

디지털 큐레이션: 미래 세대를 위한 디지털 인문학 교육

Digital Curation: Digital Humanities Education for Future Generations

김현(Kim, Hyeon)*

목차

1. 디지털 원어민, 디지털 리터러시
2. 디지털 큐레이션
3. 디지털 아카이브
4. 백과사전적 아카이브
5. 시맨틱 데이터 아카이브
6. 디지털 스토리텔링과 AI 내러티브
7. 디지털 큐레이션 교육을 위한 제언

초록

나는 이 글에서 ‘디지털 원어민’ 세대의 인문학도들이 실천할 수 있는 ‘디지털로 인문학 공부하기’의 방법을 ‘디지털 큐레이션’이라는 이름으로 제안하였다. 디지털 큐레이션은 디지털 미디어를 통해 획득한 디지털 자원을 체계적으로 조직화하여 지식과 이야기를 담은 디지털 콘텐츠를 만들어 내는 행위이다.

‘디지털 큐레이션’의 기술적인 방법론은 ‘시맨틱 데이터’의 편찬이다. ‘시맨틱 데이터’란 독립된 데이터의 나열이 아니라, 데이터 사이의 연결 방식을 표현함으로써 데이터에 담긴 의미를 명시적으로 설명할 수 있도록 구성된 정보 기술 형식을 말한다.

디지털 큐레이션은 디지털 아카이브라는 결과물을 만들어 낸다. ‘시맨틱 데이터’ 기반의 인문 지식 디지털 아카이브는 우리의 인문학적 탐구의 대상이 되는 세계를 데이터로 재현한 저작물이다. 이렇듯 큐레이팅 된 인문학 자원의 저장소인 시맨틱 데이터 아카이브에 대해 우리는 다음과 같은 활용성을 기대할 수 있다.

첫째, 여러 학문 분야에 걸친 데이터 통합: 학제간 의사소통과 협업을 촉진하여 인간의 지식과 문화를 총체적으로 이해할 수 있게 한다.

둘째, 교육적 응용 지원: 교육자와 학생은 이 아카이브를 다양한 문화 요소의 상호 연결성을 탐구하는 자원으로 사용할 수 있다.

셋째, 인공지능-준비성: 시맨틱 데이터 형식으로 큐레이팅 된 인문 지식 콘텐츠는 인공지능이 효과적으로 처리하고 이해하며 활용할 수 있는 데이터이다. 인문학 연구자와 학생들은 시맨틱 데이터 큐레이션을 통해 인공지능과 협업하는 미래 인문학을 준비해 갈 수 있다.

주제어: 디지털 인문학, 디지털 큐레이션, 디지털 아카이브, 시맨틱 데이터, 시맨틱 데이터 아카이브, 생성형 스토리텔링

* 단독저자: 한국디지털인문학협의회 고문, 초대 회장. 한국학중앙연구원 인문정보학과 교수, xuanflute@gmail.com

Abstract

In this article, I introduce a method called ‘digital curation’ designed for humanities students who are part of the digital native generation. Digital curation involves systematically organizing digital resources to create informative and narrative-rich digital content. This process utilizes ‘semantic data,’ which is information structured to clearly define its own meaning and its relationships with other data. Semantic data usually involves entities such as people, places, or things, and delineates their interactions and connections.

The primary outcome of digital curation is the creation of ‘digital archives’ that rely on semantic data. These archives serve as repositories for humanistic resources and effectively replicate the subjects studied in the humanities through structured data. The potential benefits of these semantic data-based archives include:

1. Enhancing interdisciplinary integration, which facilitates communication and collaboration across various fields, enriching our understanding of human knowledge and culture.
2. Supporting educational uses, where educators and students can utilize the archives to explore the interconnectedness of different cultural elements.
3. Preparing for future integration with AI, as the data format is well-suited for AI processing and utilization, enabling a collaborative future between humanities and artificial intelligence.

Keywords: digital humanities, digital curation, digital archives, semantic data, semantic data-based archives, semantic data-based generative storytelling

내가 생각하는 ‘디지털 인문학의 정의(定義)’는 ‘디지털 환경에서 인문지식을 탐구하고 응용하는 실천적 노력’이다. 이 글에서 이야기하려는 디지털 큐레이션은 이러한 의미의 디지털 인문학을 실현하는 구체적인 수행 방법이다.

디지털 큐레이션은 디지털 아카이브라는 결과물을 만들어 낸다. 그렇게 만들어진 디지털 아카이브는 새로운 디지털 큐레이션의 환경으로 기능하게 될 것이다. 디지털 큐레이션과 디지털 아카이브 사이의 지속적으로 확장되고 정련되는 선순환이 내가 생각하는 디지털 인문학의 세계이다.

1. 디지털 원어민(原語民), 디지털 리터러시(文識)

디지털 원어민(Digital Natives)은¹ 디지털 인문학의 고객이자 미래 인문학의 주역이다. 디지털 원어민들이 전통적인 인문학 지식으로부터 완전히 단절되는 것을 막고, 과거로부터 오늘에까지 이어져 온 인문학이 그들에 의해서 미래 사회에서 지속적으로 발전할 수 있게 하는 것이 디지털 인문학의 당면 과제이다.

디지털 원어민이 디지털 세계에서 지식과 정보를 수용하고 생산할 수 있는 능력을 ‘디지털 리터러시(Digital Literacy, 디지털 문식)’라고 한다. ‘문식(文識)’에 반대되는 개념은 ‘문맹(文盲)’이다. 시력을 잃은 장님이 사물을 식별하지 못하는 것처럼 글을 보아도 그 뜻을 알지 못하는 것을 뜻한다. 그런데 누군가가 글을 읽을 수는 있는데 쓸 줄을 모른다면?

아날로그 세계에서는 글을 읽을 줄 알면 당연히 쓸 수도 있다고 생각한다. 그러나 디지털 세계의 상황은 그렇지 않다. 디지털 원어민을 대상으로 하는 오늘날의 인문학 교육 현장에서는 ‘디지털로 읽기’와 ‘디지털로 쓰기’가 극단적으로 불균형적이다. 학생들이 지식과 정보를 얻는 소스는 이미 대부분 디지털 미디어이고, 거기에서 문자에만 국한되지 않는 다양한 형태의 디지털 데이터를 수용하는데, 그 디지털 입력 자원을 소화해서 디지털 세상에서 통하는 자기식의 표현을 해내는 능력은 디지털로 읽는 것의 양과 수준에 비해 턱없이 부족하다. 디지털 미디어에 콘텐츠를 담아 넣는 제작 기술이 부족하다는 뜻이 아니다. 읽고 생각한 것을 디지털 콘텐츠로 전환하는 방법을 모르거나 서툴다는 것이다.

‘읽기’만 할 뿐 그 읽는 것을 활용한 ‘쓰기’를 할 수 없다면, ‘읽기’를 통해 얻은 것도 피상적이거나 부분적이거나 부정확한 것에 머물기 쉽다. 입수한 정보를 정리하고, 확인하고, 검증하고, 보충하는 것, 다시 말해 읽은 것을 확실한 나의 지식으로 만드는 일은 ‘읽기’와 ‘쓰기’가 단절 없이 이어지고 순환되는 과정 속에서 이루어지는 일이다.

인문학 교육은 학생들에게 인문학적 지식을 일방적으로 제공하는 것이 아니라, 그 지식을 재료와 실마리로 삼아 자신의 지적 세계를 확장하고 심화시킬 수 있게 하려는 것이다. 아날로그 시대의 글쓰기는 그런 능력을 신장시키는 훈련이었다. 인문학 공부로서의 글쓰기는 단지 문장을 만들어 내는 것이 아니다. ‘책 읽기’의 과정에서 만난 무수한 지식의 요소 중에 자기에게 의미 있는 것을 추려내고 자기가 이해한 문맥으로 재구성해서 자기식의 이야기로 만들어 내는 훈련이었다. 인문학 공부의 요체는 사실상 이런 식의 ‘글 쓰기’이다. ‘글 읽기’는 ‘글 쓰기’의 길을 열어 주는 선행 과정일 수 있다. 그러나 정말 열의 있고 깊이 있는 ‘글 읽기’의 동기는 ‘글 쓰기’의 과정 속에서 자신의 이야기를 더 의미 있게 만들어 보고자 하는 데서 촉발된다.

디지털 원어민이 배워야 할 ‘디지털로 글쓰기’도 이와 유사한 효용성을 가질 것이다. 학구적인 주제를 가지고 디지털 데이터를 엮어내는 일을 하다 보면, 새로운 정보를 추가해서 이야기를 확장할 필요를 느끼거나 피상적인 것을 더 구체적으로, 모호한 것을 더 정확하게 하고 싶다는 생각이 든다. 지금 하고 있는 일이 자신이 속한 디지털 세상에서 내 이야기로 나를 표현하는 것이기 때문에 그와 같은 ‘공부’의 동기가 더 강하게 유발될 수 있다.²

2. 디지털 큐레이션

디지털 원어민이 문해력 증진을 위해 배워야 할 ‘디지털로 글 쓰기’는 문자로 쓰인 텍스트를 만들어 내는 일만을 의미하지 않는다. 그것은 문자 텍스트뿐 아니라 도면이나 영상, 전자지도, 3차원 모델링과 같은 자원들을 조합해서 어떤 대상이 묘사되거나 설명될 수 있게 하는 것이다. ‘디지털로 글 쓰기’는 다양한 멀티미디어 데이터들이 이야기의 문맥을 형성하도록 조직화하는 것이다. 아날로그 세계에서 이루어지던 일 가운데 이것과 가장 유사한 일을 꼽는다면, 그것은 바로 박물관에서 이루어지던 ‘큐레이션’이라고 할 수 있다.

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 것도 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 즐기기도 하는 다른 이들과 함께 공유하고 공감하는 것을 더 의미 있다고 생각하는 것은 인간의 모든 문화적 활동의 저변에 놓인 사고이다. 큐레이터의 큐레이션이나 인문학자들의 저술 활동은 모두 자신

이 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 구성요소가 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 차별화했을 뿐이다.

아날로그 시대에는 그 두 가지를 전혀 다른 일로 취급할 만한 이유가 있었다. 박물관의 수장과 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 연구 방법이 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 강의실로 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 것을 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 제작 기법이나 역사적 배경을 설명하는 연구논문이 다 같이 디지털 데이터로 존재하는 곳에서 두 세계를 넘나드는 행위는 한번의 마우스 클릭이면 족하다.

디지털 인문학계에서 논의되는 ‘근현대 인문학의 과오에 대한 반성’ 가운데 하나는 우리가 그동안 아카이브와 대학의 연구실을 둘로 나누는 잘못을 저질러 왔다는 것이다. 2009년 미국에서 발표된 「디지털 인문학 선언 2.0」은 전통적인 대학의 연구실, 도서관, 박물관이 모두 하나로 통섭되는 세계가 미래 인문학의 무대가 될 것으로 내다보고, 그곳에서 이루어지는 융합적인 교육·연구 활동을 새로운 ‘큐레이션’으로 인식하였다. 이에 대한 이해를 돕기 위해 선언문의 일부를 인용한다.

확장된 학술 활동으로서의 큐레이션

디지털 인문학자들은 큐레이션을 미래 인문학의 중심적인 특징으로 인식한다.

근대 대학이 큐레이션을 이차적 보조적 역할로 강등시키고 큐레이터를 박물관, 아카이브, 도서관으로 추방하면서 학문에서 분리시켰던 반면, 디지털 인문학 혁명은 연구와 교육의 지형도를 근본적으로 재형성하도록 촉진한다. 디지털 인문학은 **학자를 큐레이터로, 큐레이터를 학자로 재구성**하고, 이로 인한 가능성과 수요의 확장을 통해 학문 활동에 다시 활기를 불어넣으며 박물관, 도서관, 아카이브의 학술적 사명을 재개하려 한다.

.....

큐레이션은 강화된 학술 연구 활동이다. 그것은 또한 가르치는 것과 배우는 것도 강도 높게 증진시킨다.

큐레이션의 방법으로, 미래 세대의 인문학자들은 문화와 역사 연구의 가장 기본적인 원자재를 가지고 연구를 시작할 것이며, 실제 실험실과도 같은 환경에서 전문가인 연구자들의 지도하에 지

식을 수집하고 생산하는 일에 직접 참여하게 될 것이다. 큐레이션 작업이 학술 연구 “성과”의 범위에 추가됨으로써 인문학 연구 세계는 광대하게 풍요로워진다.

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

종래 아날로그 세계에서 행하던 독립적, 자기완결적 큐레이션을 디지털 미디어상에서 하는 일을 가지고 ‘디지털 큐레이션’이라고 하는 이들도 있다. 물론 박물관 안에서의 아날로그적 큐레이션에 상응하는 ‘디지털 전시실 꾸미기’ 같은 일도 ‘디지털 큐레이션’의 범주에 포함시킬 수 있다. 그러나 디지털 인문학의 입장에서 이해하는 ‘디지털 큐레이션’은 보다 넓은 지평에서 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 콘텐츠로 활용될 수 있게 된다. 그것은 다시 개방적인 클라우드 환경에서 새로운 큐레이션을 유도할 것이고 유의미한 지식의 확장을 촉진할 것이다.

