인문논총 79권 2호(2022. 5. 31), pp. 391~427
[DOI] https://doi.org/10.17326/jhsnu.79.2.202205.391

지식의 공유와 표현 그리고 디지털 인문학

해외 디지털 인문학 연구 사례 검토

김지선*
류인태**

초록 최근 몇 년 사이 국내에서도 디지털 인문학에 대한 관심이 꾸준히 증가하고 있으나, 디지털 인문학 분야에서 보편적으로 다루어지는 인문학 데이터 구축·활용 맥락과 그 학술적 의의에 대한 이해는 여전히 미흡한 상황이다. 소위 '디지털 인문학'(Digital Humanities)으로 총칭되는 영역의 연구 형식이 다양해서 일목요연한 정리가 쉽지 않다는 점, 국내에는 디지털 인문학 연구 사례가 많지 않고 해당 영역의 전문 연구자가 극소수라서 디지털 인문학 연구의 실제를 전달할 수 있는 채널이 제한적이라는 점 등이 그 원인이라 하겠다. 이와 같은 점을 고려해 이 글은 해외의 디지털 인문학 연구 가운데 지식의 공유와 표현에 초점을 둔 대표적 사례 24건을 검토함으로써, 디지털 인문학 프로젝트의 데이터 처리 과정과 결과물 형식 그리고 그에 내포된 학술적 의의를 개략적으로나마 정리해보고자 하였다. XML, RDF로 대표되는 기계가독형 데이터를 포함해 CSV, JSON 등 웹 환경에서 통용되는 여러 포맷의 데이터를 구축·개방하는 한편, D3.js, Leaflet.js, JQuery, Angular.js와 같은 다양한 맥락의 JavaScript Library와 Web Framework를 활용하는 것이 그 공통적인 특징이라 하겠다. 이 글에서 다룬 각각의 디지털 인문학 연구 사례에 관한 검토가 추후 개별적으로 이루어진다면, 디지털 인문학 연구의 실제를 살피는 데서 더 나아가 디지털 인문학 연구 프로젝트를 구체적으로 기획·진행하는 데 있어서도 유용한 단서로 삼을 수 있을 것이다. 디지털 환경과 결합한 인문학 연구가 구체적으로 어떠한 방향으로 나아가야 할지 그리고 그 과정에서 발생하는 인문학적 문제의식과 디지털 테크놀로지에 대한 이해를 어떻게 수용할 것인지를 본격적으로 고민해야 하는 시점에서, 이 글이 작게나마 기여하는 바가 있기를 바란다.

주제어 디지털 인문학, 인문 데이터, 데이터 공유, 데이터 표현

---

* 제1저자, 고려대학교 문과대학 강사
** 교신저자, 성균관대학교 국어국문학과 박사후연구원

1. 머리말

디지털 인문학을 향한 학계의 관심이 꾸준히 증가하고 있다. 디지털 환경 및 데이터 처리와 연계한 인문학 연구를 주제로 곳곳에서 학술대회가 개최되고 있으며, 디지털·데이터 기반의 방법론을 활용한 인문학 연구 논문이 지속적으로 발표되고 있다. 개인 연구나 학회 활동의 차원뿐만 아니라 대학 강단에서의 교육 및 연구기관의 학술적 지향 그리고 현실적으로는 인문학 연구자들의 취업 채널에까지 디지털·데이터 테크놀로지의 영향력이 가시화되고 있는 상황에서, 소위 '디지털 인문학'으로서 디지털 환경과 인문학 연구가 조우하는 구체적 양상에 관한 탐구는 필연적일 수밖에 없다.

이것을 단순히 Python이나 JavaScript와 같은 프로그래밍 언어, HTML(CSS)이나 XML과 같은 마크업 언어, RDB(SQL)나 GIS와 같은 데이터베이스 기술, RDF(SPARQL)와 같은 웹 기반 데이터 기술 체계를 인문학 연구·교육에 어떻게 활용할 것인가에 관한 기능적 접근으로만 이해해서는 곤란하다. 오히려 그와 같이 다양한 디지털 테크놀로지가 결합하여 만들어 내는 복잡한 데이터 처리 양상이 디지털 환경을 매개로 이루어지는 최근의 지식 생산·교환 방식에 미치는 영향과, 아날로그 환경에 익숙한 인문학 연구자들이 향후 그에 어떻게 대응해 나갈 것인가를 구체적으로 고민하기 위한 기초적 사유로 인지할 필요가 있다.

이 글은 그러한 문제의식을 바탕으로 기획된 것이다. 본론에서 소개할 해외의 대표적 디지털 인문학 연구 사례들은 단순히 인문학 데이터 처리를 위한 방법론(methodology) 차원에서 검토되어야 할 대상이 아니라, 그러한 데이터 기반의 방법론이 인문학 자원을 다루고 이해하는 방식에 미치고 있는 영향과 그러한 영향이 인문학술 환경에 야기하고 있는 변화를 짚어 보기 위한 구체적 단서라 할 수 있다.

이와 관련해 그동안 해외의 디지털 인문학 연구 동향이나 유관 학술 환경을 다룬 논문이 국내에서 꾸준히 발표되어 왔으나, 대체로 특정 '국가'나

'전공 분야'에 치우친 논의가 많으며, 개별 사례에 대한 접근 역시 기술적 차원의 검토·분석까지는 이르지 못하고 연구 목적이나 결과 및 시사점 정도를 개략적으로 제시한 경우가 대부분이다.[1] 이 글은 그와 같은 기존 사례 연구를 충실히 참고하되, 한편으로 기 연구에서 초점을 두지 않은 디지털 환경이나 데이터 처리와 관련된 기술적 차원의 접근을 조망함으로써, 디지털 인문학 연구의 실제로서 인문학 데이터 처리의 양상과 그 의미가 무엇인지를 살펴보고자 한다.

---

1 다음은 그에 해당하는 구체적 연구 결과들이다. 김바로(2014), 「해외 디지털 인문학 동향」, 『인문콘텐츠』 33, 인문콘텐츠학회, pp. 229-254; 홍정욱·김기덕(2014), 「'2014 세계 디지털인문학' 학술대회 및 한국의 디지털인문학」, 『인문콘텐츠』 34, 인문콘텐츠학회, pp. 53-75; 김종대(2015), 「독일의 디지털 인문학 동향」, 『철학과 문화』 32, 한국외국어대학교 철학문화연구소, pp. 55-76; 서경숙(2015), 「디지털 인문학 교수법의 이론 및 실제: 영미문학을 중심으로」, 『인문콘텐츠』 38, 인문콘텐츠학회, pp. 19-39; 홍정욱(2015), 「디지털기술 전환 시대의 인문학」, 『인문콘텐츠』 38, 인문콘텐츠학회, pp. 41-74; 송인재(2016), 「동아시아 개념사와 디지털인문학의 만남」, 『개념과 소통』 18, 한림과학원, pp. 97-130; 이지수(2017), 「디지털 인문학의 정의 및 연구 현황: 일본학 지식정보의 디지털화를 중심으로」, 『일본어문학』 74, 한국일본어문학회, pp. 41-62; 권윤경(2018), 「새로운 문필공화국을 향하여: 18세기 프랑스사 연구와 디지털인문학의 사례들」, 『역사학보』 240, 역사학회, pp. 35-74; 김용수(2018), 「영미문학과 디지털인문학 미국 디지털 영문학 연구 동향」, 『비평과이론』 23, 한국비평이론학회, pp. 35-58; 문수현(2018), 「독일의 디지털 역사학 현황」, 『역사학보』 240, 역사학회, pp. 1-33; 박은재(2018), 「영국 디지털 역사학의 발전과 현황」, 『역사학보』 240, 역사학회, pp. 75-98; 손정훈·김민규(2018), 「프랑스의 디지털인문학 동향과 정책」, 『프랑스문화예술연구』 64, 프랑스문화예술학회, pp. 125-151; 송인재(2018), 「대만 디지털인문학의 발자취와 진화」, 『역사학보』 240, 역사학회, pp. 127-155; 이주영(2018), 「미국에서의 디지털 역사학 발전 과정과 최근의 경향」, 『역사학보』 240, 역사학회, pp. 99-126; 이지수·이혜은(2019), 「미국 대학의 디지털 인문학 교육 프로그램 연구」, 『정보관리학회지』 36, 한국정보관리학회, pp. 107-128; 정유경(2020), 「디지털 인문학 분야의 국내외 연구 동향 분석」, 『정보관리학회지』 37, 한국정보관리학회, pp. 311-331; 이지수(2022), 「대학도서관의 디지털 인문학 연구지원 가이드에 관한 연구: 미국 대학도서관 사례를 중심으로」, 『한국문헌정보학회지』 56, 한국문헌정보학회, pp. 297-320.

2. Digital Humanities Awards의 내용과 그 의의

2.1. Digital Humanities Awards 개괄

전 세계 곳곳에서 이루어지고 있는 다채로운 형식의 디지털 인문학 연구를 전부 찾아내 검토하는 것은 불가능하다. 이에 따라 수많은 디지털 인문학 연구 가운데 대표성이 뚜렷한 사례를 선정하기 위해 Digital Humanities Awards의 시상 분야와 실제 지명·수상한 사례를 참고하였다. 디지털 인문학 연구의 보편적 경향을 토대로 대표성을 확인하기 위해서는 디지털 인문학 연구의 절차나 방법론 그리고 결과물의 형식이 갖는 공통적 특징이 무엇인지를 살펴보아야 하는데, 그것을 참고할 수 있는 다양한 리소스가 Digital Humanities Awards에 있기 때문이다. Digital Humanities Awards[2]는 매년 초 작년 한 해 동안 전 세계에서 발표된 디지털 인문학 연구 결과물을 대상으로 한 시상 행사로, 2012년부터 2021년까지 총 10차례의 시상식이 열렸다.

시상 후보 등록은 매년 초 홈페이지를 통해 공지되며, 디지털 인문학 분야의 전문 연구자뿐만 아니라 일반 연구자나 심지어 대중도 연구 결과물을 등록할 수 있다. 이후 등록된 사례 가운데 시상 후보에 대한 지명은 디지털 인문학 분야의 전문 연구자들로 구성된 '국제 지명 위원회'(International Nominations Committee)[3]의 논의를 거쳐 진행된다. 지명 절차는 국제지명위원회가 감독하며, 몇 가지 기준의 충족 여부에 따라 각 시상 분

2 Digital Humanities Awards URL: http://dhawards.org

3 2012년부터 2021년까지 총 13명의 연구자가 국제지명위원회의 일원으로 참여했다. 그동안의 경우를 살펴보면 당해 연도 국제지명위원회에 배속된 위원의 국적 및 소속 기관은 중복되지 않는 것이 특징이다. 국제지명위원회의 규모는 2019년에 두 배로 증설되었는데, 이는 위원회를 국제적으로 분산시킴으로써, 다양한 언어와 문화적 배경을 고려하기 위함이다. 한국의 경우 현재 서울대학교 자유전공학부에 소속된 차주항(Javier Cha) 교수가 2019년부터 2021년까지 세 차례 위원으로 참여했다.

