

I . 디지털 인문학이란 무엇인가?

디지털 인문학(Digital Humanities)이란 디지털로 표현하고 디지털로 소통하는 이 시대에 인문지식이 더욱 의미 있게 탐구되고 가치 있게 활용되도록 하려는 노력이다.¹⁾ 디지털 인문학의 함의는 무엇이며, 디지털 인문학 활동을 통해 얻고자 하는 것은 무엇인가?

I -1. 디지털 인문학의 함의

디지털 인문학이란 정보통신기술(Information & Communication Technology)의 도움을 받아 새로운 방식으로 수행하는 인문학 연구와 교육, 그리고 이와 관계된 창조적인 저작 활동을 일컫는 말이다. 이것은 전통적인 인문학의 주제를 계승하면서 연구 방법 면에서 디지털 기술을 활용하는 연구, 그리고 예전에는 가능하지 않았지만 컴퓨터를 사용함으로써 시도할 수 있게 된 새로운 성격의 인문학 연구를 포함한다. 단순히 인문학의 연구 대상이 되는 자료를 디지털화 하거나, 연구 결과물을 디지털 형태로 간행하는 것보다는 정보 기술의 환경에서 보다 창조적인 인문학 활동을 전개하는 것, 그리고 그것을 디지털 매체를 통해 소통시킴으로써 보다 혁신적으로 인문지식의 재생산을 촉진하는 노력 등이 ‘디지털 인문학’이라는 새로운 조어의 함의라고 할 수 있다.²⁾

I -2. 디지털 인문학의 배경

1) 세계의 디지털 인문학

미국에서는 이탈리아의 예수회 신부 로베르토 부사(Roberto Busa, 1913~2011)가 토마스 아퀴나스(Thomas Aquinas)의 저작을 중심으로 하는

1) 미국의 인문학재단(National Endowment for Humanities)이 2008년 디지털 인문학 지원단(Office of Digital Humanities)을 설치하고 각 대학의 디지털 인문학 연구에 대한 지원을 강화하면서, 디지털 인문학은 미국뿐 아니라 세계 대학의 새로운 관심사로 부상하였다. 우리나라에서도 인문학 연구와 교육의 새로운 활로를 모색하고자 하는 인문계 연구자들이 ‘한국 디지털 인문학 협의회’를 결성하는 등 (2015년 5월) 디지털 사회에서 인문학의 기능을 제고하는 방법을 찾고 있다.

2) 김현, “디지털 인문학: 인문학과 문화콘텐츠의 상생 구도에 관한 구상”, 『인문콘텐츠』 29 (2013)

중세 라틴어 텍스트 1,100만 단어의 전문 색인을 전자적인 방법으로 편찬한 것(1949~1974)을 디지털 인문학의 시작으로 보고 있다.³⁾ 이를 계기로 인문학 연구의 새로운 방법에 눈을 뜨게 된 연구자들은 컴퓨터의 활용을 여러 방향으로 모색하기 시작했다.

이 방면의 선도적인 노력 가운데 하나는 인문학 지식을 담은 텍스트를 전자적으로 기술하기 위한 마크업 언어의 개발이다. 1987년 인문학재단의 지원으로 시작된 SGML(Standard Generalized Markup Language) 기반의 인문지식 전자 텍스트 기술 언어 TEI(Text Encoding Initiative)의 개발 연구는 1990년에 초안을 발표한 이후 2007년까지 지속적으로 확장되면서 인문학 텍스트를 전자적인 연구 자원으로 전환하는 일을 선도하였다.

