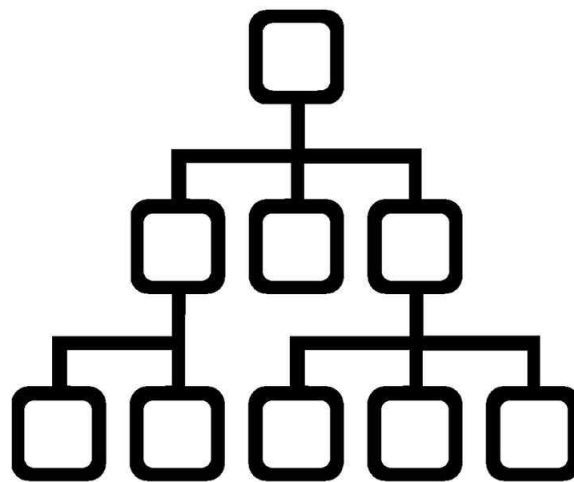
Object	Property	Domain	Range
	isComposedBy	Music	Composer
	isArrangedBy	Music	Composer
	isPolishedBy	Music	Composer
	isAttributedTo	Music	Composer
	isDerivedFrom	Music	Music
	isPerformedIn	Music	WebResource
	isBgmOf	Music	WebResource
	hasLyrics	Music	WebResource
	hasVersion	Music	WebResource
	hasLyrics	Music	WebResource
	isKnownBy	Composer	Composer
	hasMusic	Film	Music
	hasScene	Film	WebResource
	appearsIn	MovieStar	Film
.....			

② 카테고리 데이터 (Category Data)

‘사실과 문맥 데이터’ (Factual Data)가 정보 요소와 요소 사이의 사실적 관계에 주목하는 것이라면, ‘카테고리 데이터’는 관리자가 정한 분류체계 따라 정보 요소들을

그룹화 하는 것이다. 네트워크 데이터베이스 속에서 특정 주제나 유형의 정보를 검색할 수 있게 한다.

기존의 릴레이셔널 데이터베이스 기반의 정보 시스템에서는 특정 메타데이터를 가지고 전체 데이터를 분류해서 주제별, 유형별로 열람할 수 있게 하는 기능을 제공해 왔다. 백과사전적 아카이브에서도 그와 유사한 데이터 서비스가 이루어질 수 있도록 하는 것이 ‘카테고리 데이터 섹션’의 역할이다. ‘사실과 문맥 데이터’ 노드의 속성 정보 가운데 분류 카테고리 역할을 할 수 있는 것들을 특정함으로 생성한다.



카테고리 데이터 개념도

※ 온톨로지 설계



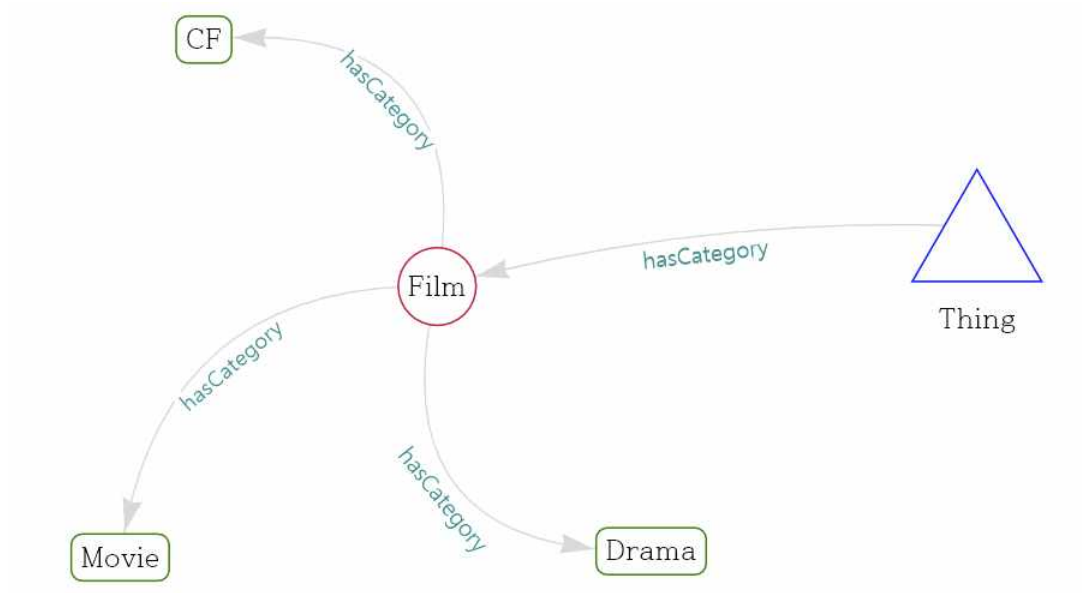
- Class: Category, Individual⁴¹⁾

- Relation:

Object Property	Domain	Range
hasMember	Category	Category
	Category	Individual

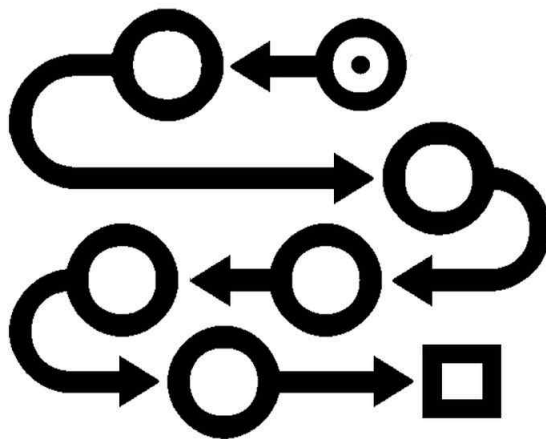
41) 여기서의 Individual Class는 ‘사실과 문맥 데이터’에 있는 모든 개체(Individual Object)를 멤버로 삼는 클래스이다. 아래의 다른 데이터 섹션에 있는 Individual Class도 같은 성격이다.

※ 카테고리 데이터 예시



③ 스토리라인 데이터 (Storyline Data)

어느 주제에 대해서 광범위하게 조사·정리된 ‘사실과 문맥 데이터’ 데이터가 있을 때, 그 네트워크의 특정 노드들을 가지고 큐레이션의 줄거리(Storyline)를 만들어서 고객에게 ‘이야기’(Storytelling)를 제공하는 데이터이다.