야별 최종 후보자를 결정한다. 지리적 배경, 언어적 배경, 그리고 인문학의 전통적 연구 영역(field)에 국한하지 않고, 그 배경이 어떠한 결과물이든 디지털 인문학 형식에 적합한 자원은 모두 지명될 수 있다.[4]

지명된 사례들에 대한 평가는 웹사이트를 방문하는 사용자들의 투표를 통해 이루어지며, 투표 결과는 홈페이지를 통해 분야별 우승과 준우승 및 기타 후보자 목록의 형식으로 공개된다. 이러한 시상식을 연례화한 까닭은, 낯설고 새로운 연구 분야로서 디지털 인문학 자원에 대한 관심을 환기하고, 디지털 인문학 분야의 전문 연구자나 디지털 인문학 분야에 관심을 둔 일반 대중을 대상으로 한 커뮤니티 참여를 제고하기 위한 것이라 할 수 있다.

2.2. Digital Humanities Awards 시상 분야와 그 의미

Digital Humanities Awards에서 특기할만한 지점이 하나 있으니, 바로 시상 분야다. 분야별로 시상이 진행된다는 것은, 각각의 분야가 디지털 인문학 연구의 근본적 특징을 반영하고 있음을 의미한다. Digital Humanities Awards의 시상 분야를 살펴보면 아날로그 환경에서 줄곧 이루어져 온 전통적 인문학 대비 디지털 인문학의 특징이 무엇인지를 유추해 볼 수 있다.

[표 1]은 지난 10년간 이루어진 시상 분야를 정리한 것이다. 2012년부터 2021년까지 10년간 이루어진 시상 분야는 매년 상이한데, 대략 그 항목을 살펴보면 '즐거움', '도구', '시각화', '웹 간행', '대중성', '교육', '실패', '데이터셋', '기타' 등 9가지 맥락으로 나누어 살펴볼 수 있다. 각각의 분야는 디지털 환경 또는 데이터 처리 기술이 인문학 연구에 접목될 때 발생할 수 있는 특수한 양태를 가리키며, 해당 지점들은 인문학 연구 환경에 대한 독특한 시사점을 전달한다.

---

4 국내 디지털 인문학 연구 사례 가운데서는, 〈조선시대표류기록시각화〉(http://soundh.net/drift)가 2019년에, 〈지암일기 디지털 인문학 연구〉(http://jiamdiary.info)가 2020년에 각기 Data Visualization 분야에 수상 후보로 지명되었다.

[표 1] Digital Humanities Awards의 10년간 시상 분야(2012~2021)

<table>
<tr><th>연도</th><th>즐거움</th><th>도구</th><th>시각화</th><th>웹 간행</th><th>대중성</th><th>교육</th><th>실패</th><th>데이터</th><th>기타</th></tr>
<tr><td>2012</td><td rowspan="11">USE OF DH FOR FUN</td><td rowspan="11">DH TOOL or SUITE OF TOOLS</td><td rowspan="2">DH VISUALI-ZATION or INFO GRAPHIC</td><td>DH BLOG/ ARTICLE or SHORT PUBLIC-ATION</td><td rowspan="2">DH PROJECT FOR PUBLIC AUDIENCES</td><td>PROFES-SIONAL RESOURCES FOR LEARNING ABOUT or DOING DH WORK</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>2013</td><td>DH BLOG/ POST/ ARTICLE or SHORT PUBLIC-ATION</td><td>-</td><td>-</td><td>-</td><td>DH CONTRIBU-TION NOT IN THE ENGLISH LANGUAGE</td></tr>
<tr><td>2014</td><td rowspan="9">DH DATA VISUALI-ZATION</td><td rowspan="7">DH BLOG/ POST or SERIES OF POSTS</td><td rowspan="4">USE OF DH FOR PUBLIC ENGAGE-MENT</td><td>-</td><td>EXPLOR-ATION OF DH FAILURE</td><td>-</td><td>-</td></tr>
<tr><td>2015</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>2016</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>2017</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>2018</td><td>-</td><td>EXPLOR-ATION OF DH FAILURE</td><td>-</td><td>-</td></tr>
<tr><td>2019</td><td>DH PUBLIC ENGAGE-MENT</td><td>-</td><td rowspan="3">EXPLOR-ATION OF DH FAILURE/ LIMITA-TIONS</td><td rowspan="3">DH DATASET</td><td>-</td></tr>
<tr><td>2020</td><td>USE OF DH FOR PUBLIC ENGAGE-MENT</td><td>-</td><td rowspan="2">SPECIAL CATE-GORYDH RESPONSE TO COVID-19</td></tr>
<tr><td>2021</td><td>DH SHORT PUBLIC-ATION</td><td>-</td><td>DH TRAINING MATERIALS</td></tr>
</table>

우선 10년간 거의 빠짐없이 시상이 이루어져 온 분야로 '즐거움', '도구', '시각화', '웹 간행', '대중성'을 거론할 수 있다. '즐거움'(fun)은 말 그대로 '즐거움'을 보여 주는 디지털 인문학 연구 사례를 시상하기 위한 분야이다. 디지털 환경의 탄성을 보여 주는 대표적 채널이자 상대적으로 아날로그 인문학의 '경직성'을 비판하는 요소로서, '즐거움'은 디지털 인문학의 방향성을 뚜렷하게 나타내는 지표라 하겠다.

'도구'(tool)는 디지털 인문학 연구를 진행하는 데 있어서 적합한 도구(모음)를 만들어 여러 연구자에게 잘 공유한 사례를 시상하기 위한 분야이다. 인문학 연구자 입장에서 연구 목적에 적합한 형태로 데이터 처리 기술을 가공하기가 쉽지 않기에, 연구에 즉각적으로 활용할 수 있는 '도구'를 개발·공유하는 것이 디지털 인문학 분야에서 중요한 지점임을 짐작할 수 있다.

'시각화'(visualization)는 데이터의 '시각화' 구현 결과가 훌륭한 사례를 시상하기 위한 분야다. 아날로그 환경에서의 글쓰기를 보완하는 커뮤니케이션 채널로서 디지털 환경에서의 재현(representation) 방식이 인문학 연구의 새로운 형식이 될 수 있는가에 대한 실험적 모색이 꾸준히 진행되고 있음을 의미한다.

'웹 간행'(web publication)은 인문학 자료나 지식을 블로그(blog), 포스트(post), 기사(article), 단편 간행물(short publication) 등 웹 자원으로 잘 편찬한 사례를 시상하기 위한 분야이다. 인문학 지식이나 유관 분야의 문제의식이 그와 같은 미디어로 편찬된다는 것은 곧 디지털 환경에서 접근 가능한 인문학 자원이 확장됨을 의미한다.

'대중성'(public engagement)은 일반 대중의 접근성과 활용도가 높은 사례를 시상하기 위한 분야다. 디지털 인문학은 디지털 환경을 매개로 연구가 이루어지기에, 연구 과정이나 결과물을 다루는 데 있어서 대중의 참여 가능성이 크다. 연구자 개인의 전문성에만 초점을 두고 이루어지는 전통적 인문학 연구와는 다른 차별적 지점이라 하겠다.

‘즐거움’, ‘도구’, ‘시각화’, ‘웹 간행’, ‘대중성’과 달리 ‘교육’, ‘실패’, ‘데이터셋’, ‘기타’는 지난 10년간 불규칙적으로 이루어져 온 시상 분야에 해당한다. 그러한 불규칙함이 ‘중요성’을 판단하는 기준이 될 수는 없겠으나 참고할 만한 지점은 될 것이다.

‘교육’(learning/training)은 디지털 인문학 방면의 ‘학습’이나 ‘훈련’을 위해 만들어진 다양한 형식의 강연 슬라이드, 실습자원 또는 안내문 등의 교육 자료 가운데 탁월한 사례를 찾아 시상하기 위한 분야다. 디지털 기술의 활용 범위와 실제 그것을 연구에 운용할 수 있는 연구자의 역량은 천차만별이기에, 그에 대한 솔루션으로서 ‘교육’은 디지털 인문학에서 매우 중요한 영역에 해당한다.

‘실패’(failure)는 디지털 인문학 프로젝트가 실패한 이유 또는 디지털 인문학 연구를 진행하는 과정에서 발생한 문제를 잘 다룬 보고서를 시상 대상으로 삼는다. ‘실패’란 기본적으로 디지털 인문학 연구 결과물을 온전히 산출하지 못한 경우를 가리키지만, 동시에 ‘결과물 산출에 실패한 이유를 찾고 개선하기 위한 단서’로서의 의미 또한 갖는다.

‘데이터셋’(dataset)은 잘 구축한 인문학 ‘데이터셋’에 대한 시상 분야이다. 일반적으로 디지털 인문학 연구에 주로 활용되는 텍스트 말뭉치나 데이터 테이블 또는 그러한 형식과 유사한 정보의 군집체를 가리켜 ‘데이터셋’이라 통칭한다. 지식을 구성하는 요소로서 다양한 맥락의 데이터셋을 구축·공유하는 것이 디지털 인문학의 중요한 학술적 과업임을 확인할 수 있다.

앞서 7가지 분야와 비교할 때 다소 독립적인 맥락에 해당하는 2가지 항목을 기타 분야로 분류하였다. 2013년에 시상이 이루어진 ‘비영어권’ 분야와 2020년에 신설되어 2년간 시상이 지속된 ‘코로나 바이러스’(COVID-19) 분야가 그것이다.

‘비영어권’ 분야는 비영어권의 디지털 인문학 자원 또는 출판물 가운데 디지털 인문학 영역에 기여가 뚜렷한 연구를 시상하기 위한 항목이다.

이 분야에 대한 시상이 2013년 이후 없는 상황에 대해서는 여러 방면의 해석이 가능하겠으나, 2012년과 2013년 이후 여러 분야의 시상 대상에 비영어권에서 구현된 디지털 인문학 자원이 자연스럽게 반영된 결과가 가장 크다고 하겠다. '코로나 바이러스'의 경우 2020년부터 전 세계에 확산된 코로나 팬데믹 상황에 대처하기 위해 이루어진 디지털 인문학 연구에 대한 시상 분야로서, 최근의 특수한 시류를 반영한 항목에 해당한다.

특수한 항목으로서 '기타'를 제외한 '즐거움'(fun), '도구'(tool), '시각화'(visualization), '웹 간행'(web publication), '대중성'(public engagement), '교육'(learning/training), '실패'(failure), '데이터셋'(dataset)' 8가지 분야는 Digital Humanities Awards의 구체적 시상 분야인 동시에 아날로그 환경의 인문학과는 상이한 디지털 인문학의 특징을 뚜렷하게 보여주는 요소들에 해당한다. 전통적 인문학 연구와는 다소 거리가 있는 8가지 분야에서 지난 10년간 지명·수상된 연구는, 전문가와 대중에 의한 복합적 검증을 거친 결과들로서, 디지털 인문학의 특장(特長)을 드러내는 대표적 연구 사례라 할 수 있다.