초기에는 컴퓨터를 인문학 연구에 활용하려는 시도가 ‘인문학 전산화’(Humanities Computing), 또는 ‘전산 인문학’(Computational Humanities)이라는 이름으로, 텍스트 및 언어 자원의 색인·통계 처리를 위주로 하였으나, 정보기술 환경의 급속한 진화와 더불어 그 활용 범위를 데이터베이스와 멀티미디어, 그리고 대규모 원시 데이터에서 전자적인 방법으로 의미 있는 사실을 찾아내는 데이터 마이닝(Data Mining), 그 결과를 그래픽으로 보여주는 시각화(visualization)로까지 넓혀 갔다.⁴⁾



로베르토 부사(Roberto Busa, 1913~2011)⁵⁾

전통적인 인문학과 디지털 기술의 교섭이 더 이상 데이터의 양적 처리에만

3) Susan Schreibman, Ray Siemens, John Unsworth, *A Companion to Digital Humanities* (Oxford: Blackwell, 2004)

4) 김현, “디지털 인문학: 인문학과 문화콘텐츠의 상생 구도에 관한 구상”, 『인문콘텐츠』 29 (2013)

5) 사진 출처: <http://stephenramsay.us/2011/08/11/father-roberto-busa/>

머물지 않는다는 것을 주목한 일군의 연구자들은 ‘인문학 전산화’(Humanities Computing)를 대신하는 표현으로 ‘디지털 인문학’(Digital Humanities)이라는 용어를 사용하였다.⁶⁾ 이 새로운 조어는 2006년 미국 인문학재단의 기획연구사업의 이름(Digital Humanities Initiative)으로 채용되었고, 이것이 2008년 인문학재단의 정규 사업단 중 하나(Office of Digital Humanities)로 편제되면서 ‘디지털 인문학’은 디지털 환경에서의 인문학 연구와 교육을 지칭하는 이름으로 자리잡게 되었다.

2) 한국의 디지털 인문학

로베르토 부사의 토마스 아퀴나스 콘코던스(Thomas Aquinas Concordance)에 비견되는 우리나라 디지털 인문학의 효시는 에드워드 와그너(Edward Wagner, 1924~2001)의 문과 프로젝트라고 할 수 있다.

미국 하버드 대학의 동아시아언어문명학과 교수를 역임한 에드워드 와그너는 미국과 서유럽의 역사학계에서 한국사 연구를 선도한 학자로 잘 알려져 있다. 와그너 교수는 1967년에 하버드 대학교 옌칭 연구소의 지원을 받아 그의 한국인 동료 송준호 교수와 함께 문과 합격자 명부인 ‘문과방목’(文科榜目) 데이터를 디지털 데이터베이스로 편찬하는 프로젝트(Munkwa Project)에 착수하였다.⁷⁾ 와그너와 송준호는 이 프로젝트를 통해 14,600명의 문과 합격자와 그의 가까운 친족(부, 조, 외조, 장인)에 관한 데이터(본관 성씨, 관직, 거주지 등)를 컴퓨터에 입력하고 이를 종합적으로 분석하였다. 여러 가지 기술적인 어려움에도 불구하고 이와 같은 방식의 연구를 수행하고 있는 이유에 대해 와그너 교수는 다음과 같이 말하였다.

“역사 연구를 위해 컴퓨터를 사용하는 데에는 근본적인 이유가 있습니다. 컴퓨터는 이미 인류의 삶의 일부가 되었고 앞으로 더욱 그러할 것이기 때문입니다. 컴퓨터를 이용한 연구에 종사하는 것은 우리로 하여금 우리의 미래에 더욱 다가갈 수 있게 하는 것입니다.”⁸⁾

6) 2004년 Susan Schreibman, Ray Siemens, John Unsworth는 *A Companion to Digital Humanities* (Oxford: Blackwell, 2004)를 공동으로 편찬하였다. 이 책은 종래 ‘인문학 전산화’(Humanities Computing)라고 불리던 학술 활동의 확장된 개념으로 ‘디지털 인문학’(Digital Humanities)이라는 용어가 쓰이게 하는 데 기여하였다.