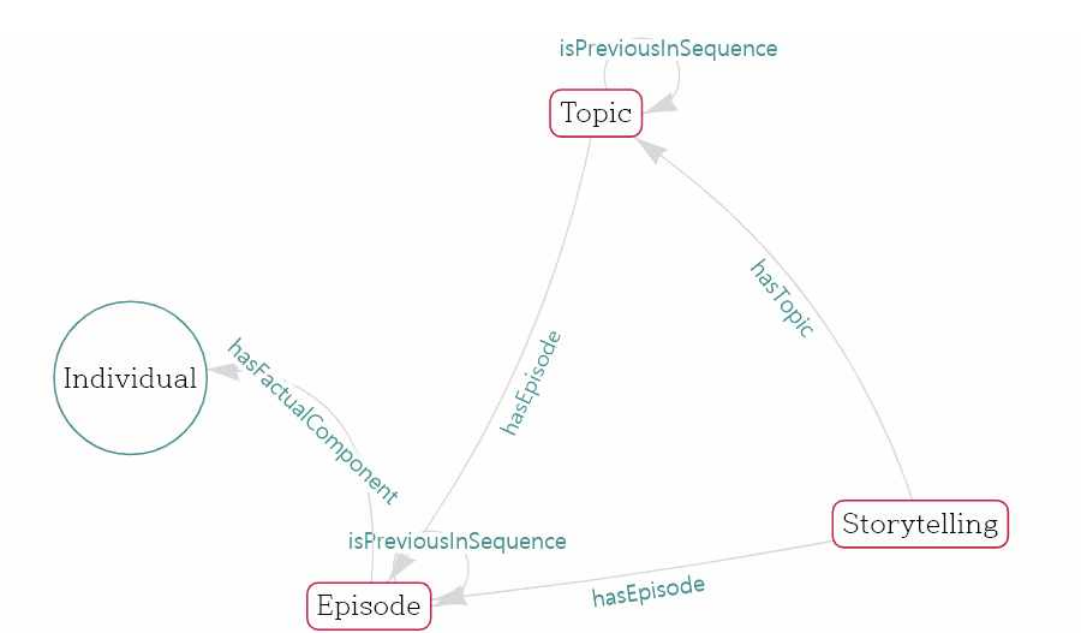


스토리라인 데이터 개념도

‘사실과 문맥’ 데이터는 하나의 대상을 중심으로 의미 있는 주변의 관계를 보여주지만, 그 관계의 확장은 너무 광범위해서 그 모든 것을 한꺼번에 들여다보거나 이해할 수 없다. 박물관의 전시 큐레이션은 기본적으로 다양한 자원 속에서 기획의도에 부

합하는 대상을 선택적으로 취해서 의미 있는 이야기의 자원으로 활용하는 것이다. 이와 유사하게 네트워크 데이터베이스 속에서 특정 노드들을 선택하여 주제가 있는 이야기의 자원으로 활용할 수 있게 하는 장치가 스토리라인 데이터이다.

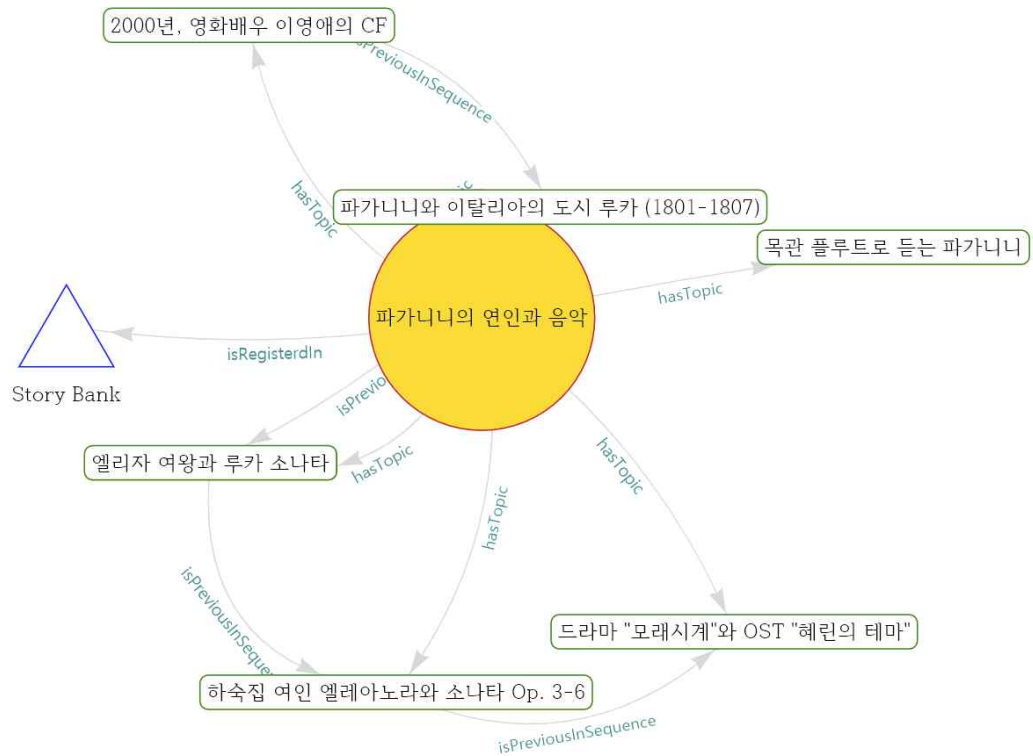
※ 온톨로지 설계



- Class: Storytelling, Topic, Episode, Individual
- Relation:

Object	Property	Domain	Range
hasTopic		Storytelling	Topic
		Topic	Topic
hasEpisode		Topic	Episode
		Episode	Episode
hasFactualComponent		Episode	Individual
isPreviousInSequence		Topic	Topic
		Episode	Episode