3. 해외의 대표적 디지털 인문학 연구 사례와 그 의미

3.1. 검토 대상 선정

앞서 살펴본 바와 같이 Digital Humanities Awards의 '시상 분야'는 디지털 환경에서 이루어지는 인문학 연구의 특징과 의미를 내포하고 있는 보편적 단서로서, 일정 이상 '디지털 인문학 연구로서의 기초 요건과 정체성'을 정의하기 위한 근거로 삼을 수 있다. 그에 따라 본 연구자는 Digital Humanities Awards의 지명·수상 사례 가운데 데이터의 공유와 표현에 있어 특장이 뚜렷한 연구들을 선정하였으며, Digital Humanities Awards에서 지명·수상되지는 않았으나 웹 환경에서 인문학 데이터를 효과적으로 공유

하기 위한 방안을 적극적으로 제시하거나 인문 지식을 대상으로 한 시각적 표현을 창의적으로 시도한 연구 사례들을 추가적인 검토 대상으로 정리하였다.[5] 구체적 검토 대상으로 삼은 디지털 인문학 연구 사례 24건은 아래와 같다.[6]

• Digital Humanities Awards 지명·수상 사례(나열: 지명·수상 연도순)

① Livingstone Online (URL: https://livingstoneonline.org)

-〈Livingstone Online〉은 파일럿 프로젝트 〈Livingstone's Letter from Bambarre: A Multispectral Critical Edition〉(2010-11)과 이후 수행된 〈Livingstone's 1871 Field Diary: A Multispectral Critical Edition〉(2010-13), 〈Livingstone's 1870 Field Diary: A Multispectral Critical Edition〉(2013-17), 〈LEAP: The Livingstone Online Enrichment and Access Project〉(2013-17) 등 4개의 프로젝트 결과물을 통합한 프로

5 '디지털 인문학 연구로서의 기초 요건과 정체성'을 정의하는 데 있어서 데이터의 공유와 표현에 초점을 둔 것은, '개방', '협업', '공유', '표현' 등이 디지털 인문학 연구의 특징을 드러내는 대표적 방향으로 거론되기 때문이다. 그에 대한 단서는 2009년 미국에서 발표된 '디지털 인문학 선언문'(Digital Humanities Manifesto)을 통해서도 확인할 수 있다. 해당 선언문의 내용을 참조할 경우 디지털 인문학의 근본적 방향성은 '웹 환경의 개방성과 협업을 기반으로 하여 다양한 표현 방식으로 지식을 생산 및 재생산'하는 데 있다. 관련된 자세한 내용은 홍정욱(2015), pp. 63-67을 참고할 수 있다.

6 24건의 연구 가운데 10건은 Digital Humanities Awards의 지명·수상 사례이며, 나머지 14건은 본 연구자가 그동안 수집·검토해 온 사례들 가운데 '공유'와 '표현'으로서 디지털 인문학 연구의 특징을 뚜렷하게 드러내는 것들이다. 이와 관련해 영국 케임브리지대학에서 구축한 'Casebooks'(https://casebooks.lib.cam.ac.uk)와 미국 스탠포드 대학에서 진행한 'The Republic of Letters'(http://republicofletters.stanford.edu) 그리고 하버드대학과 중국 북경대학의 협업으로 편찬된 'CBDB'(https://projects.iq.harvard.edu/cbdb)와 네덜란드 라이덴대학을 중심으로 구현된 'MARKUS'(https://dh.chinese-empires.eu/markus) 등 디지털 인문학을 대표하는 사례로 익히 알려진 여러 프로젝트는 검토 대상에서 제외하였다. 그와 같은 결정은, 익히 알려지지 않았으나 검토할 가치가 있는 디지털 인문학 연구 사례들을 제한된 지면 내에서 가급적 다양하게 살펴보고자 한 의도로부터 기인한 것임을 밝히는 바이다.

젝트로, 〈Livingstone's 1871 Field Diary〉(2010-13)과 〈Livingstone's 1870 Field Diary〉(2013-17)은 각각 2012 "Best professional resources for learning about or doing DH work" 분야와 2016 "Best DH Data Visualization" 분야에서 준우승을 차지했으며 〈Livingstone Online〉은 2017 "Best Use of DH For Public Engagement" 분야에 지명되었다.

② Kindred Britain (URL: https://kindred.stanford.edu)

-2013 "Best DH Visualization or Infographic" 분야에서 3위를 했다.

③ Chines Text Project (URL: https://ctext.org)

-2017 "Best DH Tool or Suite of Tools" 분야에 지명되었다.

④ Six Degrees of Francis Bacon (URL: http://sixdegreesoffrancisbacon.com)

-2017 "Best DH Data Visualization" 분야에 지명되었다.

⑤ Coins (URL: https://uclab.fh-potsdam.de/coins)

-2018 "Best DH Data Visualization" 분야에서 우승했다.

⑥ Derrida's Margins (URL: https://derridas-margins.princeton.edu)

-2018 "Best DH Tool or Suite of Tools" 및 "Best DH Data Visualization" 분야에 지명되었다.

⑦ The Pulter Project (URL: http://pulterproject.northwestern.edu)

-2019 "Best DH Data Visualization" 분야에 지명되었다.

⑧ Witches (URL: https://witches.is.ed.ac.uk)

-2019 "Best DH Data Visualization" 분야에서 우승했다.

⑨ On the Trail of Antiquities (URL: https://ventesdantiques.inha.fr)

-2019 "Best DH Data Visualization" 분야에서 준우승을 차지했다.

⑩ Tudor Networks (URL: http://tudornetworks.net)

-2020 "BEST DH Data Visualization" 분야에서 준우승을 차지했다.

• 추가 사례(나열: 프로젝트 수행 시점 기준)

⑪ Slave Voyages (URL: https://www.slavevoyages.org)

⑫ London Lives (URL: https://www.londonlives.org)

⑬ Lviv Interactive (URL: https://lia.lvivcenter.org)

⑭ The Joyce Project (URL: http://www.joyceproject.com)

⑮ Icelandic Saga Map (URL: http://sagamap.hi.is/is)

⑯ Linked Jazz (URL: https://linkedjazz.org)

⑰ O Say Can You See (URL: https://earlywashingtondc.org)

⑱ Digital Panopticon (URL: https://www.digitalpanopticon.org)

⑲ Lit Long (URL: https://litlong.org)

⑳ Novel City Maps (URL: https://uclab.fh-potsdam.de/NCM)

㉑ The Museum of the World (URL: https://britishmuseum.withgoogle.com)

㉒ Enslaved (URL: https://enslaved.org)

㉓ Codex Atlanticus (URL: http://www.codex-atlanticus.it)

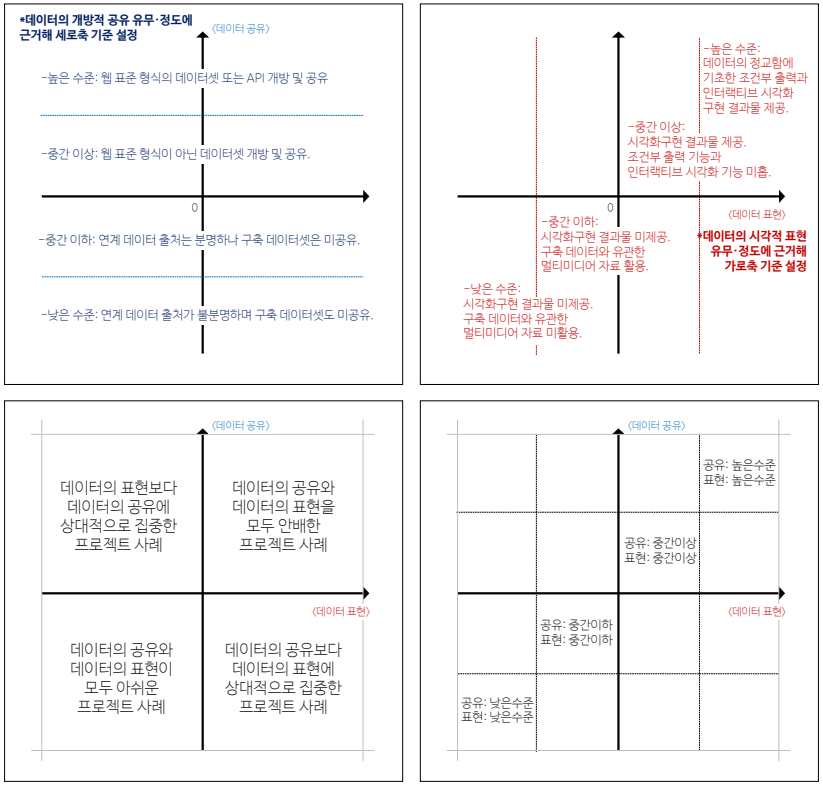
㉔ 宋元学案知识图谱 (URL: https://syxa.pkudh.org)

3.2. 검토 기준 설정

24건의 사례는 연구 분야와 목적, 기간과 대상 및 구현 결과물의 형식이나 성격이 저마다 다르기에, 단일 기준에 의거해 그 가치와 의의를 검토·평가하기 어렵다. 이를 고려해 디지털 인문학 연구로서의 보편적 요건을 공통으로 전제하되 개별 사례를 일목요연하게 비교할 수 있는 기준을 안배하고자 하였으며, 그에 따라 '데이터 공유'와 '데이터 표현' 두 가지 항목을 검토 기준으로 삼았다.[7]

---

7 앞서 각주 5를 통해서도 언급하였듯이 디지털 인문학 연구의 보편적 요건으로 거론되는 대표적 방향성은 데이터의 공유와 표현이라 할 수 있다. 이와 관련해 본 연구자 또한 디지털 인문학 연구의 근본적 방향으로서 데이터의 공유와 표현을 강조한 적이 있다. "특정한 학문 분야에서 줄곧 다루어 온 방법론 및 그와 관련된 담론에 뿌리를 두지 않고, 전통적인 인문학 연구에 대한 이해를 토대로 웹 표준에 근거한 형식의 데이터를 편찬해서 공유하고 더 나아가 해당 데이터를 디지털 형식으로 새롭게 표현하고자 하는 연구들이



[그림 1] 좌표평면에 안배한 '데이터 공유'(세로)와 '데이터 표현'(가로) 기준

[그림 1]은 '데이터 공유'와 '데이터 표현' 두 기준을 좌표평면상에 나타내었을 때, 매핑 영역으로서 각각의 분할 공간이 어떠한 기준에 의거하는지를 정리한 것이다. '데이터 공유'는 아카이브나 데이터베이스 구축의 기본 단계인 원자료 가공과 데이터 모델링 그리고 결과물 데이터셋 제공 등

다. 특정한 학문 분야에서 중심적으로 다루는 방법론 및 담론을 매개로 하는 것이 아니라, 인문학 본연의 문제의식을 웹 환경에서 통용되는 표준적 성격의 기술을 바탕으로 어떻게 '공유'하고 '표현'할 것인가에 초점을 두는 것이 그 특징이라고 할 수 있다."[류인태, 「디지털 인문학은 인문학이다」, 『인문논총』 77(3), 서울대학교 인문학연구원, 2020, p. 391]

에 관한 프로세스와 밀접한 관련이 있다. 그에 비해 '데이터 표현'은 상대적으로 구축 데이터를 본격적으로 보여 주기 위한 프로세스와 관련성이 깊은데, 데이터 시각화와 같은 분야가 그에 해당한다.