디지털 원어민이 디지털 환경에서 인문학적 지식을 탐구하고, 거기서 얻은 정보를 새로운 체계로 엮어서 자신의 이야기로 표현하는 행위는 더 이상 예전처럼 문헌 중심의 연구와 실물 중심의 큐레이션으로 구획할 필요가 없을지도 모른다. 주안점을 어디에 두느냐의 차이는 있을지라도 그 두 가지가 서로 소통하고 융합할 수 있는 길이 디지털 세계에서는 항상 열려 있다.

이러한 맥락에서 나는 앞에서 이야기한 ‘디지털로 글쓰기’ 역시 디지털 환경에서 실물과 문헌의 경계를 넘어서는 통섭적인 일이며, 그래서 이것을 ‘디지털 큐레이션’이라 할 충분한 이유가 있다고 생각한다. 책 속의 문장이든 유튜브의 영상이든 박물관 소장 유물의 3D 모델링 데이터이든 그것이 내가 얻은 정보의 원천자원이고, 내가 전달하고자 하는 지식의 증거물이라면 그 모든 것이 나의 이야기 속에서 체계적으로 엮이게 할 필요가 있다. 디지털 원어민이 디지털 데이터의 세계에서 지식의 자원을 탐구하고 자신이 발견한 것을 다른 이들도 이해하고 활용할 수 있도록 디지털 콘텐츠로 표현하는 행위를 ‘디지털 큐레이션’이라고 한다.

3. 디지털 아카이브

디지털 큐레이션의 결과물은 여러 가지 형태일 수 있다. 박물관의 전시실과 유사한 모습의 가상전시관을 꾸밀 수도 있고, 초등학생들의 지역화 수업 교재로 쓸 수 있는 ‘내 고장 이야기’ 디지털 그림책을 만들 수도 있다. 그러나 그와 같은 최종소비재 타입의 디지털 콘텐츠보다, ‘디지털

인문학'의 교육이나 연구의 관점에서 더 우선적으로 고려할 수 있는 디지털 큐레이션의 실천 영역은 '디지털 아카이브(Digital Archives)'일 것으로 생각한다. 인문학은 최종적인 결론이나 답을 정하는 일이 아니고 의미 있는 것을 부단히 찾아가며 새로운 사실의 발견과 확인을 통해 예전에 알았던 부분적인 지식을 확장해 가는 '상시 진행형'인 노력이다. 디지털 아카이브는 완성품인 고정형 저작물이 아니고, 끊임없이 확장하고 보완할 수 있는 개방형 저작물이기 때문에 '인문학 공부'라고 하는 과정의 동반자로 활용하기에 적합하다. 여기서 말하는 디지털 아카이브는 원시 자료의 보존을 목적으로 하는 실물 아카이브의 디지털 사본(디지털화의 결과물)이 아니라, 큐레이션의 과정을 거쳐 만들어지는 '지식과 자료의 융합적 아카이브'이다.

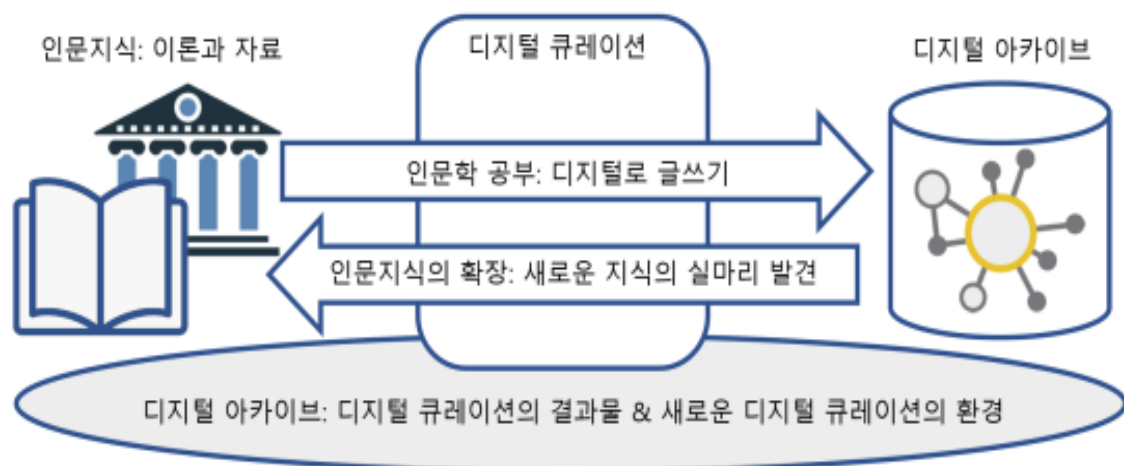


그림 1. 디지털 큐레이션과 디지털 아카이브

일반적으로, 사실의 증거 기능이 있는 기록물의 저장소를 '아카이브(Archives)'라고 한다. 실물 아카이브의 디지털 버전을 '디지털 아카이브'로 이해하는 관점에 서면, 이때의 디지털 아카이브는 '실물 아카이브에 있는 원본의 디지털 사본을 저장한 데이터베이스와 그 데이터를 서비스하는 정보 시스템' 정도로 이해할 수 있을 것이다. 그러나 앞에서 이야기한 '디지털 큐레이션'의 개념이 박물관에서 하던 아날로그 큐레이션의 디지털 버전에 국한된 것이 아니었듯이, 디지털 아카이브의 함의 역시 실물 아카이브의 디지털화(Digitalization)로만 한정할 필요가 없다.⁶

'디지털 큐레이션(Digital Curation)' 행위에 의해 만들어지는 디지털 저작물이 1회적인 이벤트로만 쓰이고 사라진다가 아니라 자기완결적인 최종결과물로 고정되는 것이 아니라, 다른 이들로 하여금 새로운 저작물을 만들 수 있게 하는 재료로 쓰일 수 있게 된다면 그것은 '디지털 아카이브'(Digital Archives)라고 불릴 수 있다. '큐레이션'이 있느냐를 디지털 아카이브의 인정 요건으로 보는 이유는 그것 없이 쌓여 있기만 한 디지털 데이터는 활용할 수도, 공유할 수도 없기 때문이다.

인문학계가 아카이브와 큐레이션을 소홀히 했던 것과 마찬가지로 예전의 아카이브들은 보유 자료들을 인문학적 지식의 세계와 연결짓는 일에 적극적이지 않았다. 소장하고 있는 물품이나 기록물에 대한 해석을 유보한 채, 원천 자료를 그대로 정리·보존하여 제공하는, '연구 지원' 역할을 수행하고, 그 자료에 대한 해석이나 의미 부여는 대학의 연구실에서 '학술 연구'의 행위로 수행한다는 입장에 머물렀던 것이다. 이러한 사고의 연장선상에서 '디지털 아카이브'도 한동안 '해석되지 않은 원천 자료의 디지털 사본'의 형태에 고집스럽게 머물고 있는 듯한 양상을 보였다. 하지만

‘디지털 아카이브가’ ‘실물 아카이브’의 ‘디지털 사본’에 머물러서는 안된다고 하는 것은 아카이브계의 종사자들과 이용자들이 이미 공감하는 사고이다. 원래의 실물 아카이브는 증거력을 가진 ‘원본’의 보존이 가장 중요한 존재 이유이지만, 복제물의 서비스에 불과한 ‘디지털 아카이브’는 태생적으로 ‘보존’보다는 ‘활용’ 가치의 증대를 추구할 수밖에 없는 것이다.⁷

디지털 아카이브가 ‘원본의 형상’을 보여 주는 데 머물기보다는 그것이 왜 ‘보존’하고 ‘기억할’ 가치가 있는 것인지 이유와 문맥을 설명할 수 있을 때 그것에 관심을 갖는 고객의 범위와 활용성이 더 커질 것임은 당연하다. 디지털 아카이브의 세계에서 큐레이션을 강화해야 할 필요성은 이미 인정되고 있지만, 실천이 더딘 이유는 자료의 학술적 문맥에 대한 이해와 데이터 조작 능력을 함께 갖춘 디지털 큐레이터가 아직은 턱없이 부족하기 때문이다.⁸ 디지털 인문학 교육 과정에서 디지털 큐레이션의 실천을 강조하는 이유는 일차적으로 그것이 인문지식을 체계적으로 탐구하고 정리하는 공부의 방법이기 때문이지만, 그 훈련을 받은 인문학도가 ‘지식과 자료의 융합적 집성체’인 미래 디지털 아카이브의 큐레이터로 활동할 수 있게 되는 것도 그 교육의 효과로 기대할 수 있는 일이다.

4. 백과사전적 아카이브

디지털 인문학 교육·연구 과정에서 만들어지는 디지털 아카이브는 백과사전적 지식정보와 그것의 증거가 되는 문헌·실물 자료의 집성체 성격을 갖는다. 나는 이러한 성격의 디지털 아카이브에 대해 ‘백과사전적 아카이브(Encyclopedic Archives, Encyves)’라는 이름을 부여했다.⁹ ‘백과사전적 아카이브’란 아날로그 세계에서 두 개의 독립된 영역에 존재했던 백과사전과 아카이브 속의 정보와 자료들이 상호간의 의미적 관련성을 좇아 긴밀하게 엮여 있는 디지털 데이터 아카이브이다.¹⁰

먼저 여기서 말하는 백과사전은 ‘세상의 모든 영역’의 지식을 망라적으로 다루는 ‘일반 백과사전’이 아니라 ‘특정 분야’의 지식을 포괄적으로 다루는 ‘전문 백과사전’의 의미임을 전제한다. 인문학 공부는 항상 주제가 있기 마련이다. 그 공부의 주제가 ‘특정 분야’일 수 있다. 일정한 주제를 가지고 인문학 공부를 하는 사람들이 이루어 낸 인문 지식 큐레이션의 결과물은 그 주제의 ‘백과사전적 아카이브’의 성격을 갖게 될 것이다.

조선왕실의 의례를 주제로 한 백과사전적 아카이브를 예시로 들어 설명하겠다. 다음의 예시(문헌 자료 디지털 아카이브 예시 1, 2, 3)는 한국학중앙연구원 장서각 아카이브, 서울대 규장각 아카이브, 국립고궁박물관 등에서 소장하고 있는 조선 왕실 의례 가운데 ‘왕실 잔치’의 기록이라고 할 수 있는 문헌들을 볼 수 있게 하는 디지털 아카이브 서비스 화면들이다.

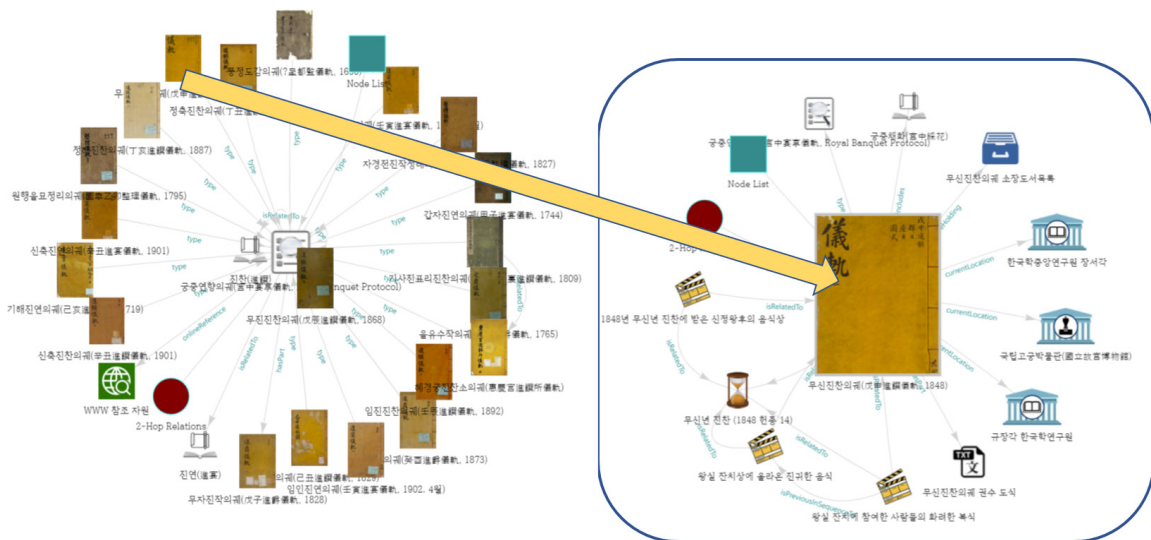
그림 2. 문헌자료 디지털 아카이브 예시 1: 조선왕실 궁중연향의례 ⇒ 무신진찬의례(1848)¹¹