※ 스토리라인 데이터 예시

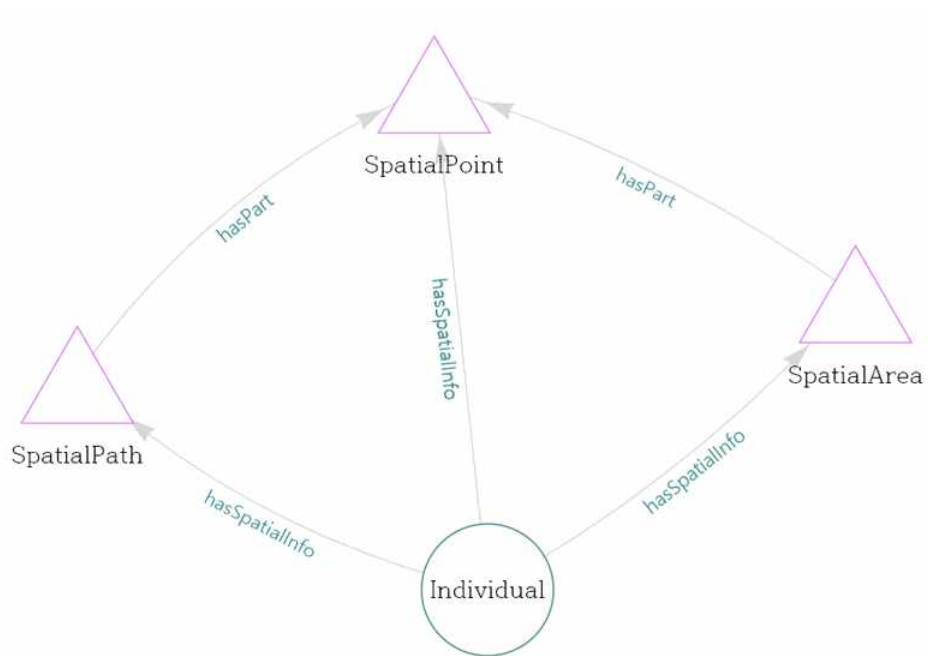


3) 부가 데이터 세트

① 공간 데이터 (Spatial Data)

‘사실과 문맥 데이터’ 속에 있는 모든 공간 정보를 전자지도 상에 표시할 수 있도록 하고, 각 지점들의 다양한 연관 관계를 탐색할 수 있게 하는 데이터이다.

※ 온톨로지 설계



- Class: SpatialPoint, SpatialArea, SpatialPath, Individual
- Relation: hasPart, hasSpatialInfo

Object Property	Domain	Range
hasPart	SpatialArea	SpatialPoint
	SpatialPath	SpatialPoint
hasSpatialInfo	Individual	SpatialPoint
	Individual	SpatialArea
	Individual	SpatialPath

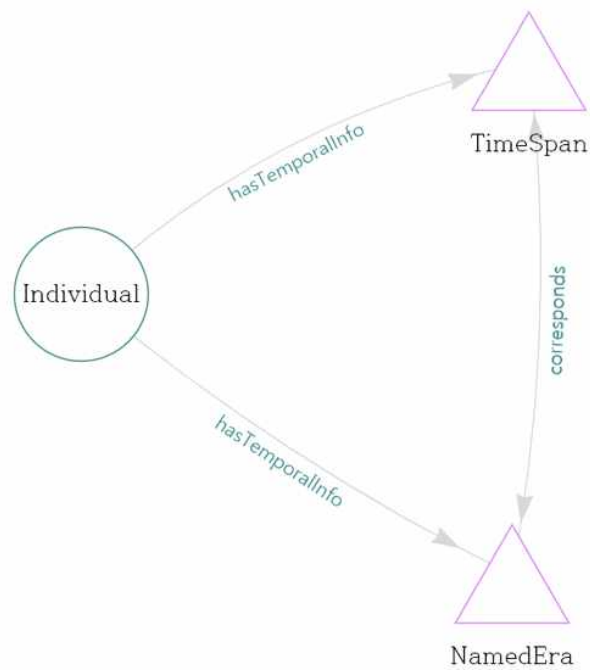
※ 공간 데이터 예시



② 시간 데이터 (Temporal Data)

‘사실과 문맥 데이터’ 속에 있는 모든 시간 정보를 전자연표 상에 표시할 수 있도록 하고, 각 시점들의 다양한 연관 관계를 탐색할 수 있게 하는 데이터이다.

※ 온톨로지 설계



- Class: TimeSpan, NamedEra, Individual

- Relation:

Object	Property	Domain	Range
	corresponds	TimeSpan	NamedEra
		NamedEra	TimeSpan
	hasTemporalInfo	Individual	TimeSpan
		Individual	NamedEra
		Individual	SpatialPath

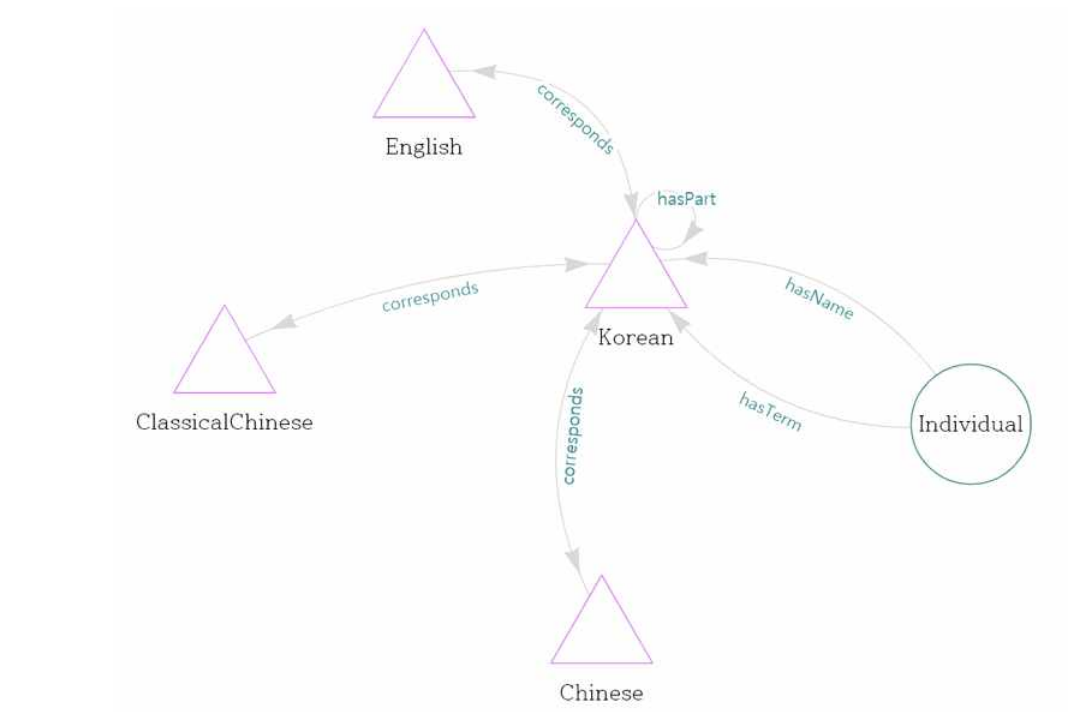
※ 시간 데이터 예시

Time Span	Image	Label	Description	Network
1390		【문헌】 고려 말 화령부 호적 관련 문서		
1394		【문헌】 진충귀를 의주목사로 임명한다는 사령서		
1397		【금속공예】 청동 은입사 청곡사 향완		
1399		【문헌】 오묘한 진리를 담은 연꽃과 같은 부처님의 가르침		
14c-15c		【불상】 금동십일면관음보살좌상		
14c-15c		【불상】 금동관음보살좌상		
14c-15c		【금속공예】 철제 금은입사 삼인검		
14c-15c		【금속공예】 동제 팔괘 무늬 네모 거울		
14c-15c		【불상】 금동보살좌상		
14c-15c		【문헌】 명태조가 세정한 의례에 관한 책		
14c-15c		【도자기】 백자 철화 파초 무늬 합		

③ 대역어 데이터 (Translation Data)

‘사실과 문맥 데이터’ 속에 있는 정보 요소(노드) 이름에 대응하는 외국어 대역어 데이터이다. 다중 언어 큐레이션 및 데이터 서비스를 위한 장치이다.