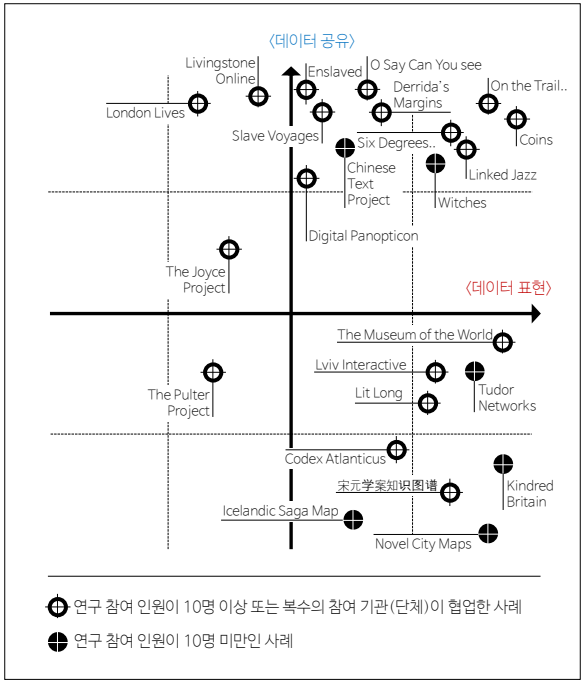
좌표평면에서 상대적으로 우측 상단에 가깝게 배치될수록 '공유'와 '표현'을 위한 데이터 처리 과정이 정교하게 진행된 연구 사례임을 의미하며, 좌측 하단에 가깝게 매핑될수록 데이터 처리 과정 및 결과물 형식에 있어 아쉬움이 많이 남는 사례에 해당한다.

3.3. 검토 결과와 전반적 해석[8]

[그림 2]는 앞서 제시한 좌표평면에 24건의 사례를 매핑한 결과이다. 24개 가운데 1/4에 해당하는 6개 사례는 참여 인원이 10명 미만인 소규모 연구이며, 나머지 3/4에 해당하는 18개 사례는 참여 인원이 10명 이상이거나 또는 2개 이상의 복수 기관(단체)이 일정 기간 이상의 협업을 진행한 경우다.

좌표평면상의 전반적 분포에 대해 살펴보자면, 우선 세로축을 기준으로 우측에 집중된 경향이 뚜렷하게 드러난다. 좌측에는 4건의 사례만 분포하는 반면, 나머지 20건의 사례가 우측에 집중적으로 배치되어 있다. 이와 같은 양상은 다수의 사례가 '데이터 공유'의 차원보다는 '데이터 표현'의 차원에서 두드러진다는 것으로 해석이 가능하다. 다만 데이터를 잘 '표현'하기 위해서는 데이터를 처리하는 과정에서의 정교한 가공과 표현된 데이터의 검증과 활용을 위한 데이터 '공유'의 안배 또한 중요한데, 20건의 사례

8 24개 연구 프로젝트 사례에 관한 소개가 본 연구의 목적이 아니며 또한 지면의 한계가 있기에, '3.3. 검토 결과와 전반적 해석', '3.4. 특기할 만한 요소 1', '3.5. 특기할 만한 요소 2' 장절에서 개별 사례에 관한 소개나 설명에 해당하는 내용을 기술하지 않았음을 미리 밝히는 바이다. 개별 프로젝트 사례에 관한 소개는 본 연구의 온라인 페이지(http://dh.aks.ac.kr/~red/wiki/index.php/OverseasDHResearchCases)에 담긴 내용을 통해 참고할 수 있다.



[그림 2] 24가지 해외 디지털 인문학 연구 사례를 좌표평면에 매핑한 결과

가운데 거의 절반인 11건은 상단에, 나머지 절반인 9건은 하단에 분포하는 양상이 그에 대한 시사점을 제공해 준다.

그에 더해 한 가지 인상적인 사실은 우측에 분포한 사례들 가운데 다수가 Digital Humanities Awards의 Visualization 분야에서 후보로 지명되거나 실제 수상한 연구라는 것이다. 〈Kindred Britain〉(2013), 〈Six Degrees of Francis Bacon〉(2017), 〈Coins〉(2018), 〈Witches〉(2019), 〈On the Trail of Antiquities〉(2019), 〈Tudor Networks〉(2020) 등이 그에 해당한다. 데이터 시각화 방면의 디지털 인문학 연구를 대표하는 많은 사례가 Digital Humanities Awards를 통해 소개되었음을 알 수 있다.

그리고 가로축을 기준으로 좌표평면 상단의 분포를 보면 '데이터 공유'와 관련된 한 가지 특징이 확인된다. 14개 사례 가운데 〈The Joyce Project〉

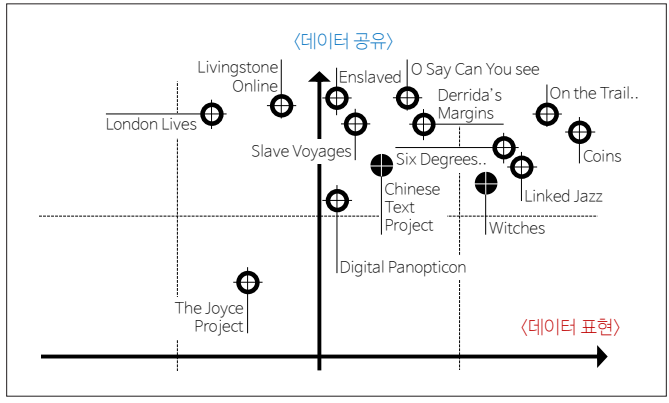
를 제외한 나머지 13개의 사례가 모두 '중간 이상'의 위치에 분포해 있다는 사실이 그것이다. 이와 같은 양상은 곧 '데이터 공유'에 초점을 둔 사례의 경우 대부분 웹 표준 형식의 데이터 가공 및 개방에 주력하였음을 알려 주는 단서에 해당한다.

전자지도 매핑에 초점을 둔 4가지 사례, 〈Witches〉, 〈Lviv Interactive〉, 〈Lit Long〉, 〈Icelandic Saga Map〉이 모두 흩어져 분포하고 있다는 것도 흥미로운 사실이다. 전자지도(Digital Atlas)라는 미디어, 즉 결과물의 공통적 형식만을 고려할 경우 좌표평면상 가까운 거리에 4가지 사례가 위치할 것이라는 유추가 가능하겠으나, 오히려 유관 데이터의 구체적 처리 방법에 따라 연구의 성격이 달라질 수 있다는 것을 생각하게끔 하는 결과다. 이와 관련해 유물 데이터 큐레이션에 초점을 둔 4가지 연구 사이의 위상도 마찬가지다. 〈On the Trail of Antiquities〉과 〈Coins〉이 우측 상단에 함께 있으며, 〈The Museum of the World〉가 우측 중단에 그리고 〈Codex Atlanticus〉가 우측 하단에 자리하고 있다. '큐레이션'이라는 동일한 목적을 지향함에도 불구하고, 유물 자원을 데이터로 옮겨 처리하는 과정에서의 문제의식과 결과물의 형식은 다양할 수 있음이 확인된다.

3.4. 특기할 만한 요소 1: 데이터 개방·공유와 협업을 위한 노력

[그림 3]에서 확인할 수 있듯이 좌표평면 좌측의 〈London Lives〉, 〈Livingstone Online〉 그리고 우측의 〈Digital Panopticon〉과 〈Derrida's Margins〉, 〈Six Degrees of Francis Bacon〉와, 〈Slave Voyages〉와 〈Enslaved〉, 〈O Say Can You See〉 등은 상대적으로 데이터의 개방과 공유에 초점을 둔 연구 사례에 해당한다.

〈London Lives〉는 기존 런던의 8개 아카이브에 소장된 24만 개 이상의 원고 및 인쇄물과 여타 프로젝트에서 구축한 15개의 데이터셋을 기반으



[그림 3] 해외 디지털 인문학 연구 사례의 좌표평면 매핑 결과: 가로축 상단

로 335만 개 이상의 인물 정보를 구축하였는데,[9] 사용자는 동일 인물과 관련된 기록을 연결하고, 가장 잘 기록된 개인의 전기를 엮어 나갈 수 있으며, 〈London Lives〉 웹 사이트에 구글 계정을 등록해 개인의 workspace[10]에서 해당 작업을 가능하도록 하였다. 검색 엔진과 웹 사이트를 개발해 키워드와 문구뿐만 아니라 XML 태깅이 된 텍스트(XML tagged text)로 다각적 검색을 용이하게 하였다. 해당 프로젝트를 통해 구축된 XML 데이터는 University of Sheffield data repository(ORDA)에서 크리에이티브 커먼즈 저작자 표시 비상업(CC-BY-NC) 라이센스에 따라 다운로드할 수 있다.[11]

〈Digital Panopticon〉은 영국과 호주에서 보유하고 있는 '죄수의 삶과 관련된 계보, 생체 인식 및 형사 사법 데이터셋'(genealogical, biometric and criminal justice datasets) 약 50개를 결합하였다. 이를 바탕으로 여러 기록을 효과적으로 연계하기 위해 데이터를 고도로 구조화하고 표준화하는 작업을 진행하였다. 데이터셋 연결을 위한 주요 정보(특히 이름, 나이, 생일, 성별, 직

9 외부 데이터셋의 활용에 관한 정보는 다음 웹페이지를 참고할 수 있다. 참고 페이지 URL: https://www.londonlives.org/static/Datasets.jsp

10 URL: https://www.londonlives.org/search.jsp?form=workspace

11 URL: https://figshare.com/articles/London_Lives_XML_Data/4797829

업, 주소, 출생지, 범죄, 재판일 및 재판장소, 형량)에 관한 치밀한 검토가 선행되었다. 그것을 바탕으로 node.js를 활용해 데이터 자동 연결을 진행하였으며, 동시에 MySQL과 Java servlet 기반의 수동 연결 작업 또한 함께 수행되었다. 결과물 데이터셋은 조건 검색을 매개로 JSON이나 TSV 포맷으로 다운로드 가능하며, 일부 데이터셋의 경우 CSV 포맷으로 다운로드할 수 있다.[12]

〈Livingstone Online〉은 2004년부터 현재에 이르기까지 총 3단계(2005~09, 2010~12, 2013~현재)의 개발 과정과 3차례에 걸친 〈Livingstone Spectral Imaging Project(2009~10, 2010~12, 2013~현재)〉를 통해 작업이 수행되었으며, 고해상도 원고 이미지와 편집된 필사본(1만 5,000개의 이미지 및 780개의 필사본)을 대상으로 디지털 컬렉션을 구축하였다. 브라우저를 통해 '소장기관', '작성연대', '기록장소' 등을 매개로 한 탐색이 가능하며, 자료에 관한 에세이와 교육 및 학습 자료로서의 리소스가 제공되고 있다. 데이터의 경우, 전체 기록에 대한 이미지 파일, 전사본의 PDF 파일, 전사본의 TEI 파일,[13] MODS(Metadata Object Description Schema, 메타데이터 개체 설명 스키마) XML 문서[14]를 제공받을 수 있다.

〈Derrida's Margins〉은 알제리 출신 프랑스 철학자 자크 데리다(Jacques Derrida, 1930~2004)의 개인 서재에 있던 약 1만 3,800권의 책과 기타 자료를 대상으로, 데리다가 책을 읽으며 육필로 써 내려간 주석과 밑줄, 인용구, 참고자료 표기 등을 데이터화함으로써 데리다의 고민과 생각의 흔적에 관한 탐색을 목적으로 한 프로젝트이다. 데리다의 사후 그가 소장했던 도서 및 자료는 2015년 프린스턴 대학교 도서관의 파이어스톤 도서관(Firestone

12 데이터셋 URL: https://www.digitalpanopticon.org/Access_to_Digital_Panopticon_Data

13 TEI 전사 파일의 경우, TEI P5 인코딩 지침에 따라 XML로 모든 전사를 제작하였다. 792개의 TEI 전체파일을 다운로드할 수 있으며, 또한, 코딩 매뉴얼, 전사 템플릿, ODD 및 RNG 스키마를 포함하는 전체 TEI 전사 자료를 제공하고 있다(관련 URL: https://livingstoneonline.org/behind-the-scenes/practices-standards-and-arrangements).