7) Kim Sun Joo “Edward W. Wagner and His Legacy: Toward New Horizons in the Research of Korean History”, The Wagner Fiftieth Appointment Anniversary Special Lecture, (Korea Institute, Harvard University, September 29, 2008)

8) Edward Wagner, "Problems in the Computerization of Materials in the Korean Studies Field: A Report on the Munkwa Project", 『한국학자료의 전산화연구』 (성남:한국정신문화연구원 1982)



에드워드 와그너(Edward Wagner, 1924~2001)

CAND NO=002486

#	1..	KSPM	002	0010	MKPM	003	026A	TKPM	005
#	2..	JUGJOG	17	1522					
#	3..	PYLSI							
#	4..	02B							
#	5..	SEGWN					MTPB	246A	
#	6..	JUH							
#	7..	0013	7720						
#	10..	2529	3266						
#	11..	52							
#	12..	51	NYN	59					
#	13..	JUH	2429	8047					
#	15..	JUH	7022	1327					
#	17..	JUH	1755						
#	19..	HWG	6210	0022					
#	22..	HA	3087						
#	34..	AN	3067	1446					
#	37..	HOCAM	HOTAG						
#	38..	WN	19						
#	41..	SAGJU							
#	42..	CILWN							
#	43..	1957	7872						
#	56..	34	0001	1172	1446				
#	57..	8427	8000	3637					
#	57..	0590	8000	3637					
#	57..	000000							
#	58..	2589	2398	6151					

와그너-송 문과 데이터. 한자 대용으로 중국의 전신 부호계 사용.

한국의 인문학자들로 하여금 와그너가 말한 ‘컴퓨터의 도움을 받는 연구’(Computer-assisted Research)에 다가가게 한 사건은 『조선왕조실록(朝鮮王朝實錄)』의 전문을 데이터베이스에 수록한 『조선왕조실록 CD-ROM』 편찬이었다. 1992년에 착수하여 1995년 10월에 첫 선을 보인 이 디지털 편찬물의 간행은 대한민국 인문학 정보화의 효시이자 가장 커다란 파급 효과를 가져다 준 획기적인 사업이었다.⁹⁾

9) 최희수, 「디지털 인문학의 현황과 과제」, 『소통과 인문학』 13 (2011)



국역 조선왕조실록 CD-ROM 초판 패키지

조선왕조실록 데이터베이스는 국내외의 한국사 연구자들, 특히 조선시대사 연구자들에게는 필수적인 연구 도구가 되었다. ‘검색’이라고 하는 제한된 기능만을 제공하는 시스템이었지만, 데이터의 양이 워낙 방대했기 때문에 기사 탐색에 많은 시간을 소비해 온 조선시대사 연구자들은 이 CD-ROM을 통해 인문학 연구 환경의 변화를 실감할 수 있었다.

조선왕조실록의 디지털화는 조선시대사 연구라는 학술 분야뿐 아니라 역사 지식을 대중적인 문화 상품의 창작에 활용하는 데 기여하였다. 어떤 의미에서는 교양서적 저술가와 TV 방송 프로그램 제작자, 연극 극작가 등이 디지털 조선왕조실록의 최대 수혜자였다고 할 수 있다. 예전 같았으면 조선시대사 전공자가 아니고서는 찾아낼 수 없었던 그 시대의 세세한 일들을 누구나 찾아볼 수 있게 된 것이다.¹⁰⁾

디지털 조선왕조실록은 인문지식이 일반인들도 즐길 수 있는 다양한 문화상품의 소재가 될 수 있다는 사실과 디지털 기술이 그것의 가교 역할을 할 수 있다는 사실을 일깨웠다. 이러한 인식은 우리나라 대학의 인문학 연구자들 사이에서도 확산되어 갔고, 그 연장선상에서 ‘인문콘텐츠학’과 ‘인문정보학’이라는 새로운 학술 활동이 시작되었다.