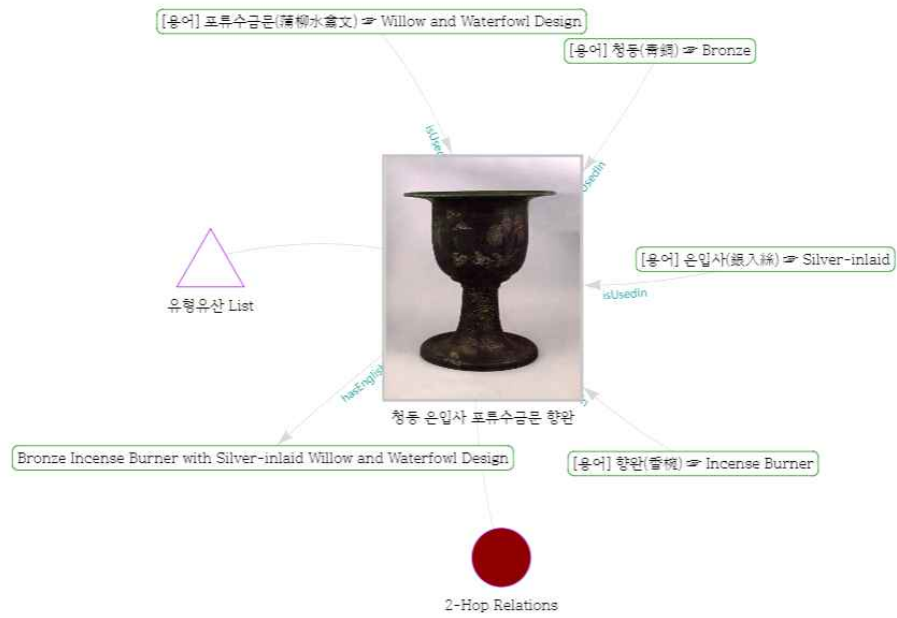
※ 온톨로지 설계



- Class: Korean, English, Chinese, ClassicalChinese, Individual
- Relation:

Object	Property	Domain	Range
	corresponds	Korean English Chinese ClassicalChinese	Korean English Chinese ClassicalChinese
	hasName	Individual	Korean
	hasTerm	Individual	Korean
	hasPart	Korean	Korean

※ 대역어 데이터 예시



6. 세계유교문화 박물관 디지털 아카이브의 디지털 큐레이션

1) 유교문화연구와 데이터 편찬의 융합적 큐레이션

백과사전적 아카이브의 데이터 모델을 앞에서 보인 것처럼 다층적인 시맨틱 데이터 베이스로 설계한 이유는 디지털 자원의 아카이빙과 그것을 활용한 큐레이션을 하나의 시스템 속에서 실현할 수 있게 하려는 것이다.

우리가 익숙하게 보고 있는 대부분의 디지털 아카이브 데이터베이스는 2차원 테이블 형식의 일정한 틀 안에 모든 대상 자원에 대한 정보를 집어넣어서, 그 틀에서 정한 분류 기준에 의해 데이터가 탐색되도록 하는 기능을 제공한다. 이것은 제시한 여러 가지 데이터 모델 중 두 번째 ‘카테고리 데이터’의 역할과 비슷한 기능이다. 이러한 수준에 머무는 데이터베이스는 그야말로 이 아카이브 속에 어떤 자원이 있는지를 알려 주는 것 이상의 기여를 하기 어렵다. 그 자원을 가지고 의미 있는 지식을 탐구하고 이야기를 만드는 것은 이용자가 완전히 새롭게 시작해야 할 일이다. 아카이브의 자원을 가지고 그와 관련한 지식의 세계를 탐구하기 시작할 때, 필수적으로 알고 이해해야 할 정보가 있다면 그것도 그 디지털 자원 속에 포함되도록 하고, 이러한 기초 자원을 가지고 새로운 발견을 얻어냈다면 그것도 다시 데이터베이스 안에 축적될 수 있도록 하는 것이 아카이브 데이터베이스에 지속가능한 생명력을 불어넣는 길이다.

디지털 아카이브의 중심 데이터가 창고 관리인의 재고 물품 목록 수준에 머물지 않고, 큐레이터와 이용자가 함께 의미 있는 지식을 탐구하는 지식 활동의 환경이 될 수 있게 하는 첫 번째 요건은 그 아카이브의 전문 영역 안에서 풍부한 학술적 지식을 가진 큐레이터들이 직접 그 데이터를 편찬의 실무를 담당하는 것이다. 백과사전적 데이터 모델 가운데, 첫 번째 ‘사실과 문맥 데이터’는 전문 연구자들이 직접 그 데이터를 다루지 않으면 만들어질 수 없는 데이터이다. 세 번째 ‘스토리라인 데이터’ 역시 물리적인 전시실 안에 꾸미는 특별기획전시 연출할 수 있는 지식과 안목이 있는 큐레이터가 직접 그 데이터를 조직할 때 가장 우수한 결과물이 나올 수 있을 것이다.

물론, 디지털 아카이브의 세계에서는 실물 아카이브에서보다 훨씬 더 개방적이고 자유롭게 대상에 대한 실험을 할 수 있기 때문에, 그 분야의 최고 전문가 아니라 하더라도 지적 탐구의 의지가 강한 학생들과 일반인들도 시맨틱 데이터의 확장과 스토리라인 개발에 참여할 수 있다. 궁극적으로는 이와 같은 클라우드 소싱의 길을 열어놓아야 그 디지털 아카이브의 생명력이 유지된다. 고객의 참여는 그들의 관심과 애정을 유발할 가장 좋은 견인 장치임이 분명하다. 클라우드 소싱이 아카이브 데이터의 품질을 떨어뜨리지 않고, 선순환적인 발전의 체제가 될 수 있게 하는 요건은 참여자들로 하여금 무엇을 어떻게 해야 하는지를 알게 ‘교육’이다. 전문 연구자들이 선도적으로 만든 우수한 콘텐츠가 바로 이들이 보고 따라할 교육 콘텐츠의 역할을 할 것이

다.