14 MODS는 버전3에 따라 제작한 3,123개의 XML 파일을 제공한다(관련 URL: https://livingstoneonline.org/behind-the-scenes/practices-standards-and-arrangements).

Library)에서 인수하였고, 현재 프린스턴 도서관의 'Rare Book Division, Department of Special Collections'에서 소장·관리 중이다. 유관 데이터는 '주석', '주석 삽입 페이지', '소장 도서', '참고 문헌' 4가지 범주를 기반으로 CSV, JSON 포맷으로 다운로드할 수 있다.[15]

〈Six Degrees of Francis Bacon〉은 베이컨, 셰익스피어, 뉴턴 등과 같은 유명 인물 사이의 사적 관계를 추적하기 위한 16~17세기 영국의 소셜 네트워크 구현 프로젝트로서, 현재(2022년 4월 기준) 1만 5,870명의 식별 인물과 그들 사이의 약 17만 1,558개의 관계 데이터를 제공하고 있다. 해당 프로젝트의 특수성은, 데이터 구축에 있어서 사용자가 참여할 수 있는 환경을 구현하였다는 점이다. 결과물 웹사이트에 접속할 수 있는 사용자는 누구나 유관 인물과 관계 데이터를 추가하고 수정·비평할 수 있다.[16] 그와 같은 환경을 토대로 유관 분야에 종사하는 여러 연구자의 작업이 종합될 수 있고, 기존 연구 결과물에 새로운 연구 산출물을 연결함으로써 상호 운용적인 협력 연구가 가능하다. 데이터는, 웹사이트의 Download 섹션에서 인물, 관계, 관계 타입, 그룹, 그룹 멤버십 등에 대한 항목을 CSV 포맷으로 다운로드할 수 있다.

〈Slave Voyages〉는 3개의 데이터베이스로 구성된다. 1514년에서 1866년 사이 대서양을 건너 노예가 된 아프리카인을 강제로 수송한 3만 6,000회 이상의 항해에 대한 정보를 담은 'Trans-Atlantic Slave Trade DB', 아메리카 대륙 내에서 노예를 인신매매한 1만 1,000건 이상의 해상 항해에 대한 정보를 담고 있는 'Intra-Americanslave Trade DB', 1808~1862년 사이 노예무역 순찰대에 의해 포획된 19세기 노예선이나 아프리카 무역 현장에서 가져

15 데이터셋 URL: https://dataspace.princeton.edu/handle/88435/dsp01gf06g579z

16 연구 결과물 웹사이트(http://sixdegreesoffrancisbacon.com)에 접속해 계정 가입 후, 'CONTRIBUTE' 모드를 활성화하여 Node 생성, Link 생성, Group 할당이 가능하다. 데이터 구축에 기여하고자 제출한 내용은 SDFB 큐레이터의 검증을 받아야 하며, 검증 전에는 승인되지 않은 기여(unapproved contribution)로 표시된다.

온 9만 1,491명의 아프리카인에 대한 세부 정보를 담고 있는 'People of the African Slave Trade DB'가 그것이다. 이름, 나이, 성별, 출신, 승선 및 하선 장소 등에 관한 정보뿐만 아니라, 일부의 경우 노예 운송, 구매 및 판매한 사람들에 대한 기록 또한 찾을 수 있다. 3가지 DB는 모두 데이터 소스를 공개하고 있으며, SAV나 CSV 포맷으로 데이터를 다운로드할 수 있다.[17]

〈Enslaved〉는 원활하고 효율적인 데이터셋 수집 및 검토를 위해, 유관 데이터셋 대상 기사를 발행하는 디지털 학술지 *Journal of Slavery and Data Preservation*(ISSN 2691-297X)[18] 플랫폼을 구축하였다. 유관 데이터를 기고하고자 하는 개인 또는 기관은 CSV 또는 xlsx 포맷으로 데이터셋을 제출하고 그에 대한 description을 제출하는데, 이후 엄격한 Peer Review 및 사내 검토를 거쳐 데이터와 해당 description 게시 및 보존이 이루어진다. 또한 개별 프로젝트 데이터셋을 다운로드할 수 있으며, 데이터 기고자에 대한 Ontology, Metadata, Controlled Vocabularies, Linked Open Data 등이 제공된다. Linked Open Data(LOD)로 구축된 데이터 활용 시스템은 'Enslaved.org'에서 연결된 모든 프로젝트 데이터에 대한 통합 검색 및 탐색을 용이하게 하는데, 현재 'Enslaved.org' 내에서 사용할 수 있는 Linked Data 프로젝트는 총 16개이다. 브라우저상에서 노예무역에 관련된 인물·사건·장소·sources를 다양한 조건으로 상세하게 검색할 수 있으며, RDF/XML, Turtle, RDF/JSON 포맷의 데이터를 다운로드할 수 있다.

〈O Say Can You See〉는 총 509개의 자유 청원서에서 4,800명 이상의

---

17 Trans-Atlantic Slave Trade DB의 데이터는 SPSS 소프트웨어와 함께 사용하도록 형식화된 SAV 포맷의 파일로 다운로드 가능하다(데이터셋 URL: https://www.slavevoyages.org/voyage/downloads#full-versions-of-the-trans-atlantic-slave-trade-database/0/en/). Intra-Americanslave Trade DB의 데이터는 CSV 포맷으로 제공된다(데이터셋 URL: https://www.slavevoyages.org/american/downloads#intra-american-database-downloads/0/en/). People of the African Slave Trade DB의 데이터도 CSV 포맷으로 제공된다(데이터셋 URL: https://www.slavevoyages.org/resources/downloads#african-names-database-downloads/0/en/).

18 *Journal of Slavery and Data Preservation* URL: https://jsdp.enslaved.org/

개별 인물과 5만 5,000개의 관계를 발견해 정리하였다. 다종다양한 기록물에 흩어져 있는 개별 인물 정보를 종합적으로 정리하기 위해, 표준적이고 확장 가능한 데이터 포맷인 TEI 형식으로 데이터를 구축하였으며, 개별 정보마다 참고 자료원의 URI를 속성으로 기술하였다. 데이터는 기본적으로 'Data Download/Query' 항에서 제공되며,[19] 'Relationship description file(.ttl): 관계 자체에 대한 파일', 'Relationship ontology file(.owl): 관계 유형 정의 파일(온톨로지)', 'Raw relationships data with case sources(.csv): 관계가 정의된 사건들에 대한 인용을 포함한 파일' 등 RDF로 구축하여 CSV 포맷으로 데이터를 다운로드할 수 있도록 하였다.

상술한 〈London Lives〉, 〈Digital Panopticon〉, 〈Livingstone Online〉, 〈Derrida's Margins〉, 〈Six Degrees of Francis Bacon〉, 〈Slave Voyages〉, 〈Enslaved〉, 〈O Say Can You See〉 사례를 살펴볼 경우, 공통적으로 웹 표준 형식의 데이터를 구축해 외부에 개방·공유하기 위한 데이터 처리에 많은 노력을 기울였다는 것을 알 수 있다.

〈London Lives〉의 데이터를 기초 자원으로 활용해 구현된 〈Digital Panopticon〉의 경우나, 〈Slave Voyages〉의 데이터를 확장·포괄하는 〈Enslaved〉의 경우를 볼 경우, 인문학 자원을 대상으로 한 데이터셋이 디지털 환경에서의 '접속'과 '연결'을 매개로 확장·심화해 나갈 수 있는 가능성이 확인된다. 그리고 그러한 데이터의 개방적 구축·활용에 있어서 웹(web)이라는 공간은 해당 작업에 참여할 수 있는 인원을 폭넓게 확대할 수 있다. 웹 기반 크라우드 소싱을 인문학 데이터 구축 채널에 실제 도입한 사례로서 〈Six Degrees of Francis Bacon〉과 〈Linked Jazz〉[20]의 경우와 일련의 데

19 데이터 URL: https://earlywashingtondc.org/about/data

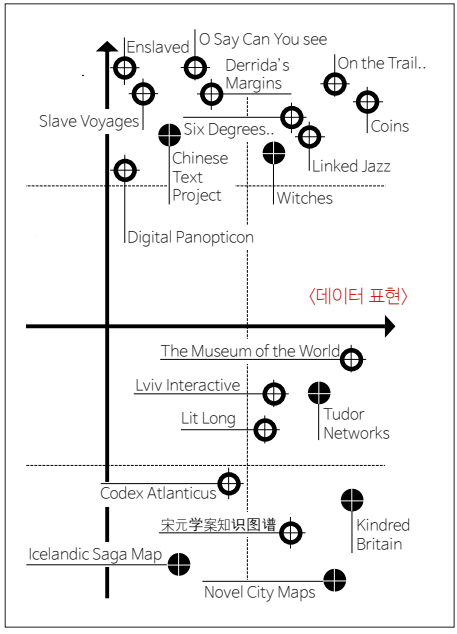
20 〈Linked Jazz〉 프로젝트에서 크라우드 소싱을 활용한 사례는 2022년 6월 10일자로 발간 예정인 『한국학』 여름호(통권 167호)에 수록될 졸고 「인문학술 데이터 프로세싱에 관한 시론」에 자세히 서술하였기에, 지면 활용의 경제성을 고려해 이 글에서는 생략하였음을 밝히는 바이다.

이터 프로세싱을 효과적으로 진행하기 위한 플랫폼으로서 디지털 학술지를 창간한 〈Enslaved〉의 경우는, 디지털 환경의 개방성을 적극적으로 활용하고자 한 새로운 인문학 연구 방법으로서, 인문학 자원을 표준적인 데이터 셋으로 구축해 개방·공유하는 데이터 프로세싱의 학술적 의미를 여실히 보여 주는 사례라 하겠다.

3.5. 특기할 만한 요소 2: 데이터 시각화를 통한 지식 표현의 시도

앞서 '3.4.절'의 경우 '데이터의 공유'에 초점을 두고 그와 유관한 여러 사례를 세부적으로 검토하였다면, 이번 장에서는 '데이터의 표현'에 주력한 여러 사례의 세부적 문제의식과 그로부터 공통으로 드러나는 학술적 시사점에 관해 정리해 보고자 한다. 실제 데이터 프로세싱 과정에 있어서는 '공유'와 '표현'으로 언명되는 지점을 엄격하게 구분하기 어렵지만, 구현된 결과물의 형식을 기준으로 할 때 그 절차적 지향이 상이한 것으로 이해할 수 있다. [그림 4]에서 드러나듯이 좌표평면 상단의 〈On the Trail of Antiquities〉, 〈Coins〉, 〈Six Degrees of Francis Bacon〉, 〈Witches〉 그리고 하단의 〈The Museum of the World〉와 〈Tudor Networks〉, 〈Kindred Britain〉, 〈Novel City Maps〉와 〈宋元学案知识图谱〉 등은 상대적으로 데이터의 시각적 표현에 초점을 둔 연구 사례에 해당한다.