그림 3. 문헌자료 디지털 아카이브 예시 2: 특정 의결 ⇒ Archive Holdings ⇒ 소장 자료 서지사항¹²



그림 4. 문헌자료 디지털 아카이브 예시 3: 소장 자료 서지사항 ⇒ 원문 텍스트 / 원문 이미지¹³

온라인상에서 이러한 고도서의 디지털 사본을 볼 수 있게 하는 것은 기존의 물리적인 아카이브에서 하던 일을 디지털 환경에서 수행하는 것이다. 그런데 그 고문헌의 온라인 서비스가 문헌자료의 제공에 그치지 않고, 그것이 어떤 행사에 대한 기록인지 그 행사는 어느 곳에서 일어났으며, 참석자는 누구인지, 그곳에서 어떠한 음악과 무용이 공연되었고, 어떤 악기와 도구가 사용되었으며, 어떤 음식이 제공되었는지를 알게 하는 백과사전적 지식정보를 함께 제공하는 서비스를 생각할 수 있다. (지식정보 데이터 예시 1, 2)

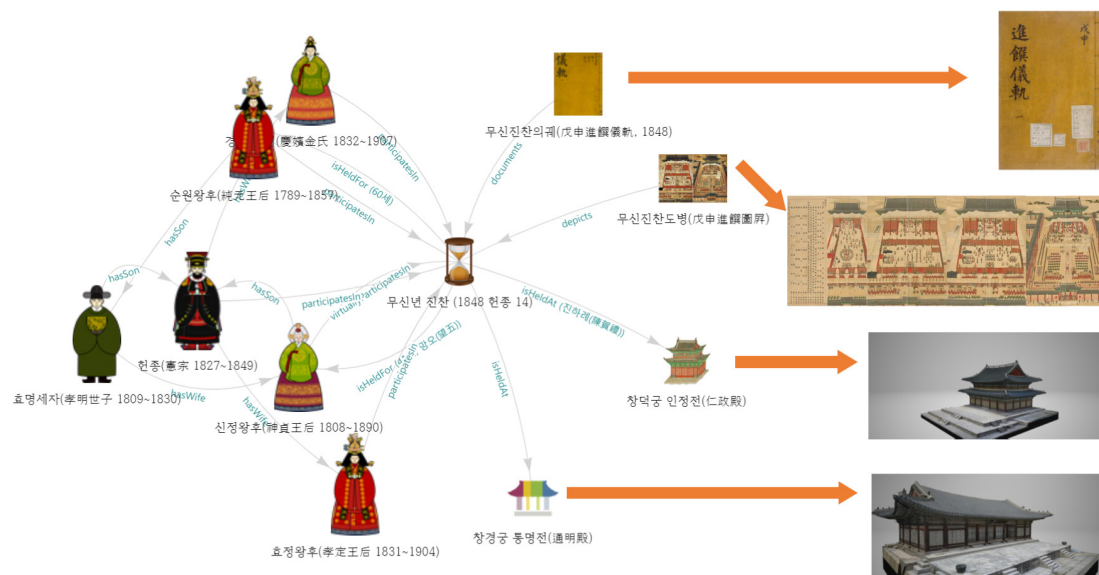


그림 5. 지식정보 데이터 예시 1: 무신년진찬(1848)¹⁴ ⇒ 행사/장소/인물/기록물 등¹⁵

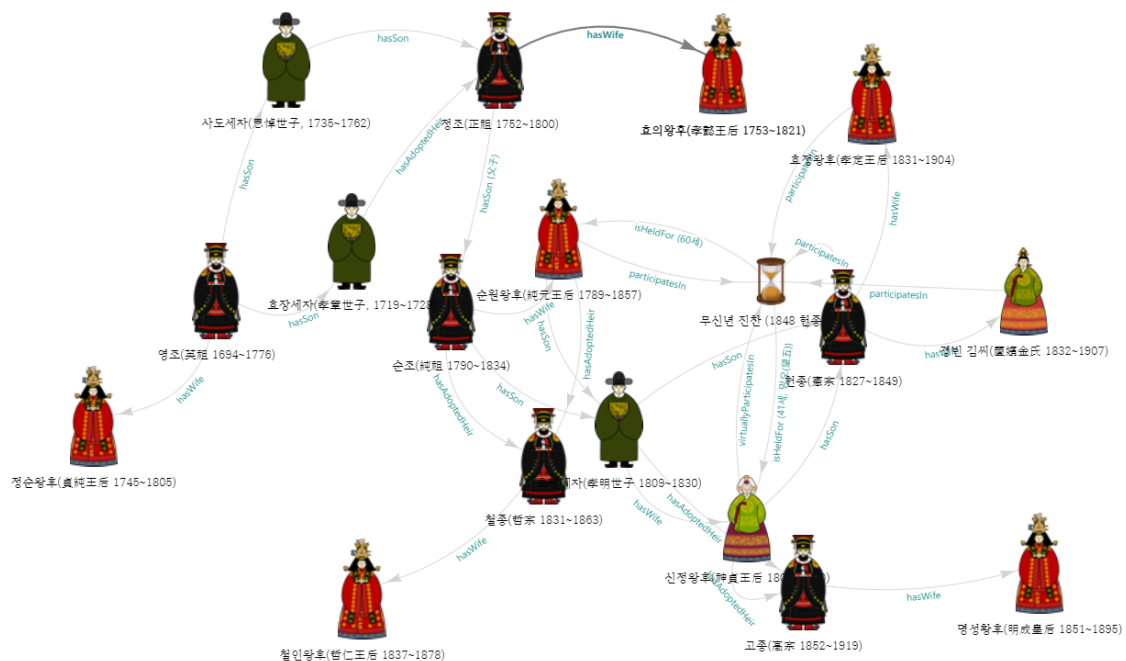


그림 6. 지식정보 데이터 예시 2: 무신년진찬 ⇒ 행사 참석자 중심 왕실 계보¹⁶

위의 예시에서 네트워크 그래프로 표현된 여러 정보의 노드들 가운데 문헌 자료의 원문을 선택하면, 해당 원문의 디지털 이미지나 문자화된 원문 텍스트를 볼 수 있고, 사건(행사), 장소, 인물, 물품 등 그 문헌에 기록된 내용과 관련된 노드들을 클릭하면 그것을 설명하는 백과사전적 데이터가 제공된다. 종래의 ‘디지털 아카이브’와 ‘디지털 백과사전’의 기능을 하나로 묶었다고 할 수 있는 서비스이다.

이것이 가능하기 위해서는 누군가 이 방면의 지식을 확보한 사람이 그 지식의 문맥을 데이터로 표현하고, 신뢰할 수 있는 인터넷상의 자원들을 조사하여 데이터 네트워크상에서 바로 접근할 수 있게 해야 한다. 이 작업이 바로 인문 지식의 디지털 큐레이션이며, 그러한 큐레이션의 결과물이 ‘백과사전적 아카이브’이다.

디지털 인문학 교육이나 연구 프로젝트를 통해서 이루어지는 인문 지식 큐레이션은 학구적인 지식 탐구의 노력 위에서 이루어지기 때문에 대체로 한꺼번에 양적으로 많은 데이터를 생산하거나, 관련 지식을 망라적으로 포괄하지는 않는다. 그래서 그 결과물로 만들어지는 디지털 콘텐츠도 망라적인 빅 데이터(Big Data)가 아니라, 주제와 이야기가 있는 스몰 데이터(Small Data)이다. 그럼에도 불구하고 그것을 ‘백과사전적 아카이브’라고 부르는 이유는 그 내용물이 어느 한 사람의 연구자나 하나의 연구팀에 의해서만 만들어지는 것이 아니라, 여러 사람이 수행한 유사한 주제 또는 관련이 있는 주제의 큐레이션이 하나의 데이터 저장소에 집적될 수 있고, 그래서 지속적인 확장이 가능하기 때문이다. 물론, 이와 같은 협업과 확장의 가능성이 보장되기 위해서는 아카이브의 기획자들에 의해 그 분야에서 지속가능한 디지털 큐레이션의 방법론이 제시되어야 하고, 아카이브 콘텐츠의 확장에 기여할 큐레이터들은 그 방법론상의 준칙들을 이해하고 준수하는 노력을 기울여야 한다.

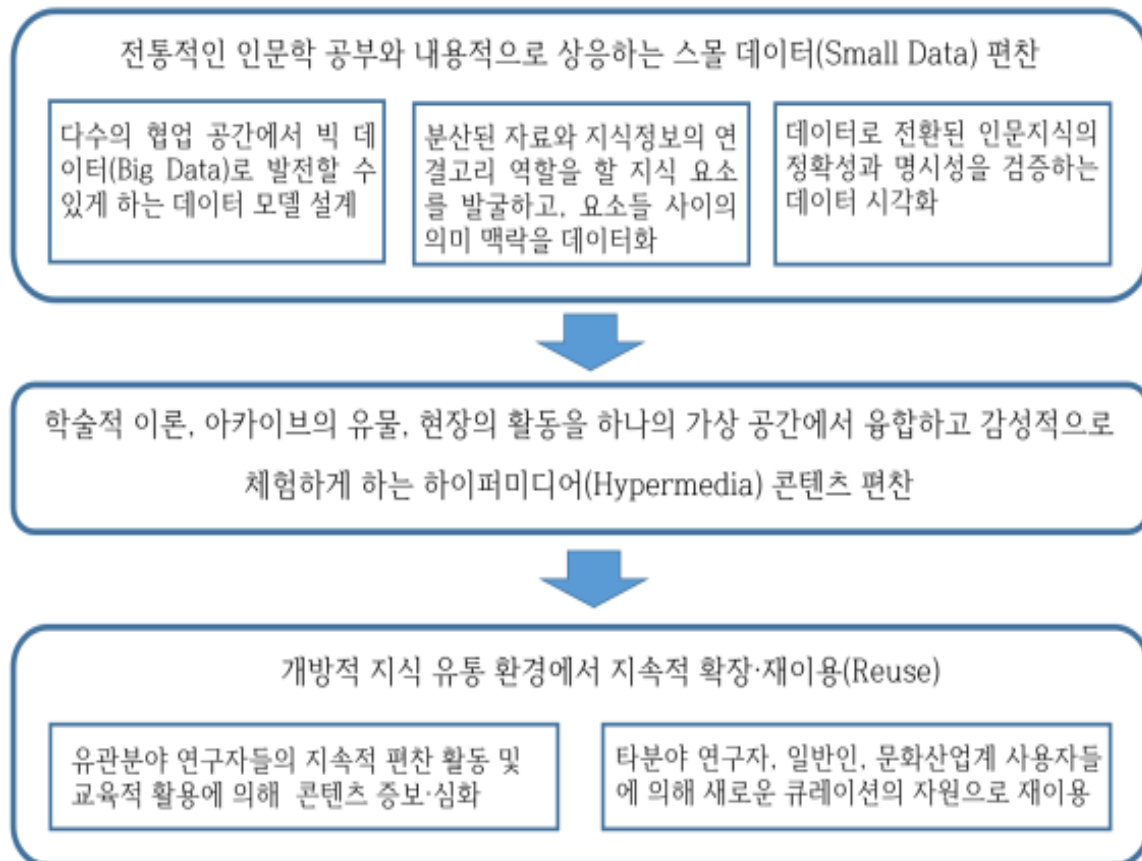


그림 7. 백과사전적 아카이브를 구현하는 ‘디지털 큐레이션’ 수행 방법과 전개 과정

백과사전적 아카이브는 그것이 개인의 저작물이 아니라 공동의 관심사를 가지고 있는 다수의 인문학도가 소통과 협업의 노력으로 만들어 가는 것이라는 점에서, 디지털 인문학 교육·연구의 중요한 성과물로 인정될 수 있다. 디지털이기 때문에 가능해진 하이퍼미디어¹⁷ 세계는 아날로그 세계에서 어렵고 불편했던 소통과 협업을 촉진하는 환경이며, 디지털 인문학은 그 새로운 환경에서 개인과 조직, 분야의 벽을 넘어서는 학제적 통섭을 지향하기 때문이다.

5. 시맨틱 데이터 아카이브

앞에서 언급한 ‘백과사전적 아카이브’가 우리가 지향하는 인문 지식 디지털 아카이브의 성격과 기능을 대표하는 말이라면, ‘시맨틱 데이터 아카이브’는 그 아카이브에 담기는 데이터의 형식에 주목하는 표현이다.

‘시맨틱 데이터(Semantic Data)’란 독립된 데이터의 나열이 아니라, 데이터 사이의 연결 방식을 표현함으로써 데이터에 담긴 의미를 명시적으로 설명할 수 있도록 구성된 정보 기술 형식을 말한다. 시맨틱 데이터의 형태는 일반적으로 대상(사람, 장소, 사건, 사물 등)에 대한 정의와 대상들 사이의 관계를 포함하며, 이들이 서로 연결되거나 상호 작용하는 방식을 명확하게 설명한다.