디지털 기술의 활용을 통해 ‘학술’과 ‘창작’ 사이에 놓였던 ‘지식 소통의 장벽’을 해소하고자 한 시도는 우리나라 ‘인문학 정보화’의 특징적인 면모의 하나이다. 이것은 1990년대 말부터 우리나라가 범국가적 차원에서 정보 산업과 문화콘텐츠 산업의 진흥을 위해 다각적인 정책을 펼친 것에도 그 원인이 있다

10) 김현, “The Korean Wave, Culture Content, and Cultural Informatics”, 『인문콘텐츠』 10 (2007)

고 할 수 있다.¹¹⁾ 이러한 배경에서 발전해 온 한국의 디지털 인문학은 전문적인 학술 연구만이 아니라 인문지식을 대중화하여 유용하게 쓰일 수 있게 하는 것, 그리고 그와 같은 일을 수행할 수 있는 응용인문학 인력을 양성하는 데에도 관심을 두고 있다.

I -3. 디지털 인문학 연구개발 사례

디지털 인문학이라고 할 수 있는 연구 활동의 구체적인 내용은 세계 여러 나라의 대학에 ‘디지털 인문학 연구 센터’(Center for Digital Humanities) 또는 이와 유사한 이름으로 설치된 조직에서 수행하는 연구개발 프로젝트와 그 산출물을 통해 살필 수 있다.

디지털 인문학 프로젝트는 교양적 수준의 인문지식 교재 개발에서부터 고도의 정보처리 기술을 동원한 빅 데이터 분석에 이르기까지 매우 다양한 성격을 드러낸다. 디지털 인문학은 인문학 분야의 새로운 분과 학문이 아니라, 인문학을 연구하고 교육하는 방법상의 혁신을 지향하는 것이기 때문에 연구의 범위나 주제를 어느 특정한 영역에 한정하기는 어려울 것이다. 다양한 성격의 연구개발 프로젝트 가운데 디지털 인문학에 대한 이해를 넓히는 데 도움이 될 만한 몇 가지 사례를 살펴보기로 한다.

1) 문화의 시각화

미국의 메사추세츠 공대(MIT)에서 수행하는 ‘문화의 시각화’(Visualizing Cultures) 프로젝트는 ‘이미지가 이끄는 학술’(Image Driven Scholarship)을 표방하는 디지털 환경의 인문 교육 교재 개발 사업이다.¹²⁾ 역사적 사실에 관한 그림, 사진 등의 이미지 자료를 디지털 영상으로 제작하고,¹³⁾ 영상 자료의 구석구석에 담긴 지식의 모티브를 찾아 학술적인 설명을 추가하는 방법으로

11) 김현, “문화콘텐츠, 정보기술 플랫폼, 그곳에서의 인문지식”, 『철학연구』 90 (2010)

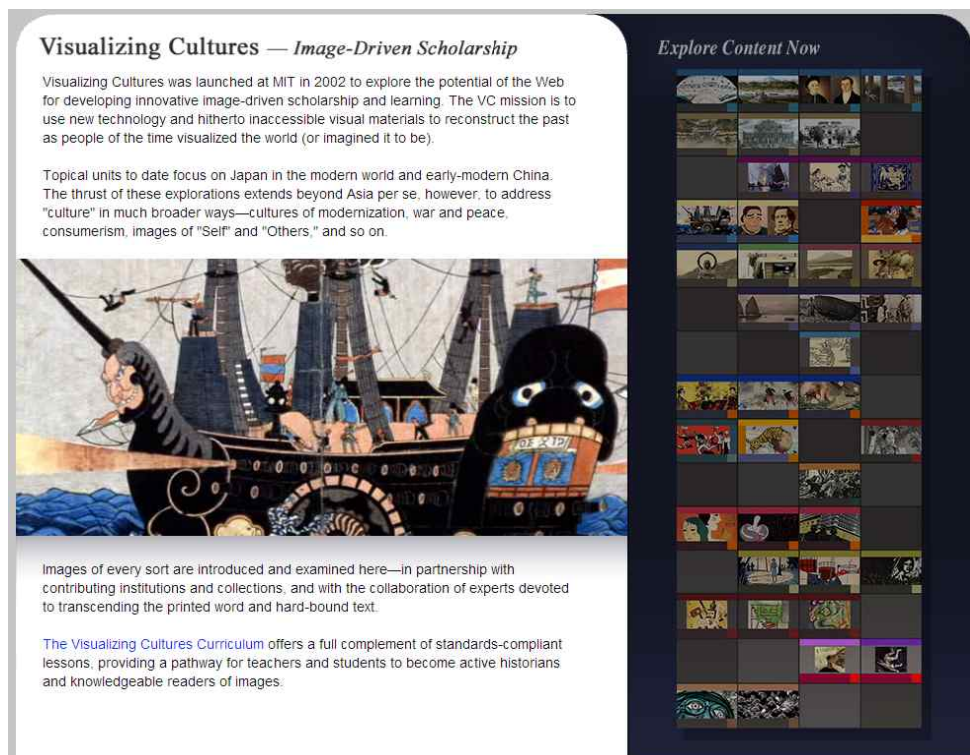
12) Visualizing Cultures: <http://ocw.mit.edu/ans7870/21f/21f.027/home/index.html>, Massachusetts Institute of Technology