〈On the Trail of Antiquities〉는 골동품 판매 카탈로그와 경매인 아카이브를 포함한 다양한 기록물을 바탕으로 개별 예술작품의 식별 정보, 소장처 정보, 예술품 설명, 생성 정보, 수집 정보, 관련 문서 등을 데이터로 구축해, 개별 미술품의 이동 경로를 동적으로 나타내는 시각화 결과물을 구현하였다. 지도에 미술품이 생성/발굴/판매/소장된 장소를 매핑하고 개별 장소들을 연결함으로써 미술품의 이동 경로를 살펴볼 수 있으며, 각각의 장소와 관련된 미술품의 메타데이터(생산 장소, 날짜, 속성, 크기 등/발견 장소 및 날짜/경매장소, 경매날짜, 딜러, 구매자, 판매가격/소장 장소/날짜) 및 연계 정보를 확인할



[그림 4] 해외 디지털 인문학 연구 사례의 좌표평면 매핑 결과: 세로축 우측

수 있다는 것이 특징이다.[21]

〈Coins〉는 뮌츠카비넷 소장 동전 약 2만 6,000개를 대상으로 구축한 다양한 속성(발견 지역, 주조 국가, 주조 공간, 주조 시기, 재질, 무게, 지름 등)을 바탕으로 웹브라우저를 통해 동전을 시각적으로 탐색하게끔 해주는 결과물을 구현하였다. 개별 동전을 클릭하면 동전의 속성을 간단하게 제공할 뿐만 아니라, 뮌츠카비넷 박물관 사이트를 연결함으로써 해당 동전에 대한 구체적인 정보에 접근할 수 있도록 해 준다. 또한 브라우저상에서 두 가지 시각화 레이아웃을 제공함으로써 동전이 내포한 여러 속성을 복합적으로 출력·비

21 데이터 시각화 결과물은 프로젝트 협업 그룹 내부의 'researchers and digital research service'팀에서 디자인하였으며, 구체적 개발은 데이터 저널리즘 전문 기업인 WEDODATA (https://wedodata.fr/)에서 응용 프로그램을 제작하였다.

교해 보는 것이 가능하다.[22]

〈Witches〉는 1563년부터 1736년 스코틀랜드에서 마녀로 기소된 3,837명의 인물들에 대한 정보를 구축한 〈Survey of Scottish Witchcraft Database〉[23]를 기반으로 교구에서 고발당한 마녀들의 거주지, 구금 장소, 재판 장소, 사망 장소 등의 지리 정보를 조사하여 전자 지도에 매핑하고 Wikidata에 연계하는 결과물을 구현하였다. 1550년부터 1750년 사이의 기간을 타임라인으로 구현하고, 마녀 3,141명에 대한 공간 데이터(거주지, 구금 장소, 재판 장소, 사망 장소)를 지도상에 매핑함으로써, 세분화된 기타 속성(성별/계층/직업 등)을 토대로, 개별 마커를 클릭할 때 그에 대한 간단한 정보와 〈Survey of Scottish Witchcraft Database〉 연계 정보 및 〈위키피디아〉 페이지 링크를 열어 볼 수 있도록 하였다.[24]

〈The Museum of the World〉는 대영박물관에 소장된 기원전 200만 년 전부터 20세기까지의 유물 가운데 약 320여 점을 대상으로 하여 WebGL (Web Graphics Library) 기반의 기술로 시각적으로 구현한 3차원 디지털 박물관이다. 공간으로서의 다섯 대륙(아프리카, 아메리카, 아시아, 유럽, 오세아니아)을 반영한 가로축과, 시간으로서의 연표를 반영한 세로축을 기준으로 각 유물을 배치하였다. 개별 유물을 클릭하면 시공간을 넘어 동일한 문화에서 형성된 유물들 사이의 연관 관계가 선으로 표현되며, 개별 유물마다 텍스트 및

---

22 데이터 시각화 결과물은 Pixi.js, D3.js, React.js 등의 JavaScript Library를 활용해 구현되었다. 관련해서 다음의 내용을 참고할 수 있다. "The visualization concept was realized as a web-based prototype using standard web technologies and several JavaScript libraries: D3.js for calculation of positions such as normal distribution and pack layout and for animations, Pixi.js for rendering the coins on the screen using WebGL, and React for rendering DOM elements." - Gortana, F. and von Tenspolde, F. and Guhlmann, D. and Dörk, M. (2018), Off the Grid: Visualizing a Numismatic Collection as Dynamic Piles and Streams, Open Library of Humanities 4(2), p. 16.

23 〈Survey of Scottish Witchcraft Database〉 URL: http://www.shca.ed.ac.uk/witches/

24 데이터 시각화 결과물은 Leaflet.js, ArcGIS Online, StoryMap.js 등의 웹 GIS Library를 활용해 구현되었다.

오디오 설명, 유물의 위치 정보 및 기타 관련 유물에 대한 정보에 접근할 수 있도록 하였다.[25]

〈Tudor Networks〉는 영국 국가 문서 보관소(United Kingdom State Papers archive)에 남아 있는 헨리 8세 재위부터 엘리자베스 1세까지의 시기 1509~1603년에 걸쳐 2만 424명의 사람들이 주고받은 12만 3,850개의 편지를 데이터로 편찬하여 튜더 왕조의 커뮤니케이션 네트워크를 시각적으로 구현한 결과물이다. 메인 시각화 구현 결과물의 세로축은 각 인물이 보낸 모든 편지 발신 연도의 평균값으로, 두 명 이상의 사람들로부터 편지를 보내거나 받은 사람들을 검은 막대점으로 표현하였다. 서신의 양이 많을수록 막대점의 길이 및 굵기가 커지며, 막대점을 클릭하면 해당 인물과 편지를 주고받은 인물들의 막대점과 연결된다. 보낸 편지가 많을수록 막대점과의 연결선 색상은 빨갛게, 받은 편지가 많을수록 연결선 색상은 파랗게 나타나며, 주고받은 편지의 양이 비슷할 경우 보라색으로 나타난다. 막대점을 클릭하면 막대점에 해당되는 인물과 그가 타인과 주고받은 편지를 시간(time), 공간(location)별로 시각화한 추가 링크가 제공되며, 해당 링크에 접속하면 해당 편지들의 수발신 시간과 공간 정보를 기준으로 전체 편지의 외연과 개별 편지에 관한 정보를 확인할 수 있다.[26]

〈Kindred Britain〉은 영국인 또는 영국인과 가족 관계인 약 3만 명(남자 1만 5,000여 명, 여자 1만 5,000여 명)의 인물을 대상으로 데이터베이스를 구축하였으며, 해당 인물들 사이의 사적 네트워크를 통해 생성되는 수억만 건 이상의 상이한 경로 정보를 제공한다. 시각화 구현 결과물의 특징은 데이터를 네트워크/타임라인/전자지도 3가지 형태로 종합하였다는 점이다. 네트워크

---

25 데이터 시각화 결과물은 WebGL(Web Graphics Library) 기술 기반 JavaScript 3D Library인 Three.js와 JavaScript 기반 오픈 소스 프론트엔드 web framework인 Angular.js를 활용해 구현되었다.

26 데이터 시각화 결과물은 하버드대학교 metaLAB 소속 디자이너 Kim Albrecht에 의해 구현되었으며, D3.js와 TopoJSON을 활용하였다.

는 사람들 사이의 혈연·혼인 관계를 보여 주는 것으로, 인물에 해당하는 개별 노드를 클릭할 때마다 가족 연결 경로가 나타난다. 타임라인은 여타 가족 구성원이 겪은 큰 사건이나 역사적으로 중요하게 여겨지는 사건을 매개로 개인의 삶을 맥락화하여 보여 준다. 그와 관련해 사용자는 개인이 경험한 비극의 정도와 같은 추상적 성격의 데이터 값을 필터링할 수 있다. 전자 지도는 인물이나 가족과 관련된 국가 또는 대영 제국의 일부를 강조 표시한다.[27]

〈Novel City Maps〉는 독일 소설 가운데 각기 다른 시대의 베를린을 배경으로 한 유명 작품 3편[Alfred Döblin, *Berlin Alexanderplatz*(1929), Hans Fallada, *Alone in Berlin*(1947), Sven Regener, *Berlin Blues*(2001)]을 대상으로, 세 소설에 등장한 베를린 공간을 개별 데이터로 만들고, 해당 공간이 소설 내에서 언급되는 구절의 위치를 함께 데이터로 구축하였다. 공간은 실재하는 위치와 픽션으로서의 장소가 혼재되어 있는데, Narrative View와 Map View 두 가지 형식의 디지털 지도에 공간 데이터를 매핑하였다. 작가, 시대, 서사에 따라 소설 속 도시에 대한 인상이 어떻게 달라지는지에 관한 정보는 꼭 문학 연구자가 아니라 하더라도 일반 독자들도 궁금해할 수 있는 정보이며, '공간'을 매개로 소설을 새롭게 탐색하고자 하는 시도는 학술 연구의 외연을 넓히는 접근이라는 측면에서 풍부한 영감을 준다.[28]

〈宋元学案知识图谱〉는 송나라에서 원나라에 걸친 시기에 활약한 중국 유학자 2,700여 명에 관한 방대한 인적 정보와 저술 정보를 수록한 문헌 『송원학안』(宋元學案)을 대상으로, 의미적으로 연결된 데이터를 구축해 그것을 다양한 형식의 시각화 결과물로 구현한 사례이다. 가계·사승관계·붕우

---

27 데이터 시각화 결과물은 디자이너 Scott Murray에 의해 구현되었으며, 기본적으로 D3.js를 활용하였다.

28 데이터 시각화 결과물은 독일 포츠담 응용과학대학 'Urban Complexity Lab' 소속의 Jan Erik Stange에 의해 구현되었으며, 기본적으로 D3.js를 활용하였다(참고 URL: https://uclab.fh-potsdam.de/projects/novel-city-maps/).

관계 등 인적관계를 매개로 표현한 인물 관계도, 시간(시대)를 기준으로 활동한 학자들의 규모를 좌표평면상의 그래프로 나타낸 학술흐름도, 연원에 관한 정보를 확인할 수 있는 사승관계도, 학파 간 선후 관계와 계승 정보를 나타낸 생키다이어그램(Sankey Diagram) 등 다양한 형식의 시각화 결과물을 보여준다는 점이 특징이다.[29]

상술한 〈On the Trail of Antiquities〉, 〈Coins〉, 〈Witches〉, 〈The Museum of the World〉, 〈Tudor Networks〉, 〈Kindred Britain〉, 〈Novel City Maps〉, 〈宋元学案知识图谱〉, 〈Six Degrees of Francis Bacon〉[30] 사례를 살펴볼 경우, 모두 데이터 처리 과정에서 각각의 인문학적 문제의식을 효과적으로 표현하기 위해 다양한 맥락의 JavaScript Library를 활용했다는 것을 확인할 수 있다.

예컨대 〈Witches〉, 〈Tudor Networks〉, 〈宋元学案知识图谱〉는 전근대기 고문헌 자료에 담긴 복잡한 정보를 어떻게 효과적으로 재현할 것인가에 주목했다는 점에서, 역사 문헌을 대상으로 한 전통적 인문학 연구가 데이터 시각화 작업과 만나는 지점을 간접적으로 보여 주는 사례에 해당한다.