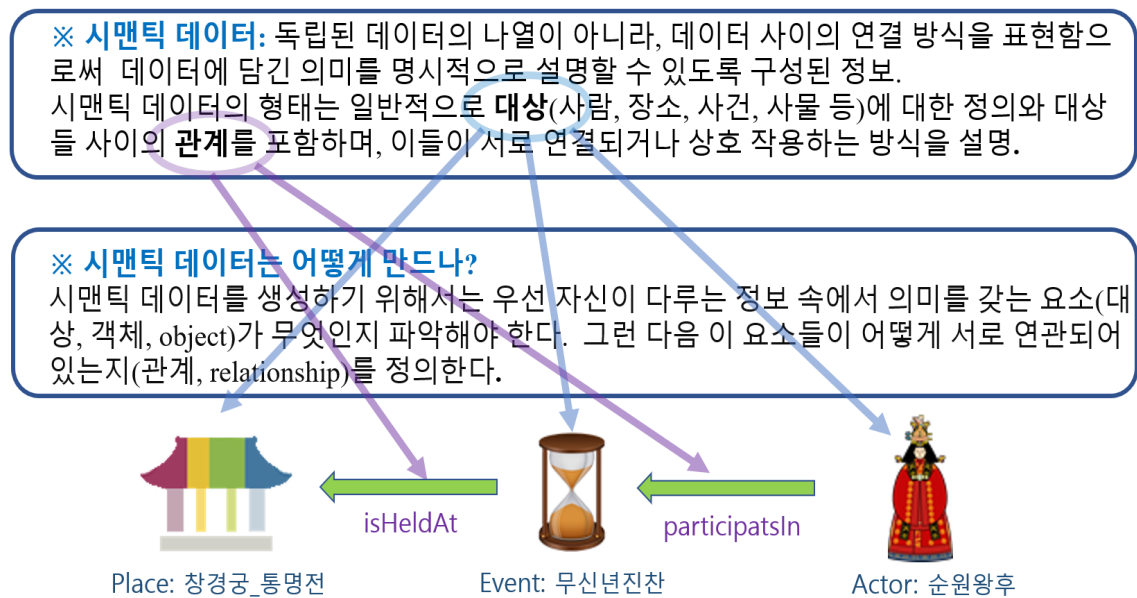


그림 8. 시맨틱 데이터(Semantic Data)

백과사전적 아카이브는 단순히 인문학 연구의 자료만 집적한 저장소가 아니고, 자료와 지식의 관계, 더 나아가 수많은 지식의 요소들 사이의 관계를 디지털 데이터로 기록한 연구 저작물이다. 백과사전적 아카이브의 기술적 요건은 여러 가지로 이야기될 수 있겠지만, 가장 중요한 것은 바로 그 ‘관계’의 데이터화가 합리적, 효과적으로 이루어져야 한다는 것이다. ‘시맨틱 데이터’는 지식 요소들 사이의 ‘관계성’을 표현하는 데이터이며, ‘시맨틱 데이터 아카이브’는 그와 같은 시맨틱 데이터의 집합, 다시 말해 특정 분야의 다양한 정보와 멀티미디어 어셋¹⁸들이 시맨틱 데이터의 형식으로 결집된 네트워크 구조의 지식정보 아카이브이다.

백과사전적 아카이브를 편찬하는 ‘디지털 큐레이션’에 대해서도 사실상 그 행위의 기술적인 방법론은 ‘시맨틱 데이터’의 편찬이 중심이 된다고 할 수 있다. 결국 미래 인문학을 위한 준비로서 학생들에게 디지털 큐레이션을 가르치는 일은 시맨틱 데이터를 만들 수 있는 능력을 배양하는 일이라고 해도 과언이 아니다.

시맨틱 데이터 아카이브의 구축이 인문 지식 디지털 큐레이션의 실천이라는 전제하에, 이 일을 하기 위해 배워야 할 도구와 방법은 다음과 같다.

- 첫째, 큐레이션의 내용을 저장하여 아카이브로 쓰일 수 있게 하는 소프트웨어인 데이터베이스¹⁹
- 둘째, 큐레이션의 대상이 되는 지식 요소들 사이의 관계를 개념적으로 설계하는 시맨틱 모델링²⁰,
- 셋째, 시맨틱 모델링을 통해 만들어지는 가상 세계의 설계도인 온톨로지²¹,
- 넷째, 온톨로지의 설계 내용에 따라 대상 세계의 사실과 문맥을 디지털 데이터로 전환하는 시맨틱 데이터 프로세싱²²

시맨틱 데이터 아카이브를 만드는 데 필요한 기술적 과제, 다시 말해 인문 지식의 디지털 큐레이션을 가능하게 해 주는 이 방법들은 실은 디지털 시대에 인문지식을 공부하고 응용하기 위해서라면 반드시 배워야 할 도구적 지식이다. 마치 동아시아의 고전을 배우는 사람들에게 ‘한문(漢

文)’이라는 특수한 언어가 지식 소통의 수단이듯이 데이터베이스나 시맨틱 모델링, 온톨로지와 같은 기술은 디지털 세계에서 인문 지식에 다가가는 소통의 언어 역할을 할 것이기 때문이다.

시맨틱 데이터 아카이브 구축의 세부 주제들은 디지털 인문학 교육 커리큘럼이나 교재 안에서 다루어질 내용이므로 이 글에서는 구체적으로 언급하지 않는 대신, 그 교육 과정을 이수한 학생들이 만든 데이터를 예시로 보여 이 주제에 대한 대체적인 이해를 돕도록 하겠다.

앞에서 백과사전적 아카이브의 예시로 보였던 ‘무신년진찬(戊申年進饌)’ 콘텐츠는 바로 시맨틱 데이터의 구조로 만들어진 디지털 저작물이다. 이것은 ‘한양도성 타임머신’ 프로젝트²³의 일환으로 편찬된 시맨틱 데이터인데, 조선시대에 왕실에서 열린 여러 차례의 성대한 연회 중 하나인 1848년 무신년의 궁중연향에 관한 데이터 일부를 보인 것이다. 한양도성 타임머신 시맨틱 데이터 아카이브²⁴는 조선왕실 의궤(儀軌)²⁵가 전하는 여러 차례의 궁중연향에 관한 데이터와 디지털 어셋을 수록하고 있다. 이 주제의 데이터 편찬 프로세스와 시맨틱 모델링 사례 몇 가지를 ‘시맨틱 데이터 아카이브 편찬’의 예시로 보이도록 하겠다.

○ 자료 조사: 시맨틱 데이터 아카이브 편찬을 위한 조사 연구 사례

이 데이터를 편찬하는 작업은 조선왕실의 궁중연향에 관한 기록물들을 조사하는 일에서부터 시작하였다. 편찬 과정 중에 적지 않은 시간이 학생들과 함께 조선왕실의궤를 읽는 일에 소요되었는데, 텍스트를 해석하는 데 그치지 않고 텍스트 속에서 데이터화 할 요소를 찾아내고 그것이 다른 요소들과 어떤 관계에 있는지 조사한 일을 병행하였다. 아래의 그림은 ‘무신년진찬’ 데이터 프로세싱의 중심 텍스트인 『무신진찬의궤(戊申進饌儀軌, 1848, 헌종14)』의 내용 속에서 시맨틱 데이터의 요소들을 발굴하는 과정을 예시로 보인 것이다.



그림 9. 문헌 자료로부터 시맨틱 데이터 요소 발굴

○ 데이터 요소의 범주(클래스) 설계 예시

문헌자료, 박물관 소장 유물, 웹 자원 등 기초자원 조사를 통해 발굴된 데이터 요소들에 대해서는 다른 요소들과 구별할 수 있는 식별자를 부여하고 그 성격에 따라 특정 범주에 소속되게 한다. 이 일을 시맨틱 모델링의 한 과정으로서 ‘클래스(Class) 설계’라고 한다. 무신년진찬 등 궁

궁중연향과 관련이 있는 수많은 데이터 요소들은 다음과 같은 10개의 클래스에 속하는 것으로 정리하였다.



그림 10. 궁중연향 데이터에 적용한 '한양도성 시맨틱 데이터 아카이브' 클래스 설계

표 1. '한양도성 시맨틱 데이터 아카이브' 주요 클래스 설명

class	description
Actor	역사적인 인물. 집단적 행위 주체로서의 단체, 운영 주체로서의 기관
Event	한양도성에서 일어난 사건, 궁중 의례 및 이를 재현하는 행사, 기념제
Place	역사적 배경(사건, 인물)과 관련이 있는 장소. 유물의 소장처
Architecture	한양도성 건조물 및 주변의 유관한 건축/조형물
Clothing	역사적 인물들이 착용했던 다양한 성격의 복장과 복식 요소들
Food	궁중 의례·행사에서 차려졌던 궁중음식, 식재료와 상차림
Object	각종 의례/행사의 도구, 그 시대의 문화를 보이는 물품들
Work	음악, 연희, 무용 등의 공연예술 작품이나 시, 악장가사 등의 문학작품
Record	문헌, 사진, 도면, 금석문 등 지식 데이터의 원천 자료가 되는 기록물
Concept	제도, 의례, 풍속 등을 설명하는 데 필요한 용어와 개념

이어서 보일 4개의 예시는 궁중연향의 내용에 따라 부분적으로 시행한 시맨틱 모델링의 사례이

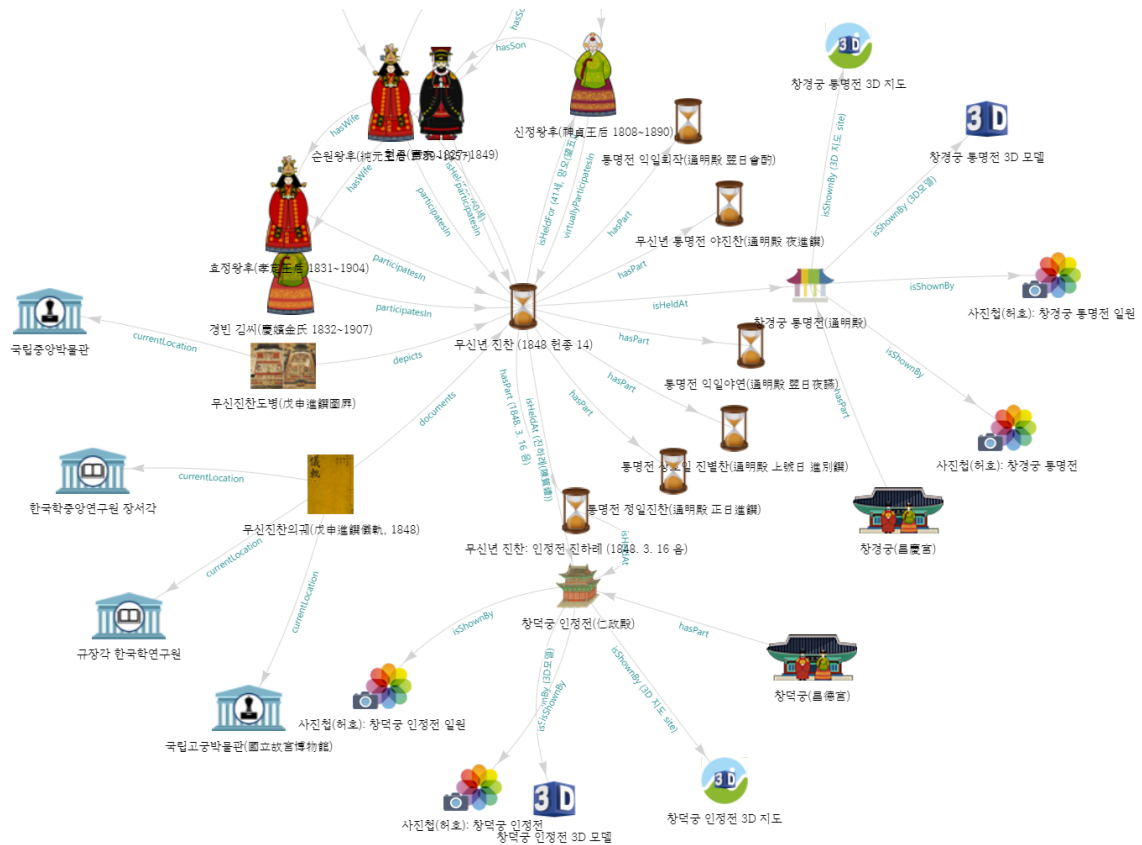
다. 궁중의 연향은 다양한 성격이 일들이 복합적으로 일어난 역사적 이벤트이기 때문에 그것을 데이터화 하는 작업도 여러 가지 관점에서 수행할 필요가 있었다.