대부분 전통적인 인문학 분야의 교육 훈련을 받은 큐레이터와 연구원들은 지식을 탐구하고 체계화 하는 일을 아날로그 환경에서만 수행해 왔기 때문에 디지털 데이터를 다루는 일에 익숙하지 못하다. 하지만 지금 이들에게 주어진 임무가 디지털 아카이브를 구축하는 일이라면 당연히 그들의 전문성을 디지털 환경에서 발휘할 수 있는 방법을 배워서 실천해야 한다. 디지털 아카이브를 위한 지식 콘텐츠를 과거에 해오던 식의 문자언어로만 생산하고, 그것을 아카이브 자원에 대해 전혀 문외한인 정보기술자들이 알아서 디지털화 하게 하는 것은 그 지식의 활용성을 높이기보다는 그것을 디지털 매체 속에 박제화 하는 일이 되기 쉽다. 디지털 데이터 편찬은 그 내용에 대한 이해와 발전적인 활용에 관한 비전을 가지고 하는 일인데, 그 안목이 없다면 주어진 것을 고정화하는 것밖에는 할 일이 없기 때문이다.

2) 세계유교문화 박물관 디지털 큐레이션 실천 방안

2020년 세계유교문화 박물관의 개관과 함께 본격적으로 가동하게 될 이 박물관의 디지털 아카이브는 지금 이곳의 대표 콘텐츠를 어떠한 성격과 방향으로 확충해 나갈 것인지에 대해 구체적인 실천 방안을 방책을 강구해야 할 시점에 있다. 이에 대한 기획의 출발점은 2019년에 편찬에 착수한 세계유교문화 디지털 오브젝트의 해설 콘텐츠이다.

이 해설 콘텐츠는 문명관, 정신관, 미래관 등 세계유교문화 박물관의 전시 주제 영역에 맞추어, 유교 사회의 제도와 규범, 일상문화, 예술, 유교 지식인, 한국·중국·일본의 유교 경전, 고문서, 그리고 현대 사회에서의 유교적 가치 문제 등에 관한 4,000여 항목을 개발하고, 전문연구자들이 이에 대한 해설문 텍스트를 집필하는 방식으로 생산되고 있다. 이것은 일차적으로 박물관에 전시될 실물, 또는 디지털 오브젝트의 해설 자료로서 일차적인 기능을 수행하겠지만, 그 이상으로 중요한 활용 가치는 이 기초 콘텐츠가 박물관 컬렉션 전체에 더 큰 생명력을 불어넣는 통섭적 스토리텔링의 토대가 될 수 있다는 점이다.

이 4,000여 건의 기사는 하나하나 개별적인 대상에 대한 해설문이기도 하지만, 그 전체는 4,000여 항목으로 이루어진 세계유교문화 백과사전이라고 보아도 무방할 것이다. 이 백과사전의 개별 기사 항목들은 현재 서로 서로에 대해 어떠한 연관관계가 있는지, 그 의미 맥락을 드러내지 못한 채 단편적인 지식의 조각으로 존재한다. 박물관의 전시실이 몇 가지 중요한 사실을 이야기로 꾸며서 전달하겠지만, 그 한정된 공간에서 보일 수 있는 것은 깊이와 다양성 면에서 명백한 한계가 있다. 디지털 아카이브의 활용가치를 높이는 디지털 큐레이션이 필요하다.

한국의 유교와 세계의 유교, 과거의 유교와 현재, 미래의 유교에 대한 4,000여 편의 지식정보 데이터 안에는 여러 가지 유물, 유적, 인물, 문헌, 역사적 사건, 철학적 개념에 대한 설명이 들어있을 것이고, 그것들을 세밀하게 들여다보면 그 중의 많은

부분이 서로 서로 거미줄처럼 얽혀서 이 세계가 품어온 유교적 문맥을 엿볼 수 있게 될 것이다. 본 연구를 통해 제시한 백과사전적 아카이브의 데이터 모델은 바로 그와 같은 문맥적 지식의 탐구, 지식의 스토리텔링을 이룰 수 있게 하는 디지털 큐레이션의 프레임워크이다.

국학진흥원에서 이러한 취지의 디지털 큐레이션을 실제적으로 추진하고자 한다면, 그 첫 번째 실천과제는 이 디지털 아카이브의 콘텐츠 특성에 부합하는 데이터 온톨로지를 설계하는 일이다. 필자는 이 지면을 통해 한국문화유산 엔사이브의 온톨로지를 하나의 예시로 제시했지만, 세계유교문화 박물관 디지털 아카이브는 나름대로의 지향점과 특성이 있을 것이므로, 이 부분을 반영한 발전적 설계가 필요하다. 물론 이 온톨로지 설계는 향후 아카이브 컬렉션의 확충에 따라 지속적으로 변화되어 갈 것이다.

그 다음에 따르는 과제는 당연히 4,000여 편의 지식정보 데이터를 현재의 평면적 텍스트에서 광대한 문맥의 탐색의 가능한 시맨틱 데이터로 가공하는 일이다. 사실 미리 준비가 되었다면, 이 작업의 상당 부분은 텍스트 집필 시에 함께 이루어질 수 있는 일이었다. 초기 작업은 이렇게 단계적으로 수행한다 하더라도, 향후의 후속 콘텐츠 생산은 디지털 환경에서 텍스트와 콘텍스트의 데이터를 한꺼번에 만들어낸 방식으로 추진되어야 할 것이다. 이러한 단계를 거쳐 세계유교문화 디지털 아카이브의 기초 콘텐츠가 시맨틱 데이터로 편성되게 되면, 그 속에서는 박물관 전시실의 벽면을 장식하는 몇 가지 이야기보다 훨씬 다양하고 깊이 있는 스토리텔링 자원들이 발굴될 것이다.

그 다음에는 아카이브의 큐레이터뿐 아니라 이 디지털 아카이브를 방문하는 고객들이 그 이야기 자원을 공유하고, 거기에 자신의 탐구한 지식, 자기 집안에서 전래되어 온 유물, 그들의 고장에 있는 유적의 데이터를 덧붙이는 길이 열려야 할 것이다. 이 시점에서 강구해야 할 중요한 과제는 유교문화유산에 진지한 관심을 갖고 있는 고객들에게 유교유산 디지털 큐레이터로 봉사할 수 있는 역량을 교육하는 것이다. 클라우드 소싱에 의한 큐레이션의 확장은 세계가 주목하는 디지털 아카이브의 미래 과제이다.