〈On the Trail of Antiquities〉, 〈Coins〉, 〈The Museum of the World〉는 박물관이 소장한 유물을 대상으로 한 데이터 큐레이션 사례로서 시각적 표현의 규모와 지향이 여타 디지털 인문학 연구 사례와는 다르다. 이와 관련해 디지털 환경의 저변 확대로 인해 소위 '강단'과 '이론'으로 대표되는 대학(원)과 '수장고'와 '자료'로 대표되는 도서관·박물관 등 아카이브 기관의

---

29 현재(2022년 4월 기준) 정보가 출력되지 않는 경우가 많아 시각화 구현 결과물의 완성도가 높다고 보기는 어렵다. 뿐만 아니라 의미 검색 페이지(https://syxa.pkudh.org/semantic)에서 검색이 온전히 되지 않는 정보들 또한 발견된다. 시각화 구현 결과물과 별개로 기초 데이터베이스 구축이 완결되지 않은 것으로 짐작되는데, 공개된 정보가 없어 그에 대한 내용을 정확히 알기는 어렵다.

30 〈Six Degrees of Francis Bacon〉의 데이터 시각화 결과물은 기본적으로 D3.js(forcelayout)와 JavaScript 기반 오픈 소스 프론트엔드 web framework인 Angular.js를 활용해 구현되었다.

역할 구분에 관한 인식이 옅어져 가는 최근의 학술 환경을 고려할 때, 원자료에 관한 정보를 어떻게 다루고 보여 줄 것인가에 대한 큐레이션의 관점은 인문학 연구의 많은 지점에서 새로운 통찰(insight)을 던져 준다.

대상이 복합적으로 내포한 정보를 입체적으로 다룰 수 있어야 한다는 통합적 시각에 입각할 경우, 인문학의 제반 문제의식이 모두 데이터 기반 큐레이션 연구의 대상이 될 수 있다. 〈Kindred Britain〉은 그와 같은 문제의식에 입각해 추상적 성격의 인문 지식을 대상으로 데이터 큐레이션을 시도한 본격적 연구 사례라 해야 할 것이다.

4. 맺음말

디지털 인문학(Digital Humanities)을 표방하는 수많은 연구가 세계 여러 곳에서 꾸준히 기획·진행되고 있다. 여타 국가와 비교할 때 한국의 경우 인문학 자원을 대상으로 한 디지털 리소스가 적지 않은 상황이며, 단순한 플레인 텍스트(Plain Text) 상태를 넘어 '인물'이나 '공간'과 같은 고유한 대상에 관해 마크업 처리가 된 텍스트 데이터(XML tagged text) 또한 적지 않은 편이다. 그러나 그와 같이 정교하게 마크업 처리된 텍스트 데이터는 전근대기 분야의 일부 자료에 한정된 경향이 뚜렷하며, 그 외의 경우 단순한 형식의 메타데이터만을 기재하고 있는 경우가 보편적이다.

무엇보다 가장 큰 문제는, 인문학 연구자들 사이에 인문학 데이터를 가공·처리하는 경험과 이해의 채널이 열려 있지 않다는 사실이다. 최근 10년 사이 급격히 발달한 데이터 처리 기술을 적용한 혁신적 연구를 인문학 분야에서 시도하기가 쉽지 않은 이유이다. 이러한 문제의 근본적 원인을 짚기 위해서는 데이터 처리 기술이나 발달된 웹 환경과 같은 디지털 테크놀로지의 시점에서 응시하기보다, 원천 대상으로서의 인문학 자원을 데이터로 가공하는 입장에서 바라볼 필요가 있다.

예컨대 첨단의 프로그래밍 기술이나 데이터 가공 알고리즘이 준비되어 있다 하더라도 그것을 적용할 수 있는 기초 자원으로서 정교한 데이터가 준비되지 않는다면, 의미 있는 결과물을 얻기 어렵다. 약 20년 전 여러 기관을 중심으로 인문학 자원을 대상으로 한 데이터 구축을 고민하고 실행했던 것처럼, 현 시점에서 첨단의 기술을 적용할 수 있는 인문학 데이터의 형식을 디자인하고 그에 의거한 데이터 고도화 또는 데이터베이스 리뉴얼 연구가 소규모라도 이루어질 필요가 있다.

그와 관련해 최근 몇 년 사이 국내에서도 〈삼일운동 데이터베이스〉[31]나 〈지암일기 데이터 아카이브〉[32]와 같이 인문학 자원을 대상으로 표준적 형식의 데이터를 구축해 웹상에서 공유하고 유관 지식을 시각적으로 표현하는 새로운 형식의 연구가 이루어지긴 하였으나, 그저 이례적 성격의 연구 프로젝트 시도로만 여겨질 뿐, 인문학 연구에 데이터 처리 기술을 본격적으로 적용하는 연구는 여전히 그 저변 확장을 기대하기가 어려운 상황이다. 그와 같은 기획이 구체적으로 실행되기 위해서는 디지털 환경에 대한 능동적 시각을 인문학 연구자 스스로 갖추어야 함에도, 기술 환경에 대한 기초 이해를 토대로 그것을 인문학에 적극적으로 활용해야 한다는 생각을 지닌 연구자를 찾기란 여전히 쉽지 않다. 인문학은 본질적으로 '아날로그 환경'에 토대하는 것이지 그와 대조적인 '디지털 환경'과는 멀다 여기는 구습(舊習)이 크게 작용하는 탓이다.

최신의 데이터 처리 기술을 활용해 이루어진 해외의 첨단 인문학 연구 사례들을 검토·정리하고 그 시사점을 살펴보는 데 이 글이 주력한 이유 가운데 하나는, 바로 그러한 전통 인문학과 디지털 인문학 사이에 있다고 여겨지는 눈에 보이지 않는 단절이 실체가 없는 것임을 밝히고자 함이기도 하다. '최신의 기술', '첨단의 연구'라는 표현을 쓰긴 하였으나, 실제 인문학

31 〈삼일운동 데이터베이스〉 URL: http://db.history.go.kr/samil/

32 〈지암일기 데이터 아카이브〉 URL: http://jiamdiary.info/

연구 자원과 데이터 처리 기술이 만나는 지점의 중추적 고민은 '근본'으로서 인문학적 문제의식이다. 그리고 '공유'와 '표현'으로 대표되는 데이터 프로세싱의 일환은 발달한 테크놀로지의 우위를 드러내기 위함이 아니라, 인문학의 근본적 문제의식을 디지털 환경에서 효과적으로 나타내고 전달하고자 하는 데 있다.

인문학 데이터 프로세싱의 시도는 '디지털에 의한 인문학'이 아니라 '디지털을 통한 인문학'으로 이해해야 한다. 인문학적 문제의식이 깊은 연구자들이 중심이 되어 테크놀로지에 능숙한 연구자가 참여하는 형식으로 인문학 데이터를 함께 디자인하고 구축·공유·표현하는 연구 프로세스가 이루어질 필요가 있다. 국내의 디지털 인문학 인프라는 여전히 열악하지만, 향후 개선과 발전이 이루어질 것임을 기대하며, 인문학 연구와 데이터 처리 기술이 만나는 접점으로서 디지털 인문학의 실제를 알리는 데 이 글이 기여하는 바가 작게나마 있기를 바란다.

참고문헌

자료

「About」, The Journal of Slavery and Data Preservation 웹페이지, 2022.4.9. https://jsdp.enslaved.org/about/#undefined

「About」, O Say Can You See 웹페이지, 2022.4.9. https://earlywashingtondc.org/about

「ABOUT SIX DEGREES OF FRANCIS BACON」, Six Degrees of Francis Bacon 웹페이지, 2022.4.9. http://sixdegreesoffrancisbacon.com/about

「About The Project」, Witches 웹페이지, 2022.4.9. https://witches.is.ed.ac.uk/about

「Access to Digital Panopticon Data」, Digital Panopticon 웹페이지, 2022.4.9. https://www.digitalpanopticon.org/Access_to_Digital_Panopticon_Data

「Additional Datasets」, London Lives 웹페이지, 2022.4.9. https://www.londonlives.org/static/Datasets.jsp

「African Names Database」, Slave Voyages 웹페이지, 2022.4.9. https://www.slavevoyages.org/resources/downloads#african-names-database-downloads/0/en/

「Behind the scenes」, On the Trail of Antiquities 웹페이지, 2022.4.9. https://ventesdantiques.inha.fr/backstage.php#en

Casebooks 웹페이지, 2022.4.9. https://casebooks.lib.cam.ac.uk

China Biographical Database Project (CBDB) 웹페이지, 2022.4.9. https://projects.iq.harvard.edu/cbdb

Chines Text Project 웹페이지, 2022.4.9. https://ctext.org

Codex Atlanticus 웹페이지, 2022.4.9. http://www.codex-atlanticus.it

Coins 웹페이지, 2022.4.9. https://uclab.fh-potsdam.de/coins

CORRIE GOLDMAN(2013), 「New Stanford website takes a digital approach to an ancient topic: families」, Stanford 웹페이지, Stanford University, 2022.4.9. https://news.stanford.edu/news/2013/august/kindred-britain-database-082613.html

「Data」, Enslaved 웹페이지, 2022.4.9. https://enslaved.org/data/

Derrida's Margins 웹페이지, 2022.4.9. https://derridas-margins.princeton.edu

「Data Download & Query」, O Say Can You See 웹페이지, 2022.4.9. https://earlywashingtondc.org/about/data

「Derrida's Margins datasets」, DataSpace 웹사이트, 2022.4.9. https://dataspace.princeton.edu/handle/88435/dsp01gf06g579z

「DH Awards 2012 Nominations」, Digital Humanities Awards 웹페이지, 2022.4.9. http://dhawards.org/dhawards2012/nominations/

「DH Awards 2013 Nominations」, Digital Humanities Awards 웹페이지, 2022.4.9. http://dhawards.org/dhawards2013/nominations/

「DH Awards 2021 – Frequently Asked Questions」, Digital Humanities Awards 웹페이지, 2022.4.9. http://dhawards.org/dhawards2021/faqs/

Digital Humanities Awards 웹페이지, 2022.4.9. http://dhawards.org

「Digital Panopticon」, DHI. The University of Sheffield 웹페이지, 2022.4.9. https://www.dhi.ac.uk/projects/digital-panopticon/

Digital Panopticon 웹페이지, 2022.4.9. https://www.digitalpanopticon.org

Digital Panopticon Project Blog 웹페이지, 2022.4.9. https://blog.digitalpanopticon.org/

Emma Carroll(2019), 「WEEK 1 – AN INTRO TO THE WITCHCRAFT OF INFORMATION SERVICES AND WIKIDATA」, WITCHFINDER GENERAL – DATA VISUALISATION INTERNSHIP 웹페이지, The University of Edinburgh, 2022.4.9. https://blogs.ed.ac.uk/ecarroll3_witchcraft_visualisation/2019/06/10/week-one-an-intro-to-the-witchcraft-of-information-services-and-wikidata/

Emma Carroll(2019), 「WEEK 13 – THE FINAL WEEK OF WITCHFINDING」, WITCHFINDER GENERAL – DATA VISUALISATION INTERNSHIP 웹페이지, The University of Edinburgh, 2022.4.9. https://blogs.ed.ac.uk/ecarroll3_witchcraft_visualisation/2019/09/13/week-13-the-final-week-of-witchfinding/

Enslaved 웹페이지, 2022.4.9. https://enslaved.org

「Full versions of the Trans-Atlantic Slave Trade Database」, Slave Voyages 웹페이지, 2022.4.9. https://www.slavevoyages.org/american/downloads#intra-american-database-downloads/0/en/

Icelandic Saga Map 웹페이지, 2022.4.9. http://sagamap.hi.is/is

「Interactive witchcraft map」, The University of Edinburgh 웹페이지, 2022.4.9. https://www.ed.ac.uk/information-services/about/news/interactive-witchcraft-map