○ 시맨틱 모델링 예시 1: 궁중연향의 이벤트별 개요

첫 번째 예시는 궁중연향에 속하는 행사를 개관(概觀)하는 데이터의 모델링이다. 그 행사는 누구를 위해 열렸고, 누가 참석했는지, 장소는 어디이고, 어떤 연회와 공연이 있었는지, 그 행사에 대한 기록물은 무엇이고 소장처는 어디인지 등을 알릴 수 있는 틀(모델)을 설계하고 그 틀에 따라 각각의 궁중연향에 해당되는 데이터를 정리하였다.

표 2. ‘궁중연향’ 개관(概觀) 시맨틱 모델링을 위한 관계 속성 설계

Name space	Relation	Domain	Range	Inverse Relation	Attribute	Description
ekc:	administrates	Actor	Event			~를 주관하다
ekc:	isHeldAt	Event	Place			~에서 개최되었다
ekc:	isHeldFor	Event	Actor			~를 위해 개최되었다
ekc:	isOperatedIn	Agent	Event			(시설이) 이벤트 ~에서 운영되었다
ekc:	participatesIn	Actor	Event	hasParticipant		~에 참여하다
ekc:	virtuallyParticipatesIn	Actor	Event			~에 참여한 것으로 간주하다 ※ 기록물에 참여한 듯이 기록
ekc:	depicts	Object/Record	Any	isDepictedIn		~를 묘사하다
ekc:	documents	Record	Event			~를 문서화하다
edm:	currentLocation	Object	Place/Actor			소재지(소장처)는 ~이다
edm:	isShownBy	Any	WebResource		유형	웹 자원 ~에 보인다
dcterms:	hasPart	Any	Any	isPartOf	수량	부분 ~이 있다

그림 11. 무신년진찬 개관 시맨틱 데이터의 시각화²⁶

○ 시맨틱 모델링 예시 2: 궁중연향에 참석한 인물의 복식

두 번째 예시는 궁중연향에 참석한 왕실의 주요 인물들은 어떤 의상을 입었으며, 그 의상을 구성하는 복식 요소들은 무엇인지 알게 하는 시맨틱 모델링이다. 격식을 갖춘 궁중의상 일습을 중심으로 그것을 착용한 인물, 그 옷을 입고 참석한 행사, 의상의 구성 요소, 의상을 묘사하는 기록물과 소장처 등의 관계를 표현하는 틀을 설계하였다.

표 3. 궁중연향에 참석한 왕실 인물의 복식(服飾) 시맨틱 모델링을 위한 관계 속성 설계

Name space	Relation	Domain	Range	Inverse Relation	Attribute	Description
ekc:	wears	Actor	Clothing			복장 ~를 입다
ekc:	isWornIn	Clothing	Event/Concept			(복장을) 행사 ~에서 입다
ekc:	isHeldFor	Event	Actor			~를 위해 개최되었다
ekc:	participatesIn	Actor	Event	hasParticipant		~에 참여하다
ekc:	virtuallyParticipatesIn	Actor	Event			~에 참여한 것으로 간주하다 ※ 기록물에 참여한 듯이 기록
ekc:	depicts	Object/Record	Any	isDepictedIn		~를 묘사하다
edm:	currentLocation	Object	Place/Agent			소재지(소장처)는 ~이다
edm:	isShownBy	Any	WebResource		미디어	웹 자원 ~에 보인다
dcterms:	hasPart	Any	Any	isPartOf	수량	부분 ~이 있다



그림 12. 무신년친찬에 참석한 대비(순원왕후)의 복식 시맨틱 데이터의 시각화²⁷

○ 시맨틱 모델링 예시 3: 궁중연향에서 제공된 음식

세 번째 예시는 궁중연향에서 제공된 음식에 관한 시맨틱 모델링이다. 연회별로 누구에게 어떤 음식상이 제공되었고, 그 음식의 식재료는 무엇이며, 어디에서 조리하였는지를 알게 하는 틀을 설계하였다.

표 4. 궁중연향서 제공된 음식(飮食) 시맨틱 모델링을 위한 관계 속성 설계

Name space	Relation	Domain	Range	Inverse Relation	Attribute	Description
ekc:	serves	Actor	Food			음식 ~을 차린다
ekc:	cooks	Agent	Food			음식 ~을 조리한다
ekc:	prepares	Agent	Food			음식, 식재료 ~을 준비한다
ekc:	hasIngredient	Food	Food			식재료 ~를 쓴다
ekc:	isGarnishedWith	Food	Object			상화 ~로 장식하다
ekc:	isProvidedIn	Food	Event			(상차림이) 이벤트 ~에서 제공되다
ekc:	isServedIn	Food	Object			(음식이) 그릇 ~에 담겨 제공되다
ekc:	isServedOn	Food	Object			(음식이) 음식상 ~에 차려지다(음식)
ekc:	isServedTo	Food	Actor			인물 ~에게 제공되다

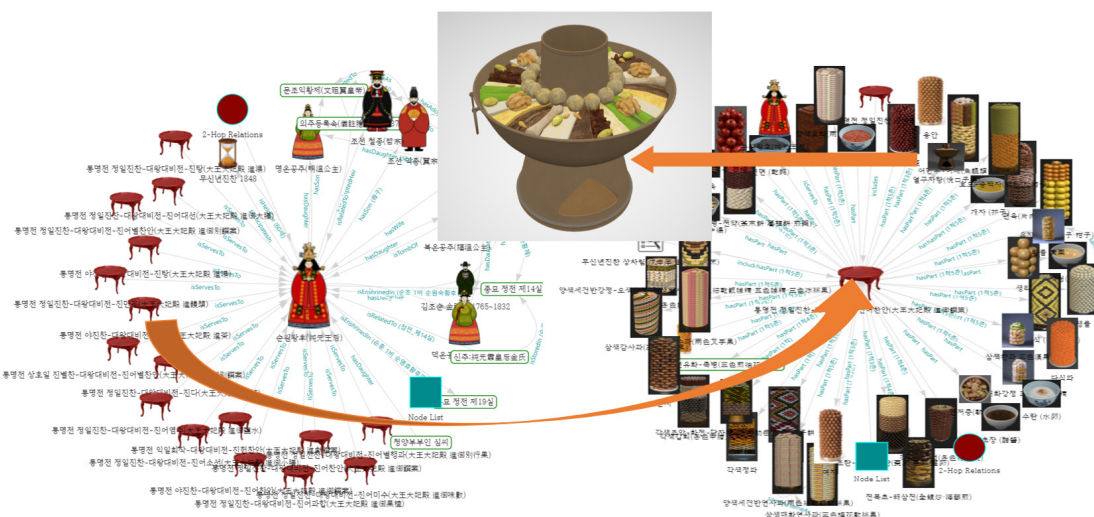


그림 13. 무신년진찬에서 대비(순원왕후)를 위해 차려진 음식상 시맨틱 데이터의 시각화²⁸

○ 시맨틱 모델링 예시 4: 궁중연향의 정재(呈才) 공연

네 번째 예시는 궁중연향에서 다양하여 펼쳐졌던 정재 공연의 시맨틱 모델링이다. ‘재주를 남에게 보인다’는 뜻의 정재(呈才)는 악(樂)·가(歌)·무(舞)의 종합예술을 뜻한다. 궁중연향의 각 연회에

서는 어떤 정재 작품이 공연되었고, 하나의 정재 작품 안에서는 어떤 춤과 노래가 있었으며, 어느 재인이 어떤 복장으로 무슨 역할을 했고, 어떤 무용도구가 사용되었는지를 설명하는 틀을 설계하였다.

표 5. 궁중연향서 제공된 음식(飮食) 시맨틱 모델링을 위한 관계 속성 설계

Name space	Relation	Domain	Range	Inverse Relation	Attribute	Description
ekc:	appearsIn	Actor	Concept			작품, 공연 ~에 출연하다
ekc:	hasPerformance	Event	Concept	isPerformedIn		(이벤트에) 공연 ~이 있다
ekc:	isAwarded	Actor	Object		수량	~을 상으로 받다
ekc:	isPerformedAt	Event/Concept	Place			(공연/의식이) ~에서 열렸다
ekc:	isPerformedBy	Concept	Actor			~가 공연했다
ekc:	isPerformedIn	Concept	Event	hasPerformance		(공연이) 이벤트 ~에서 공연되었다
ekc:	plays	Actor	Object			악기 ~를 연주하다
ekc:	isUsedIn	Object	Event			~에 쓰인다
ekc:	goesWith	Object	Object			~를 수반하다
ekc:	wears	Actor	Clothing			복장 ~를 입다
ekc:	isWornIn	Clothing	Event/Concept			(복장을) 행사 ~에서 입다
dcterms:	creator	Record/Object	Actor			(문헌/작품의) 창작자는 ~이다
ekc:	isInfluencedBy	Any	Any			~에 의해 영향받다

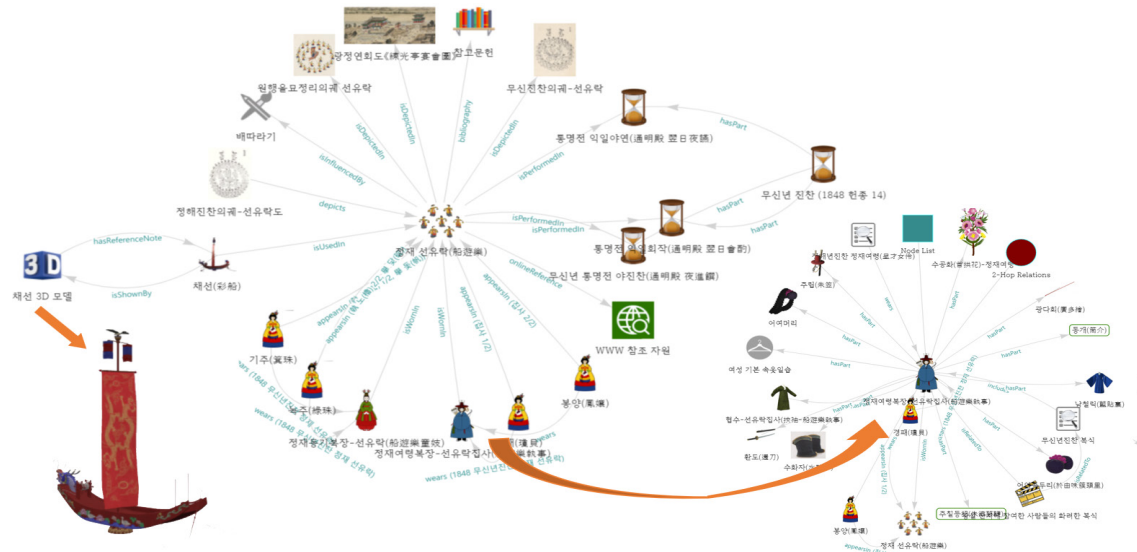


그림 14. 정재 작품 「선유락(船遊樂)」의 시맨틱 데이터의 시각화²⁹

위에서 보인 예시들은 역사, 예술, 문학, 민속학 등을 전공하는 인문학도들이 시맨틱 데이터 아카이브를 만드는 과정에서 직접 수행한 일들이다. 아카이브의 편찬 과정에서 연구자 개개인은 그들의 전공과 관심에 따라 특정 분야에 대해 조사하고, 데이터를 정리한다. 그러나 그들이 만든 데이터는 의미의 관계성에 따라 함께 연결되어 종합적인 네트워크를 형성하게 되며, 네트워크화된 데이터는 정보화의 대상을 전체적으로 조망할 수 있게 한다. 이것은 연구자들이 시맨틱 모델링 작업을 전공별로 독립적으로 수행한 것이 아니라 서로 협의하면서 통섭적인 온톨로지(Ontology)'를 정의하고, 그 온톨로지에 따라 데이터를 만들었기 때문에 가능해진 일이다.