《부록 1》

한국문화유산 엔사이브 온톨로지

1) 문맥 요소의 클래스(Class) 및 클래스별 데이터 속성(Attribute)⁴²⁾

Main Class	Sub Class	Description	Examples	Attribute Wiki Template
Heritage	문화유산	문화유산으로 지정, 등록되었거나 문화유산으로서 보존 가치가 있다고 판단되는 대상	학봉 안경 , 장흥 보림사 철조비로자나불좌상 , 진도 남도석성	틀:유물유적정보 틀:승탑정보
Actor	인물	대상자료 관련 인물	조선 영조 , 정약용 , 최치원 , 김성일	틀:인물정보
	단체	집단적 행위 주체로서의 단체, 운영주체로서의 기관	집현전 , 가례도감 , 독립협회	틀:단체정보
Event	사건	대상 자료와 직간접적으로 연관되는 역사적 사건	진주전투	틀:사건정보
	행사	기념제, 역사적 사건이나 세시풍속을 재현하는 행사	최무선장군추모제	
Place	장소	자료, 유물과 역사적 연계성을 갖는 장소	창덕궁 인정전 , 백운동 별서 정원	틀:장소정보
	소장처	자료, 유물의 현대 소장처	한국학중앙연구원 , 새마을운동중앙연수원	틀:소장처정보
Object	문헌	고문헌 또는 텍스트 위주의 작품·기사	고산구곡가 , 백운첩	틀:문헌정보
	물품	일상 물품 외에 자연물 포함	겹취건 , 백동화	틀:물품정보
Concept	개념	대상자료 관련 주요 개념	가례 , 나당연합 , 사천왕 , 파스파 문자	틀:개념정보

42) 클래스별 데이터 속성은 이 표의 Wiki Template Link를 통해 확인 가능

2) 문맥요소 사이의 릴레이션 (Relation, Object Property)

NameSpace	Relation	Inverse Relation	Notes	Example
dcterms:	A creator B	B isCreatorOf A	창작자(저자) 로 인정받는 제작자	도산십이곡 - 이항
ekc:	[s] A writer B	B isWriterOf A	글 쓴 사람. 비문 지은이.	김언경 - 장흥 보림사 보조선사탑비
ekc:	[s] A calligrapher B	B isCalligrapherOf A	글씨 쓴 사람. 비문/편액/제 액 글씨 쓴 사람.	김원(신라) - 장흥 보림사 보조선사탑비
ekc:	[s] A inscriber B	B isIncriberOf A	비문 새긴이.	
ekc:	A translator B	B isTranslatorOf A	번역자; 언해자.	태교신기언해 - 유희
ekc:	A annotator B	B isAnnotatorOf A	주석자	용비어천가 - 성삼문
ekc:	A founder B	B isFounderOf A	기관/단체/장 소 설립자. 권위로써 무엇을 만들도록 함.	현대미술가협회 - 하인두
ekc:	A constructor B	B isConstructorOf A	건축을 지휘	소양강댐 - 현대건설
ekc:	A reconstructor B	B isReconstructorOf A	수축(修築)을 지휘	백마사성 - 강감찬
ekc:	A renovator B	B isRenovatorOf A	중창·중건 추진	고성 유점사 - 유정
dcterms:	A contributor B	B isContributorOf A	기여자	장흥 보림사 - 신라 현강왕
dcterms:	A publisher B	B isPublisherOf A	필사/간행/중 간/영인	어제계주윤음 - 교서관
dcterms:	A rightsHolder B	B isRightsHolderOf A	저작권 보유자	세종대왕 표준영정 - 운보문화재단
edm:	A isDerivativeOf B		파생작; 번역물	삼강행실도 언해 - 삼강행실도
edm:	A isSuccessorOf B		속편 작품	속장경 - 초조대장경
ekc:	A hasOldName B	B isOldNameOf A	구 명칭; 전신(前身)	대우중공업 - 한국기계공업
ekc:	A isNamesakeOf B	B isEponymOf A	명확하게 B의 이름을 따온 경우.	홍무정운 - 홍무제
ekc:	A administrates B	B isAdministratedBy A	의례·제도에 서의	가례도감 - 가례

			주관행위	
ekc:	A participatesIn B	B hasParticipant A	의례·제도에 참여	백관 - 아헌례
ekc:	A documents B	B isDocumentedIn A	A:기록물(예: 의궤), B: 기록되는 대상(예:의례)	영조정순왕후가례도감의궤 - 가례
ekc:	A goesWith B		의식에서 물품 간의 수반 관계	외궤 - 내함
ekc:	A isUsedIn B		의례·사건에서의 사용	근배 - 동뢰
edm:	A isNextInSequence B	B isPreviousInSequence A	사건, 항목 간의 직접 연속 관계. B 다음이 A인 순서임.	납폐 - 납채
ekc:	A performed B	B isPerformedBy A	해당 사건을 시작부터 끝까지 주도. 그 외에는 edm:isRelated To.	민족대표 33인 - 독립선언식
ekc:	A isPerformedAt B		사건 실행 장소	독립선언식 - 태화관
ekc:	A hasExhibitionAt B		대여 전시, 소장처 외의 전시 포함.	
edm:	A happenedAt B		사건이 발생한 장소	명량대첩 - 명량해협
ekc:	A depicts B	B isDepictedIn A	묘사의 주된 대상	고산구곡도설 - 고산구곡도
ekc:	A mentions B	B isMentionedIn A	내용 중 언급	연해납약증치방 - 청심원
dcterms:	A references B	B isReferencedBy A	실제 내용의 참조	고산구곡도 - 고산구곡가
ekc:	A isSteleOf B	B hasStele A	인물·사건의 기념비	장흥 보림사 보조선사탑비 - 체징
ekc:	A isStupaOf B	B hasStupa A	승탑	장흥 보림사 보조선사탑- 체징
ekc:	A isEnshrinedIn B	B enshrines A	제향	이순신 - 현충사
edm:	A currentLocation B	B isCurrentLocationOf A	현재 소장처	고성 유점사 종 - 고성 보현사
edm:	A formerLocation B	B isFormerlocationOf A	과거 소장처	고성 유점사 종 - 고성 유점사
dcterms:	A provenance B	B isProvenanceOf A	출처; 원 소장처;	왕오천축국전 - 돈황석굴