13) Visualizing Cultures 프로젝트에서 활용하는 이미지는 대부분 미국과 일본의 유명 박물관에서 유물로 보존하고 있는 것들이다. 이러한 희귀 자료에 대한 접근성과 활용성을 높이는 것도 이 프로젝트가 지향하는 목표 가운데 하나이다.

이미지 자료 제공 기관: Arthur M. Sackler Gallery / Smithsonian Institution / Hiroshima Peace Memorial Museum / Honolulu Academy of Arts / Hood Museum of Art, Dartmouth College / Museum of Fine Arts, Boston / Peabody Essex Museum / Ryosenji Treasure Museum / Shiseido Corporation / Smith College Museum of Art

시각적인 스토리텔링을 구현하고 있다. 이 저작물은 모두 월드와이드웹(World Wide Web)을 통해 공개되고 있으며 MIT의 온라인 교육 프로그램으로 활용되고 있다. ‘중국 광둥무역의 흥망’(Rise & Fall of the Canton Trade System), ‘흑선과 사무라이’(Black Ship & Samurai) 등 20개 주제에 관한 45개의 코스웨어 유닛이 만들어져 있는데, 흥미로운 사실은 그것이 모두 아시아의 근대에 관한 내용이라는 것이다.

MIT의 Visualizing Cultures Project는 2010년부터 예일 대학(Yale University)의 동아시아학위원회(Council of East Asian Studies)와 함께 ‘Visualizing Asia in the Modern World’라는 이름의 컨퍼런스를 개최해 오고 있다.¹⁴⁾ 시각적인 자료를 통해 아시아의 역사와 문화를 연구한 것이나 아시아의 문화와 자연을 담은 멀티미디어 데이터베이스 등이 이 컨퍼런스를 통해 소개되고 있는데, 이러한 주제들은 앞으로 ‘문화의 시각화’(Visualizing Cultures)에 관한 또 다른 연구 프로젝트로 발전해 갈 것으로 예상된다.



Visualizing Cultures¹⁵⁾

14) Visualizing Asia in the Modern World: <http://www.visualizingasia.com>, 이 컨퍼런스는 2010년에 시작하여 해마다 열리고 있으며, Yale, Harvard, Princeton 세 대학이 각각의 연차 대회를 주관하고 있다.

15) <http://ocw.mit.edu/ans7870/21f/21f.027/home/index.html>

2) 편지 공화국의 지형도

‘편지 공화국의 지형도’(Mapping the Republic of Letters)¹⁶⁾는 미국의 스탠포드 대학에서 수행한 디지털 인문학 프로젝트이다. ‘편지 공화국’(Republic of Letters)이란 17, 18세기 유럽과 미국에서 원거리 편지 교신으로 지식과 감성의 공감대를 형성해 온 문화적 공동체를 지칭하는 표현이다. Mapping the Republic of Letters는 볼테르(Voltaire), 라이프니츠(Leibniz), 루소(Rousseau), 뉴턴(Newton), 디드로(Diderot) 등 계몽주의 시대의 인물들이 남긴 수많은 편지의 발신지와 수신지, 발신 날짜로 기록된 공간, 시간 정보를 시각적으로 재현한 다이내믹 디지털 콘텐츠이다.