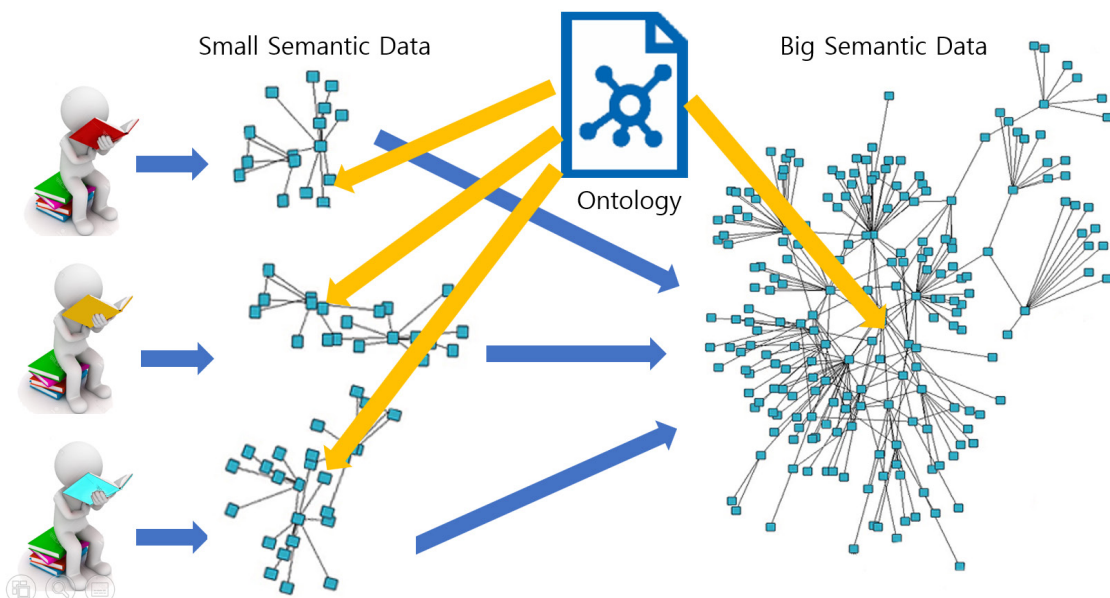


그림 15. 온톨로지를 기반으로 한 분야별 시맨틱 데이터의 통합³⁰⁾

어떤 대상 세계에 대한 시맨틱 모델링의 결과, 다시 말해서 대상 세계의 구성요소와 그것들 사이의 문맥이 디지털 세계에서 일정한 형식으로 표현될 수 있도록 만든 틀을 ‘온톨로지

(Ontology)’라고 한다.³¹ 이 온톨로지는 물리적인 현실 세계(또는 작품 속에 있는 가상의 세계)를 디지털 정보의 세계로 옮겨 짓는 설계도 역할을 하게 된다. 그 설계도에 따라 만들어지는 디지털 데이터는 현실 세계의 일들이 가졌던 의미를 데이터의 관계로써 표현하기 때문에 ‘시맨틱 (semantic, 의미론적, 의미 기반의)’이라는 수식어를 붙여서 ‘시맨틱 데이터’라고 한다. 아래의 그림은 (1) 시맨틱 모델링, (2) 시맨틱 모델링의 명세서인 온톨로지, (3) 온톨로지에 따라 만들어진 시맨틱 데이터의 관계를 도시한 것이다.

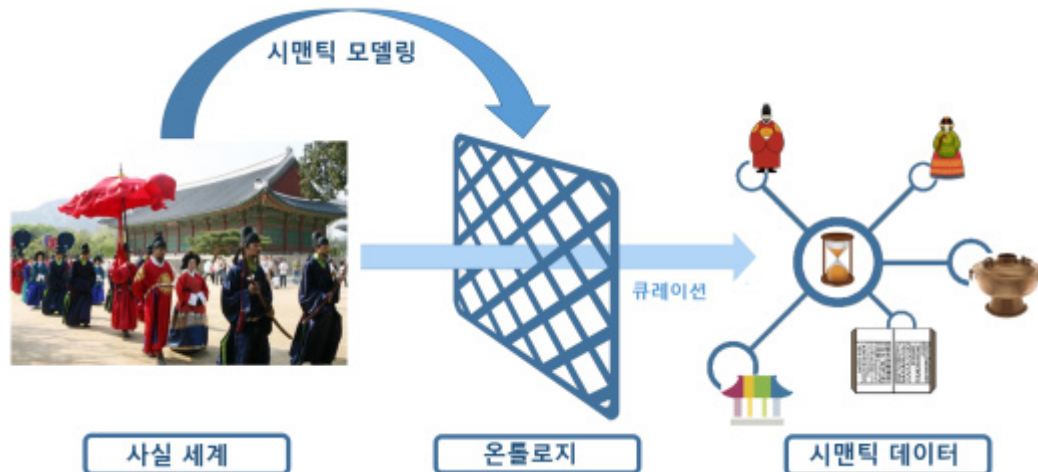


그림 16. 시맨틱 모델링, 온톨로지, 시맨틱 데이터의 관계

학제적 소통을 가능하게 하는 빅 데이터의 온톨로지를 설계하고 운용하는 것은 엄밀한 데이터 관리의 훈련이 필요한 일이다. 인문학도들은 디지털 인문학의 교과 수업을 통해 이 방법론을 습득하고 훈련의 기회를 얻는 것이 바람직하다. 온톨로지 설계의 기술적인 방법은 정보과학 분야에서 운용하는 것을 차용하지만, 우리가 관심을 두는 대상 세계에서 의미의 연결 체계를 파악해 내는 것은 전적으로 인문학적 지식과 통찰로 이루어 내는 것이라고 할 수 있다. 그런 의미에서 온톨로지에 대해 배우는 것은 그 자체로 인문학 공부라고 해도 무방하다. 온톨로지 (Ontology)라는 말부터가 철학의 존재론(存在論, Ontology)이라는 용어에서 유래한 것이다.³²

6. 디지털 스토리텔링과 AI 내러티브

‘시맨틱 데이터’를 기반으로 만들어진 인문 지식 디지털 아카이브는 우리의 인문학적 탐구의 대상이 되는 세계를 데이터로 재현한 저작물이다. 이것은 대상 세계의 구성 요소 하나하나에 대응하는 디지털 어셋들이 그 요소들 상호간의 의미론적 맥락까지도 알 수 있게 하는 형태로 집적된 것이라고 설명할 수 있다. 이렇듯 큐레이팅 된 인문학 자원의 저장소인 시맨틱 데이터 아카이브에 대해 우리는 다음과 같은 활용성을 기대할 수 있다.

첫째, 여러 학문 분야에 걸친 데이터 통합: 학제간 의사소통과 협업을 촉진하여 인간의 지식과 문화를 총체적으로 이해할 수 있게 한다.

둘째, 교육적 응용 지원: 교육자와 학생은 이 아카이브를 다양한 문화 요소의 상호 연결성을 탐구하는 자원으로 사용할 수 있다.

셋째, 인공지능-준비성: 시맨틱 데이터 형식으로 큐레이팅 된 인문 지식 콘텐츠는 인공지능이 효과적으로 처리하고 이해하며 활용할 수 있는 데이터이다. 인문학 연구자와 학생들은 시맨틱 데이터 큐레이션을 통해 인공지능과 협업하는 미래 인문학을 준비해 갈 수 있다.

아래의 그림은 디지털 세계의 미래 인문학 교실에서 시맨틱 데이터 아카이브가 수행할 역할을 개념적으로 도시(圖示)한 것이다.



그림 17. 융합적 디지털 환경의 미래 인문학 교실 개념

위의 그림에서 보였듯이 나는 미래 인문학 교실의 구성 요소를 메타버스와 인공지능 그리고 시맨틱 데이터 아카이브 세 가지로 상징하고 있다.

1) 시맨틱 데이터 아카이브(Semantic Data Archive):

- 대상 객체의 개별적인 속성 정보뿐 아니라 객체 사이의 관계에 대한 정보를 데이터로 기록함으로써, 대상 정보의 의미적 맥락을 표시할 수 있게 한 데이터베이스.
- 인문학 교육·연구 활동으로서의 인문 지식 디지털 큐레이션의 산출물이자, 새로운 교육과 연구의 도구로 쓰일 수 있는 지식 공유 환경

2) 메타버스(Metaverse):

- 3차원 가상현실 환경에서 하이퍼미디어 콘텐츠를 구현하고, 이용자들의 참여를 통해 유통, 확장, 재생산되도록 하는 개방적인 플랫폼
- 공급자→소비자의 단방향 세계가 아니라 프로슈머(Prosumer: Producer+ Consumer, 참여형 소비자)들에 의해 진화하는 쌍방향 세계

3) 인공지능(Artificial Intelligence):

- 인간의 지적, 언어적 활동의 효율성을 높일 수 있도록 만들어진 데이터 기반 소프트웨어.
- 반복적이고 데이터 양이 방대한 작업을 대신함으로써, 인간이 창의적이고 전략적인 활동에 더 깊이 몰두할 수 있도록 하는 촉진자.

디지털 원주민인 청소년들 사이에서 메타버스는 이미 떼어놓을 수 없는 그들의 놀이공간이 되고 있다. 관심도와 몰입감을 한껏 높이는 3차원 가상현실의 세계가 ‘놀이공간’만 되고 ‘교육공간’은 아니어야 할 이유는 없을 것이다. 미래세대를 그 세계에서 억지로 끌어내어 현실 세계의 교실에 가두려 하기보다는 우리의 인문교육이 그 가상 세계 안에서도 실현될 수 있는 방법을 찾아야 한다.



그림 18. 한양도성 타임머신 메타버스 제작을 위한 3D 썬³³ (창덕궁 부용지 일원)

메타버스는 사회적 관심이 큰 만큼, 그것을 위한 시각적 어셋 제작과 운용 기술 개발은 급속도로 발전해 갈 것이다. 하지만 그 가상 세계의 시각적 완성도가 아무리 높다 해도 이용자들이 그 거기에서 얻을 수 있는 지식의 내용이 빈약하거나 지식 소통의 수단이 불비하다면 그것은 우리가 기대하는 미래 교실일 수 없다.

가상현실로 재현된 역사의 현장에서 그 장면을 채우고 있는 건축물과 물건들이 무엇인지를 알려 주고, 그것이 어떠한 역사적 이야기의 맥락을 전하는 것인지 설명해 줄 수 있는 기계적 장치. 일방적으로 설명문을 포스팅하는 것이 아니라 이용자의 관심과 이해 수준에 따라 상호교감적인 지식의 소통을 이루어 낼 수 있는 소프트웨어. 이러한 기능의 지능적 커뮤니케이터가 우리가 기대하는 ‘미래 교실’ 메타버스에 필수적으로 존재해야 할 요소이다. 나는 인문 교육 메타버스에 요청되는 이 과제를 ‘시맨틱 데이터 아카이브’와 ‘인공지능’의 조합으로 풀어가야 할 것으로 생각한다.

한국학중앙연구원 디지털인문학연구소에서는 한국의 역사·문화에 관한 여러 가지 주제의 시맨틱 데이터 아카이브를 구축하는 일과 그 아카이브의 콘텐츠가 메타버스 상에서 인공지능과 협업

하는 교육자원으로 쓰일 수 있게 하는 방안을 다각도로 모색하고 있다. 다음은 그러한 실험의 하나로서, 인문 지식의 디지털 큐레이션을 통해 만들어진 시맨틱 데이터를 그 데이터 구조에 대한 정보와 함께 LLM 인공지능에게 제공하고 자연어 내러티브를 생성하도록 한 것이다. 예시 속의 내용은 앞에서 보았던 무신년진찬 개관 데이터이다.³⁴

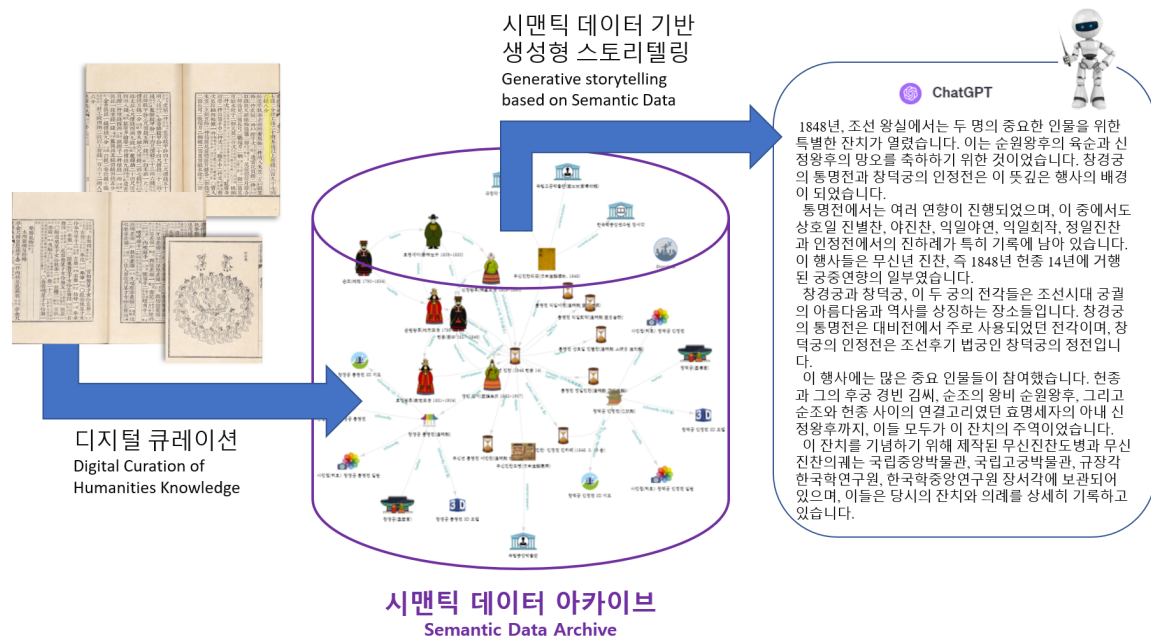


그림 19. 디지털 큐레이션에서 시작하는 생성형 스토리텔링 프로세스³⁵

챗지피티(ChatGPT)와 같은 ‘거대 언어 모델(Large Language Model, LLM)’ 인공지능은 사람에 견주어도 크게 손색이 없을 정도의 언어 구사력을 보여 세상을 놀라게 하고 있다. 그렇지만 말 잘하는 인공지능이 그 말을 배우는 과정에서 터득한 지식은 엄밀성을 중시하는 학술적 교육의 소산이 아니기 때문에 부정확하고 모호한 부분이 혼재되어 있다. 당연히 ‘교육적 역할’을 맡기기에 부족하다. 인공지능의 언어능력과 추리능력, 잡다한 정보력을 ‘정확한 문맥의 지식 정보 데이터’로 통제하고 검증할 수 있을 때 비로소 인공지능은 신뢰할 수 있는 교사의 역할을 할 수 있다.