			발견·발굴 장소	
ekc:	A hasWife B (=isHusbandOf)	B isWifeOf A (=hasHusband)	처	사도세자 - 혜경궁 홍씨
ekc:	A hasConcubine B (=isHusbandOf)	B isConcubineOf A (=hasHusband)	첩	
ekc:	A hasSon B (=isFatherOf)	B isSonOf A (=hasFather)	부자관계	김주국 - 김광찬
ekc:	A hasSon B (=isMotherOf)	B isSonOf A (=hasMother)	모자관계	혜경궁 홍씨 - 조선 정조
ekc:	A hasDaughter B (=isFatherOf)	B isDaughterOf A (=hasFather)	부녀관계	장흥효 - 장계항
ekc:	A hasDaughter B (=isMotherOf)	B isDaughterOf A (=hasMother)	모녀관계	
ekc:	A hasAdoptedHeir B	B isAdoptedHeirOf A	양자(가계상 속 목적)	윤이석 - 윤두서
ekc:	A hasBrother B		형제	조선 세종 - 효령대군
ekc:	A hasSister B		자매	
ekc:	A hasSonInLaw B (=isFatherInLawOf)	B isSonInLawOf A (=hasFatherInLaw)	장인-사위	
ekc:	A hasSonInLaw B (=isMotherInLawOf)	B isSonInLawOf A (=hasMotherInLaw)	장모-사위	
ekc:	A hasDaughterInLaw B (=isFatherInLawOf)	B isDaughterInLawOf A (=hasFatherInLaw)	시부-며느리	
ekc:	A hasDaughterInLaw B (=isMotherInLawOf)	B isDaughterInLawOf A (=hasMotherInLaw)	시모-며느리	
ekc:	A hasDescendant B	B isDescendantOf A	세대 불특정 선조-후손 관계	김성일 - 김주국
ekc:	A isLineageKinOf B		직계친족	
ekc:	A isAffinalKinOf B		인척(혼인으로 맺어진 친척)	

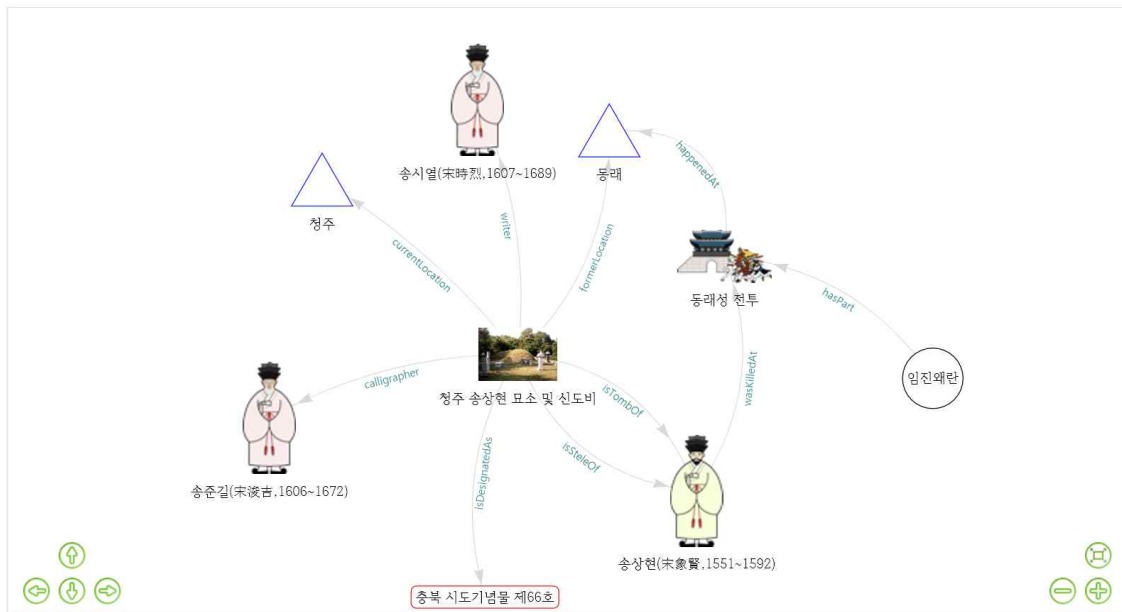
ekc:	A hasDisciple B (=isMasterOf)	B isDiscipleOf A (=hasMaster)	스승-제자 관계	이항 - 김성일
ekc:	A hasOwner B	B isOwnerOf A	노비-주인 관계	일복 - 윤두서
ekc:	A hasSubject B	B isSubjectOf A	임금-신하	고려 우왕 - 최영
ekc:	A servesAs B		(관직/직무/역 할)을 수행	서유문 - 서장관
ekc:	A wasOrdainedBy B	B wasPreceptorOf A	계를 받았다(승려 간)	수철 - 윤법
foaf:	A knows B		일반적인 교유관계	신숙주 - 성삼문
ekc:	A isFellowOf B		동문수학/동 류	
dcterms:	A hasPart B	B isPartOf A	전체-부분 관계	도산십이곡 - 도산십이곡 발문
foaf:	A member B	B isMemberOf A	조직-구성원 관계	집현전 - 정인지
owl:	A sameAs B		동일 또는 유사 관계	훈민정음 - 한글
ekc:	A isNear B		물리적 인접	금산 칠백의총 - 금산 칠백의총 기념관
ekc:	A wears B	B isWornBy A	궁중복식에서 입는 대상	궁녀 - 당의
dcterms:	A type B		특정 분류체계 배속 시. 그 외의 성격/유형은 edm:isRelated To.	세종대왕 - 표준영정 - 표준영정
edm:	A isRelatedTo B		기타 포괄적 연관 관계	