「JAN-ERIK STANGE」, Novel City Maps 웹페이지, 2022.4.9. https://uclab.fh-potsdam.de/people/jan-erik-stange/

Jan-Erik Stange and Marian Dörk(2016), 「Visualizing the spatiality in fictional narratives」, Urban Complexity Lab 웹페이지, 2022.4.9. https://uclab.fh-potsdam.de/wp/wp-content/uploads/vis4dh2016.pdf

Kindred Britain 웹페이지, 2022.4.9. https://kindred.stanford.edu

Linked Jazz 웹페이지, 2022.4.9. https://linkedjazz.org

Lit Long 웹페이지, 2022.5.5. https://litlong.org

Livingstone Online 웹페이지, 2022.04.09. https://livingstoneonline.org

「Livingstone Online」, World History Commons 웹페이지, 2022.4.9. https://worldhistorycommons.org/livingstone-online

「Livingstone Online: An Introduction」, Livingstone Online 웹페이지, 2022.4.9. https://livingstoneonline.org/about-this-site/livingstone-online-introduction

London Lives 웹페이지, 2022.4.9. https://www.londonlives.org

「London Lives」, DHI. The University of Sheffield 웹페이지, 2022.4.9. https://www.dhi.ac.uk/projects/london-lives/

「London Lives XML Data」, figshare 웹페이지, 2022.4.9. https://figshare.com/articles/London_Lives_XML_Data/4797829

Lviv Interactive 웹페이지, 2022.4.9. https://lia.lvivcenter.org

Mapping the Republic of Letters 웹페이지, 2022.4.9. http://republicofletters.stanford.edu

MARKUS 웹페이지, 2022.4.9. https://dh.chinese-empires.eu/markus

Nicholas Jenkins, 「Originating Kindred Britain」, Notes on Kindred Britain 웹페이지, 2022.4.9. https://kindred.stanford.edu/notes.html?section=originating

Notes on Kindred Britain 웹페이지, 2022.4.9. https://kindred.stanford.edu/notes.html

「NOVEL CITY MAPS」, Urban Complexity Lab 웹페이지, 2022.4.9. https://uclab.fh-potsdam.de/projects/novel-city-maps/

Novel City Maps 웹페이지, 2022.4.9. https://uclab.fh-potsdam.de/NCM

On the Trail of Antiquities 웹페이지, 2022.4.9. https://ventesdantiques.inha.fr

O Say Can You See 웹페이지, 2022.4.9. https://earlywashingtondc.org

「Practices, Standards, and Arrangements」, Livingstone Online 웹페이지, 2022.4.9. https://livingstoneonline.org/behind-the-scenes/practices-standards-and-arrangements

「Project History」, Enslaved 웹페이지, 2022.4.9. https://enslaved.org/projectHistory/
「Répertoire des ventes d'antiques en France au XIXe siècle」, AGORHA 웹페이지, 2022.4.9. https://agorha.inha.fr/ark:/54721/47
「Répertoire des ventes d'antiques en France au XIXe siècle」, INHA 웹페이지, Institut national d'histoire de l'art, 2022.4.9. https://www.inha.fr/fr/recherche/le-departement-des-etudes-et-de-la-recherche/domaines-de-recherche/histoire-de-l-art-antique-et-de-l-archeologie/repertoire-des-ventes-d-antiques.html
「semantic search」, 宋元学案知识图谱 웹페이지, 2022.4.9. https://syxa.pkudh.org/semantic
Shilo Rea(2015), 「Six Degrees of Francis Bacon Launches」, Carnegie Mellon University 웹페이지, Carnegie Mellon University, 2022.4.9. https://www.cmu.edu/news/stories/archives/2015/october/francis-bacon-launch.html
Six Degrees of Francis Bacon 웹페이지, 2022.4.9. http://sixdegreesoffrancisbacon.com
Slave Voyages 웹페이지, 2022.4.9. https://www.slavevoyages.org
The Journal of Slavery and Data Preservation 웹페이지, 2022.4.9. https://jsdp.enslaved.org/
The Joyce Project 웹페이지, 2022.4.9. http://www.joyceproject.com
「The Library of Jacques Derrida, Studio Series, 1686-2010」, Finding Aids 웹페이지, 2022.04.09. https://findingaids.princeton.edu/catalog/RBD1
「The Livingstone Online Code」, Livingstone Online 웹페이지, 2022.04.09. https://livingstoneonline.org/behind-the-scenes/the-livingstone-online-code
The Museum of the World 웹페이지, 2022.4.9. https://britishmuseum.withgoogle.com
The Pulter Project 웹페이지, 2022.4.9. http://pulterproject.northwestern.edu
The Survey of Scottish Witchcraft 웹페이지, 2022.4.9. http://www.shca.ed.ac.uk/witches/
Tudor Networks 웹페이지, 2022.4.9. http://tudornetworks.net
WEDODATA 웹페이지, 2022.5.14. https://wedodata.fr/
Witches 웹페이지, 2022.4.9. https://witches.is.ed.ac.uk
「Workspace」, London Lives 웹페이지, 2022.4.9. https://www.londonlives.org/search.jsp?form=workspace
삼일운동 데이터베이스 웹페이지, 2022.4.9. http://db.history.go.kr/samil/
조선시대 표류 기록 시각화 웹페이지, 2022.4.9. http://soundh.net/drift
지식공유연대 웹페이지, 2022.4.9. https://knowledgecommoning.org/
지암일기 데이터 아카이브 웹페이지, 2022.4.9. http://jiamdiary.info
한국사데이터베이스 웹페이지, 2022.4.9. http://db.history.go.kr/
宋元学案知识图谱 웹페이지, 2022.4.9. https://syxa.pkudh.org
「关于」, 宋元学案知识图谱 웹페이지, 2022.4.9. https://syxa.pkudh.org/about
「关系图谱」, 宋元学案知识图谱 웹페이지, 2022.4.9. https://syxa.pkudh.org/distant/

relation
「师承关系」, 宋元学案知识图谱 웹페이지, 2022.4.9. https://syxa.pkudh.org/distant/tree
「学术流变」, 宋元学案知识图谱 웹페이지, 2022.4.9. https://syxa.pkudh.org/distant/river
「学派传承」, 宋元学案知识图谱 웹페이지, 2022.4.9. https://syxa.pkudh.org/distant/school

논저

권윤경(2018), 「새로운 문필공화국을 향하여: 18세기 프랑스사 연구와 디지털인문학의 사례들」, 『역사학보』 240, 역사학회.
김바로(2014), 「해외 디지털 인문학 동향」, 『인문콘텐츠』 33, 인문콘텐츠학회.
김용수(2018), 「영미문학과 디지털인문학 미국 디지털 영문학 연구 동향」, 『비평과이론』 23, 한국비평이론학회.
김종대(2015), 「독일의 디지털 인문학 동향」, 『철학과 문화』 32, 한국외국어대학교 철학문화연구소.
류인태(2020), 「디지털 인문학은 인문학이다」, 『인문논총』 77(3), 서울대학교 인문학연구원.
문수현(2018), 「독일의 디지털 역사학 현황」, 『역사학보』 240, 역사학회.
박숙자(2020), 「학술지식은 커먼즈다: 지식공유연대와 OA 플랫폼」, 『문화과학』 101, 문화과학사.
박은재(2018), 「영국 디지털 역사학의 발전과 현황」, 『역사학보』 240, 역사학회.
서경숙(2015), 「디지털 인문학 교수법의 이론 및 실제: 영미문학을 중심으로」, 『인문콘텐츠』 38, 인문콘텐츠학회.
손정훈·김민규(2018), 「프랑스의 디지털인문학 동향과 정책」, 『프랑스문화예술연구』 64, 프랑스문화예술학회.
송인재(2018), 「대만 디지털인문학의 발자취와 진화」, 『역사학보』 240, 역사학회.
송인재(2016), 「동아시아 개념사와 디지털인문학의 만남」, 『개념과 소통』 18, 한림과학원.
이주영(2018), 「미국에서의 디지털 역사학 발전 과정과 최근의 경향」, 『역사학보』 240, 역사학회.
이지수(2022), 「대학도서관의 디지털 인문학 연구지원 가이드에 관한 연구: 미국 대학도서관 사례를 중심으로」, 『한국문헌정보학회지』 56, 한국문헌정보학회.
이지수(2017), 「디지털 인문학의 정의 및 연구 현황: 일본학 지식정보의 디지털화를 중심으로」, 『일본어문학』 74, 한국일본어문학회.
이지수·이혜은(2019), 「미국 대학의 디지털 인문학 교육 프로그램 연구」, 『정보관리학회지』 36, 한국정보관리학회.
정유경(2020), 「디지털 인문학 분야의 국내외 연구 동향 분석」, 『정보관리학회지』 37, 한국정보관리학회.
홍정욱(2015), 「디지털기술 전환 시대의 인문학」, 『인문콘텐츠』 38, 인문콘텐츠학회.

홍정욱·김기덕(2014), 「'2014 세계 디지털인문학' 학술대회 및 한국의 디지털인문학」, 『인문콘텐츠』 34, 인문콘텐츠학회.

Gortana, F., F. von Tenspolde, D. Guhlmann, and M. Dörk (2018), "Off the Grid: Visualizing a Numismatic Collection as Dynamic Piles and Streams," *Open Library of Humanities* 4(2).

Stange, Jan-Erik and Marian Dörk (2016), "Visualizing the Spatiality in Fictional Narratives," VIS4DH: 1st Workshop on Visualization for the Digital Humanities at IEEE VIS.

원고 접수일: 2022년 4월 19일, 심사 완료일: 2022년 5월 2일, 게재 확정일: 2022년 5월 5일

ABSTRACT

Sharing and Representation of Knowledge in Digital Humanities

Reviewing Overseas Digital Humanities Research Cases

Kim, Ji-sun*
Ryu, Intae**

In recent years, interest in digital humanities has grown steadily in Korea, but understanding of the context of constructing·using humanities data that is universally dealt with in the digital humanities field, as well as its academic significance, remains limited. The reasons for this are that it is difficult to organize at a glance the field collectively referred to as 'digital humanities' due to the variety of research formats in exist within. In addition, there are not many cases of digital humanities research in Korea and there are very few professional researchers in the field, as a result of which the channels that can convey the reality of digital humanities research are limited. Considering these points, this article attempts to organize the data processing process and result format of digital humanities projects by reviewing 24 representative cases focusing on sharing and representation of knowledge among overseas digital humanities research examples, as well as to clarify the academic significance

* (First Author) Lecturer, College of Liberal Arts, Korea University
** (Corresponding Author) Postdoctoral Fellow, Department of Korean Language and Literature, Sungkyunkwan University

contained therein. It is a common feature to create and open various data formats used in web environments, such as CSV and JSON, as well as machine-readable data represented by XML and RDF, while utilizing various formats of JavaScript libraries and Web Framework such as D3.js, Leaflet.js, JQuery and Angular.js. If each digital humanities research case discussed in this article is reviewed individually later, it can be used as a useful clue not only to examine the reality of digital humanities research, but also to specifically plan and carry out digital humanities research projects. At a time when we need to think seriously about where humanities research combined with the digital environment should go in detail, as well as how to accommodate the humanistic awareness and the understanding of digital technology that arises in the process, I hope this article will contribute in some small way.

Keywords Digital Humanities, Humanities Data, Open Data, Data Visualization