인공지능에 정보를 전달할 때 데이터의 형식은 인공지능의 성능과 능력에 상당한 영향을 미친다. 복잡한 사건과 인물의 관계 속에서 특별한 맥락을 탐색해 내거나 그 맥락을 사람들이 쉽게 이해할 수 있는 자연어 내러티브로 변환하는 작업 등을 인공지능에게 요구할 때 구조화된 시맨틱 데이터를 제공하는 것은 자연어 텍스트 데이터를 제공하는 것에 비해 훨씬 더 우수한 결과를 얻을 수 있다.

예시로 보인 내러티브 텍스트에서 확인할 수 있듯이, 이 인공지능의 산출물은 유려한 문장 속에 시맨틱 데이터에 담긴 내용을 정확하게 반영하고 있다. 이 예시는 데이터를 문장으로 옮기는 정도의 역할을 지시한 결과이지만, 독자의 이해 수준과 관심 영역에 따라 응답의 내용을 변형, 확장하거나, 데이터 간의 관계성에 대한 논리적 추론도 신뢰할 수 있는 결과를 확인할 수 있었다.

이 글의 주제인 인문 지식의 ‘디지털 큐레이션’은 전통적인 인문학의 세계에서 말과 글로 다루어지던 인문학 지식을 명시적인 문맥의 디지털 데이터로 전환하는 일이다. 큐레이션 행위 자체가 전통적인 인문학 교육의 책 읽기와 글 쓰기에 해당하는 교육적 실천이지만, 디지털 아카이브로 남은 큐레이션의 결과물은 더욱 확장된 이차적 큐레이션을 촉진할 것이다. 인공지능은 이 이차적

큐레이션의 과정을 효율화하고 그 성과를 고도화할 수 있는 기재이다.³⁶

인문교육의 입장에서 인공지능에 대응하는 태도는 그것의 부정적 영향에 대해 걱정하고 외면하기보다, 인공지능이 효과적인 교육의 도구가 될 수 있도록 훈련시키는 노력을 기울여야 한다. 인공지능의 관점에서 보자면, 디지털 큐레이션은 전통적인 인문학 지식을 인공지능 가독형 데이터로 전환하는 수단이다. 인문지식의 디지털 큐레이션 교육은 인공지능을 학습의 동반자로 삼을 우리의 미래 세대가 전통적인 인문학 지식을 계승하고 창조적으로 발전시킬 수 있게 하는 과업이기도 하다.

7. 디지털 큐레이션 교육을 위한 제언

전통적인 인문학 교육에서 ‘글 쓰기’ 훈련이 중요시되었던 이유는 그것이 자신의 주체적인 지적 세계를 형성하는 데 도움을 준다고 여겼기 때문이다. 호기심을 일으킨 대상에 대해서 더 알아야 할 것을 찾으려 하고, 거기서 발견한 것은 신뢰할 수 있는 것인지 의심하며, 내가 수용한 지식의 조각들을 나의 문맥으로 재구성해 보려는 노력. 디지털 시대라고 해서 이 기본적인 인문적 소양을 외면한다면 그것은 ‘인문학’일 수 없으며 ‘디지털 인문학’이라는 이름도 적절하지 않을 것이다. 디지털 인문학은 인간 지성을 컴퓨터 프로그램으로 대체하는 것이 아니라, 컴퓨터의 도움을 통해 인간의 사고력과 통찰력을 더욱 강화하려는 노력이다.

나는 이 글에서 ‘디지털 미디어’ 상에서 ‘디지털 데이터’와 ‘디지털 어셋’의 형태로 지식의 재료를 수용하는 ‘디지털 원어민’ 세대의 인문학도들이 실천할 수 있는 ‘디지털로 글쓰기’의 방법을 ‘디지털 큐레이션’이라는 이름으로 제안하였다.

이 제안의 일차적인 오디언스는 현재 대학에서 디지털 인문학 교육을 시행하고 있거나 앞으로 도입하려 하는 교육자, 그리고 이 일에 협력하고 있거나 앞으로 그 역할을 주도적으로 하게 될 대학원생들이다. 이분들이 나의 생각에 얼마나 공감해 주실지 나로서는 알 수 없으나, 설사 충분히 공감한다고 해도 구체적인 실천이 뒤따르지 않으면 그 공감은 그저 ‘문제의식’의 공유 수준에 머무는 일이 될 것이다.

이제 ‘디지털 큐레이션’ 교육이 우리의 디지털 인문학 교육에서 실제로 ‘실행’될 수 있게 하기 위한 제언을 드리는 것으로 이 글을 마무리하고자 한다. 그것은,

첫째, ‘디지털 큐레이션 입문 교육 과정’은 지난 20년간 이 교육을 실험해 온 한국학중앙연구원 인문정보학과 및 디지털인문학연구소의 교육 경험 및 기술적 인프라를 충분히 활용하라는 제안,

둘째, ‘디지털 큐레이션 심화/응용 과정’은 입문 과정에서 확보한 기술적 프레임워크와 인프라 위에서 각 대학의 관심 교원들이 전문성 있는 특정 주제의 큐레이션 교과목을 개발하여 자기 대학 학생뿐 아니라 타 대학 학생들에게도 온/오프라인의 수강 기회를 제공하라는 권고,

셋째, 이미 여러 대학의 디지털 인문학 교수들이 참여하고 있는 ‘한국디지털인문학협의회(KADH)³⁷’에서 이와 같은 단계적, 협업적 교육 플랫폼에 대한 권고안을 제시하여 대학별로 교육 커리큘럼을 마련할 때 이를 참조할 수 있도록 하자는 것이다.

한국학중앙연구원 디지털인문학연구소는 여러 해 전부터 대학의 디지털 큐레이션 교육을 지원 하는 프로그램을 운영하고 있다. 하계/동계 방학 기간 중의 단기 특강, 대학 대학원 정규 교과목

의 강사 파견, 협동 강의 지원 등의 형태로 수행하며, 교육생들에게는 서버/데이터베이스 계정, 위키 S/W, 디지털 큐레이션 S/W 등 기술적 환경을 제공한다.

수강생들의 교육 만족도는 매우 높았으며, 특히 디지털 콘텐츠의 소비자에 머물렀던 자신들이 이제는 생산자가 될 수 있는 길을 열게 되었다는 점, 디지털 환경에서 동료들과 협력하는 방법을 알게 되었다는 점을 의미 있게 평가했다.

그러나 나의 관점에서 볼 때 이러한 교육이 성공적이라거나 지속가능하다고 이야기할 수 없다. 가장 큰 문제는 이러한 교육을 받은 학생들이 배운 것을 응용하고 실천할 기회를 얻기 어렵다는 점이다.

나의 학생들의 경우, 관련 교과나 연구 프로젝트를 통해 대학원 수학 기간 내내 디지털 큐레이션의 실천 기회를 갖는 반면, 다른 대학 학생들의 경우 교육 기간 동안에는 열심히 배웠어도 그것을 응용할 후속 프로그램이 없기 때문에 애써 배운 역량을 더 발전시키지 못하는 아쉬움을 남겼다.

위에서 밝힌 세 가지 제안 중 두 번째 것은 바로 이러한 문제점을 디지털 인문학계 공동의 노력으로 해소해 가자는 것이고, 세 번째 제안은 이러한 실천적 노력의 경험으로 가지고 디지털 인문학 교육을 학교 차원에서 도입하려는 대학들에 대해 정책적 조언을 드리자는 것이다.

디지털 인문학은 디지털 세계에서 인문 지식에 다가가는 방법을 배우고[學], 나의 관심거리가 생길 때마다[時] 그것을 디지털 세상에 표현하는 실천[習]이며, 그래서 인문학이 즐거워지는[悅] 공부이다. “學而時習之, 不亦說乎!” (『論語』, 「學而」)

¹ 미국의 교육학 저술가 마크 프렌스키(Marc Prensky, 1946-)는 어렸을 때부터 디지털 환경에서 자라나 디지털 기술의 활용을 예상롭게 생각하는 젊은이들을 ‘디지털 원어민’(Digital Natives)이라는 용어로 호칭했다. 이것과 상대되는 개념인 ‘디지털 이주민’(Digital Immigrants)은 성인이 된 이후에 디지털 기술을 접하고, 의지를 갖고 배워서 그 기술을 활용할 수 있게 된 기성세대를 지칭한다. (Marc Prensky, ‘Digital Natives, Digital Immigrants.’ On the Horizon 9(5). Bradford, UK: MCB University Press. 2001. <https://www.marcprensky.com/writing/Prensky%20-%20Digital%20Natives,%20Digital%20Immigrants%20-%20Part1.pdf>)

디지털 이주민은 그가 아무리 디지털 기술에 가까이 다가갔다 하더라도 필요할 때만 디지털 세계를 출입하고, 그때마다 적지 않은 스트레스를 받는 이방인이다. 반면 디지털 원어민은 디지털 세계가 곧 생활의 무대이고 그곳에서의 삶이 편안한 존재이다.

² 디지털 원어민은 새로운 지식과 정보를 디지털적인 방법으로 수용할 뿐 아니라, 그것을 모방하고 응용하여 만들어 낸 자신의 콘텐츠를 디지털 세상에서 표현하고, 그러한 행위를 통해 자신의 존재를 인정받고 싶어 한다. 모방과 창조, 표현하기와 인정받기에 대한 욕구는 가장 강력한 학습의 동기이다. (김현(2015), “디지털 인문학과 선비문화 콘텐츠”, <儒學研究> 33. <http://dh.aks.ac.kr/~tutor/Documents/PDF/선비문화.htm>)

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 으레 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.

⁵ “Curation as Augmented Scholarly Practice

Digital Humanists recognize curation as a central feature of the future of the Humanities disciplines.

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.

.....

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

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.

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." (Jeffrey Schnapp and Todd Presner (principal authors), *The Digital Humanities Manifesto 2.0*, 2009. https://www.humanitiesblast.com/manifesto/Manifesto_V2.pdf)

⁶ 아날로그 세계의 실물 아카이브가 소장하고 있는 자료를 온라인상에서도 열람할 수 있게 하는 것은 ‘아카이브 자료의 디지털화’(Digitalization of Archival Materials)이다. 이 시대에는 이것을 ‘디지털 아카이브’(Digital Archives)라고 부르기도 하지만, 그것이 디지털 아카이브가 지향하는 궁극의 모습은 아닐 것이다. 아카이브의 세계에서 지향하는 ‘디지털 전환(Digital Transformation)’ 목표는 전통적인 실물 아카이브의 ‘디지털화’에 머물지 않고, 아카이브의 ‘실물’, 그것과 관련한 지식의 탐구, 그 결과의 활용이 통섭적으로 하나가 될 수 있게 하는 것이다.

⁷ 현대에는 태생부터 디지털로 만들어지는 기록물이나 오브젝트도 많기 때문에 이와 같은 Born-Digital Asset 의 원형을 보존하는 것을 ‘디지털 아카이빙’이라고 하는 입장도 있다. 단, 이러한 개념의 디지털 아카이브에 관한 논의는 이 글의 주제와는 다른 이야기이므로 이 글에서는 논외로 한다.

⁸ 사실상 인터넷상에서 ‘디지털 아카이브’(Digital Archives)라는 이름으로 운영되고 있는 여러 사이트의 디지털 콘텐츠는 어떤 수준이건 큐레이션의 과정을 거친 결과물이라고 할 수 있다. 온라인 서비스를 통한 활용성의 증대를 목적으로 사이트를 개설한 이상, 고객이 어떠한 방법으로 그 데이터에 접근할지, 고객이 찾은 결과물을 어떠한 형태로 보여 줄 것인지에 대한 설계와 그에 따른 구현 노력이 어느 정도로는 반드시 있었을 것이기 때문이다. 유감스러운 점은 그러한 큐레이션이 자료의 성격과 의미를 잘 아는 해당 지식의 전문가에 의해 이루어지기보다 기술적인 데이터 조작 능력과 코딩 능력을 가진 기술자들에게 맡겨지는 경우가 많다는 점이다. 큐레이터의 역할을 할 사람이 정작 아카이브의 콘텐츠에 대해 잘 모르거나 무지하다면 그 큐레이션의 수준은 당연히 기대할 것이 못 된다. 이런 일이 반복되다 보니, 디지털 아카이브는 디지털 매체에 담긴 원시 데이터의 저장소일 뿐이고, 그 자료들의 의미와 맥락을 찾는 것은 이용자들이 개별적인 노력으로 해야 할 일이라는 생각이 당연시되기도 하였다.