《부록 2》

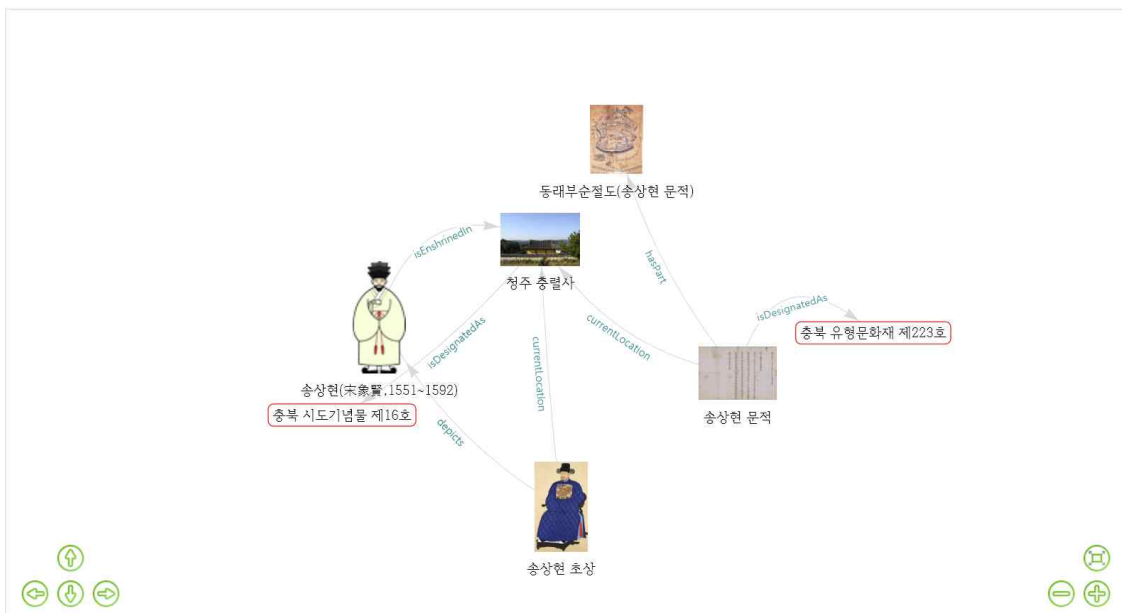
개별 문화유산의 시맨틱 데이터 및 확장 네트워크 예시

1) 스몰 데이터 예시: 개별 문화유산의 시맨틱 데이터

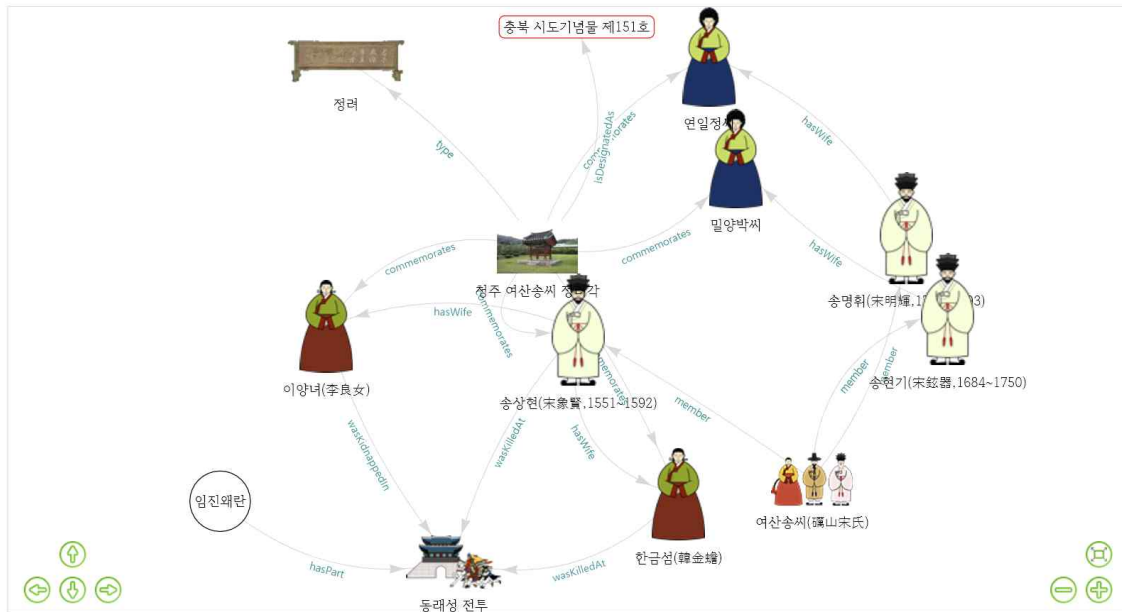
① 충북 시도기념물 제66호: 청주 송상현 묘소 및 신도비



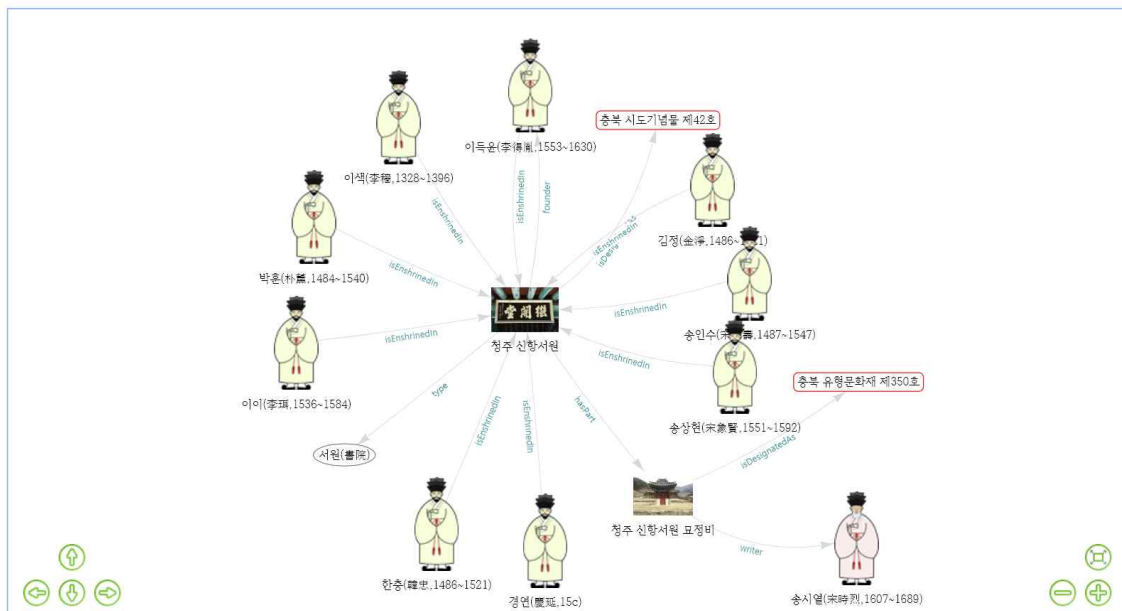
② 충북 시도기념물 제16호: 청주 충렬사 / 충북 유형문화재 제223호: 송상현 문적



③ 충북 시도기념물 제 151호: 청주 여산송씨 정려각



④ 충북 시도기념물 제42호: 신행서원 / 충북 유형문화제 제 350호: 신행서원 묘정비



○ 송상현 이야기 디지털 큐레이션