⁹ “백과사전적 아카이브(Encyclopedic Archives, Encyves)는 인문 지식의 ‘원천 자료’이자 그 지식의 진실성을 입증하는 ‘증거’인 ‘실물 자료’ (기록물, 유물 등) 데이터가 광대한 인문 지식 네트워크의 노드(node)로 존재하는 세계이다.” (김현, 임영상, 김바로 (2016). “디지털 아카이브” <디지털 인문학 입문>. 한국외국어대학교 지식출판부.)

¹⁰ 필자가 ‘백과사전적 아카이브(Encyclopedic Archives, Encyves)’를 처음 구상하게 된 것은 한국학중앙연구원의 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. 일부러 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 알고도 하지 않는 듯했다. 나 역시 처음에는 그 두 가지 일이 기술적 프레임워크 면에서 유사할 뿐 별개의 일이라고 생각했다. 그러나 시간이 흐르면서 콘텐츠의 내용에 대한 이해가 깊어질수록 그 두 가지 일은 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 콘텐츠가 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. 그때, 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 염려했다. 둘째, 양쪽의 데이터를 문맥이 통하게 연계시켜서 종합적인 지식 콘텐츠를 만들려면 기존에 해 오던 사전편찬이나 영인물 간행의 수준을 넘어서서 훨씬 더 치밀하고 깊이 있는 조사 연구가 필요한데, 그 일을 위한 전문 인력의 충원이나 조직

개편이 용이하지 않았다. 셋째, 디지털 백과사전 편찬과 디지털 아카이브 구축 사이의 연계 시스템에 대한 구체적인 설계도나 참조 모델이 없었다. 나의 머릿속에만 있는 구상을 가지고 디지털 세계에 대한 이해가 부족한 부서원들의 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.

¹¹ <https://dh.aks.ac.kr/~sandbox/cgi-bin/GuestQuery.py?db=hanyang2023&project=hanyang&key=궁중연향의궤>

¹² https://dh.aks.ac.kr/~sandbox/cgi-bin/GuestQuery.py?db=hanyang2023&project=hanyang&key=무신진찬의궤:1848_도서목록

¹³ https://jsg.aks.ac.kr/dir/view?catePath=&dataId=JSG_K2-2874

¹⁴ 무신년진찬: 1848 년(무신년) 조선 헌종이 할머니 순원왕후(純元王后, 순조비, 1834-1849)와 어머니 신정왕후(神貞王后, 효명세자빈, 1789-1857)의 장수를 기원하기 위해서 행한 궁중연향. 행사에 대한 자세한 내용은 조선왕실 의궤(『戊申進饌儀軌』)에 기록되어 있고, 잔치의 장면들은 병풍 그림(「戊申進饌圖屏」)에 묘사되었다.

¹⁵ <https://dh.aks.ac.kr/~sandbox/cgi-bin/Story03.py?db=hanyang2023&project=hanyang&key=RoyalFeast1848-E1b>

¹⁶ <https://dh.aks.ac.kr/~sandbox/cgi-bin/Story03.py?db=hanyang2023&project=hanyang&key=RoyalFeast1848-E0b>

¹⁷ 하이퍼미디어(hypermedia): 디지털 환경에서 문자 텍스트와 함께 다른 미디어의 콘텐츠를 종합적으로 엮어내는 것. 의미의 연결고리를 좇아 무수한 텍스트 조각들이 자유롭게 연결되어 새로운 이야기를 만드는 하이퍼텍스트(hypertext)의 연결성이 문자 텍스트에 한정되지 않고 오감으로 체험하는 멀티미디어 콘텐츠로까지 확장되는 것을 뜻한다.

¹⁸ 어셋(Asset): 하나의 상품처럼 재화(財貨)의 가치를 인정할 수 있는 데이터 세트. 사진, 영상, 3D 모델, 그래픽 디자인, 사운드 클립 등의 디지털 저작물.

¹⁹ 데이터베이스(Database): 컴퓨터상의 여러 프로그램이 사용할 수 있도록 체계적으로 편성·저장된 데이터의 집합을 데이터베이스라고 한다. 이 ‘편성’과 ‘저장’의 기능은 ‘데이터베이스 관리 시스템’(DBMS)이라고 부르는 소프트웨어를 통해 수행된다.

²⁰ 시맨틱 모델링(Semantic Modeling): 정보화하려는 대상 세계의 구성 요소들 사이의 다양한 관계에 대해서 그 성격과 내용이 명시적으로 표시될 수 있도록 관계성 정의의 틀을 설계하는 것

²¹ 온톨로지(Ontology): 시맨틱 모델링의 과정을 거쳐서 만들어진 데이터 구조 정의. 다시 말해서 대상 세계의 구성요소와 그것들 사이의 문맥이 디지털 세계에서 일정한 형식으로 표현될 수 있도록 만든 틀. 온톨로지는 물리적인 현실 세계(또는 작품 속에 있는 가상의 세계)를 디지털 정보의 세계로 옮겨 짓는 설계도 역할을 하게 된다.

²² 시맨틱 데이터 프로세싱(Semantic Data Processing): 온톨로지 설계에 따라 구조화된 데이터를 편찬하는 공정. 이 디지털 데이터는 현실 세계의 일들이 가졌던 의미를 데이터의 관계로써 표현하기 때문에 ‘시맨틱(semantic, 의미론적, 의미 기반의)’이라는 수식어를 붙여서 ‘시맨틱 데이터’라고 한다.

²³ 한양도성 타임머신 사업 (2020~2022): 정부 콘텐츠산업 3 대 혁신전략 과제로 문화재청이 주도한 문화유산 디지털 콘텐츠 개발 사업. 한양도성 권역 내 600 년 조선 문화유산의 ‘실감 콘텐츠’ 제작을 위한 디지털 어셋 개발 및 지식정보 자원 개발 프로젝트로 수행되었다. (문화유산 실물 3D 모델링 + 시맨틱 데이터 아카이브 구축)

²⁴ 한양도성 타임머신 시맨틱 데이터 아카이브: 한국학중앙연구원 디지털인문학연구소에서 한양도성 타임머신 사업의 일환으로 편찬한 디지털 아카이브. 조선시대 한양도성의 역사와 문화를 심층적으로 이해하는 데 필요한 지식 정보를 광범위하게 조사·추출하여 데이터화 하고, 데이터 요소 상호간의 관계를 명시적으로 표현함으로써 대상 자료의 내용을 분석·응용·확장할 수 있도록 한 시맨틱 데이터베이스 기반의 디지털 아카이브이다. (한국학중앙연구원 디지털인문학연구소, “한양도성 시맨틱 데이터 아카이브”. <https://dh.aks.ac.kr/hanyang2/wiki/>)

²⁵ 조선시대에 왕실이나 국가에 큰 행사가 있을 때 후세에 참고할 수 있도록 일체의 관련 사실을 그림과 문자로 기록한 책. 궁중연향을 기록한 의궤는 인조~고종 대에 걸쳐 모두 18 종이 전한다.

²⁶ <https://dh.aks.ac.kr/~sandbox/cgi-bin/Story03.py?db=hanyang2023&project=hanyang&key=RoyalFeast1848-E1>

²⁷ <https://dh.aks.ac.kr/~sandbox/cgi-bin/Story03.py?db=hanyang2023&project=hanyang&key=RoyalFeast1848-E2>

²⁸ <https://dh.aks.ac.kr/~sandbox/cgi-bin/Story03.py?db=hanyang2023&project=hanyang&key=RoyalFeast1848-E3>

²⁹ <https://dh.aks.ac.kr/~sandbox/cgi-bin/Story03.py?db=hanyang2023&project=hanyang&key=RoyalFeast1848-E3>

³⁰ 한양도성 타임머신 프로젝트에 적용한 시맨틱 데이터 온톨로지의 주요 내용은 다음의 웹 페이지에서 확인할 수 있다. 한국학중앙연구원 디지털인문학연구소, 'EKC 데이터 모델', 『한양도성 타임머신 시맨틱 데이터 아카이브』. https://dh.aks.ac.kr/hanyang2/wiki/index.php/Ontology:EKC_2022

³¹ 온톨로지(Ontology): 정보화의 대상이 되는 세계를 전자적으로 표현할 수 있도록 구성된 데이터 기술 체계.

³² 원래 온톨로지(Ontology)라는 말은 철학에서 ‘존재론’이라고 번역되는 용어로서 ‘존재에 대한 근본적이고 보편적인 규정을 연구하는 학문’의 의미를 갖는 말이었다. 그러한 용어가 정보과학 분야에서 중요한 개념으로 등장하게 된 것은 인간이 존재의 세계를 이해하는 틀과 컴퓨터가 정보화 대상(콘텐츠)을 해석하는 틀 사이에 유사성이 있다고 보았기 때문이다.

³³ 메타버스 씬(Scene): 메타버스 제작에 쓰일 수 있도록 특정 주제에 따라 미리 만들어 놓은 ‘재사용 가능(Reusable) 3D 공간 패키지’. 유적지, 거리, 마을 등의 열린 공간이나 교실, 전시실, 콘서트, 홀과 같은 실내 공간을 연출하고, 메타버스 방문자들이 그 주제의 콘텐츠를 확장할 수 있도록 하는 용도의 디지털 콘텐츠.

³⁴ 이 예시는 한양도성 시맨틱 데이터 아카이브에 저장된 시맨틱 데이터 가운데 특정 주제와 관련이 있는 노드들을 선택한 후, 그 노드들 사이의 관계 정보와 개별 노드의 속성 정보를 API(Application Programming Interface)를 통해 Chat GPT 에 제공하고 지정 언어의 내러티브를 생성하도록 요청한 결과이다.

³⁵ 이 내러티브 생성을 위해 Chat GPT 에게 제공한 시맨틱 데이터와 지시 프롬프트는 다음의 웹 문서에서 확인할 수 있다. <https://dh.aks.ac.kr/hanyang2/wiki/index.php/RoyalFeast1848-E1>

³⁶ 시맨틱 데이터 기반 인공지능 운용의 장점은 다음과 같이 정리해 볼 수 있을 것이다.

1. 정밀도 및 명확성 (Precision and Clarity): 시맨틱 데이터는 개체와 그 관계를 명시적으로 정의하여 모호성을 줄이고 데이터의 맥락과 의미에 대한 AI 의 이해를 향상시킨다.
2. 데이터 처리의 효율성 (Efficiency in Data Processing): 구조화된 시맨틱 데이터는 AI 가 정보를 빠르게 탐색할 수 있도록 하여 데이터 검색 및 분석을 보다 효율화 한다.
3. 향상된 추론 (Enhanced Reasoning): 시맨틱 데이터가 기술하는 개체 간의 명시적인 관계는 AI 가 더 복잡한 추론을 수행할 수 있게 하고, 이를 통해 새로운 통찰력의 발견을 지원한다.

³⁷ 한국디지털인문학협의회. <https://www.kadh.org/>

참고문헌

김현, 임영상, 김바로(2016). <디지털 인문학 입문>. 한국외국어대학교 지식출판부.

김현(2015). “디지털 인문학과 선비문화 콘텐츠”. <儒學研究> 33.
<http://dh.aks.ac.kr/~tutor/Documents/PDF/선비문화.htm>

한국디지털인문학협의회. <https://www.kadh.org/>

한국학중앙연구원 디지털인문학연구소. “EKC 데이터 모델”. <한양도성 타임머신 시맨틱 데이터 아카이브>. https://dh.aks.ac.kr/hanyang2/wiki/index.php/Ontology:EKC_2022

한국학중앙연구원 디지털인문학연구소. “시맨틱 데이터 아카이브 편찬 연구”. <한양도성 시맨틱 데이터 아카이브>. <https://dh.aks.ac.kr/hanyang2/wiki/index.php/대문>

한국학중앙연구원 장서각. <디지털 장서각>. <https://jsg.aks.ac.kr>

Prensky, Marc (2001). *Digital Natives, Digital Immigrants*. On the Horizon 9(5). Bradford, UK: MCB University Press.
<https://www.marcprensky.com/writing/Prensky%20-%20Digital%20Natives,%20Digital%20Immigrants%20-%20Part1.pdf>

Schnapp, Jeffrey. Presner, Todd (2009). *The Digital Humanities Manifesto 2.0*.
https://www.humanitiesblast.com/manifesto/Manifesto_V2.pdf