Mapping the Republic of Letters¹⁷⁾

예시 화면에서 보이듯이 특정 시간대에 주고받은 편지의 수발신 위치가 세계지도상에 표시되고 있다. 이 네트워크 그래프의 한 노드를 클릭하면 그것에 해당하는 편지의 목록이 나오고, 다시 클릭하면 개별 편지의 내용을 확인할 수

16) Mapping the Republic of Letters: <http://republicofletters.stanford.edu/>, Stanford University

17) <http://www.stanford.edu/group/toolingup/rplviz/rplviz.swf>

있게 된다. 이러한 수준의 고기능 멀티미디어 데이터베이스의 구현은 시각화 기술만으로 이루어질 수 없으며, 그 내용을 이루는 인문학적 콘텐츠가 존재해야 가능한 일이다. 이와 같은 지리적 시각화 프로젝트가 가능했던 것은 방대한 규모의 유럽 계몽기 편지 데이터베이스가 인문학자들에 의해 먼저 구축되었기 때문이다.

‘Mapping the Republic of Letters’에서 사용한 모든 데이터는 영국의 옥스퍼드 대학에서 만든 ‘Electronic Enlightenment’¹⁸⁾ 데이터베이스에서 끌어온 것이다. 이 데이터베이스는 17세기 초에서 19세기 중반까지 약 200년의 기간 동안 7,476명의 사람들에 의해 쓰인 60,647건의 역사적인 기록물을 담고 있다. 단순히 원문을 디지털화 한 데 그치지 않고, 본문에 270,000여 건의 주석을 부가하고, 관련 있는 사람들을 연결시키고, 중요한 키워드는 옥스퍼드 인명사전 등 50여 개의 다른 데이터베이스에 하이퍼링크로 연결되도록 하였다.

3) 런던 사람들의 생활

디지털 인문학적 연구 성과로 주목받는 또 하나의 인문지식 데이터베이스는 영국 셰필드 대학과 허트포드셔 대학이 편찬한 ‘런던 사람들의 생활’(London Lives)¹⁹⁾이다. 이곳에는 1690년부터 1800년 사이에 영국 런던 거주민의 삶에 관계된 고문서 240,000여 건이 집적되어 있다.²⁰⁾ 교회 교구의 기록물을 비롯해 범죄와 재판에 관한 기록, 병원의 진료 기록과 검시 보고서, 상공인 조합의 기록, 빈민 구제에 관한 기록 등이다. 이 데이터 속에는 모두 3백 35만 개의 인명이 포함되어 있는데, 데이터베이스 편찬자는 그 가운데 동일 인물들을 추적하여 18세기 런던의 하층민으로 살았던 수많은 사람들의 생애를 재구성할 수 있게 하였다. 이 데이터베이스의 효용을 가늠하는 사례로, 그 속에서 찾은 한 고아 소년에 관한 기록을 살펴보기로 하자.

18) Electronic Enlightenment Project: <http://www.e-enlightenment.com/>, Bodleian Libraries, University of Oxford

19) London Lives, 1690~1800: <http://www.londonlives.org>, version 1.1, 24 April 2012

20) London Lives는 런던시의 8개 아카이브에 소장된 자료를 담은 15개 기존 데이터베이스를 이용하여 만들어졌다.

John Conway는 1775년 6월 7일 St. Clement Danes 교구회에 맡겨졌다. 교구 기록에는 그의 나이가 3년 6개월이고, 글을 읽고 주기도문을 읽 수 있다고 적혀 있었다. 간호사 Hill이 주급 2실링 6펜스를 받고 그를 보육하였다. 1778년 John은 가난한 소년들에게 일거리와 숙식을 제공하는 구빈원(救貧院)으로 옮겨진다. (이 때 그의 나이는 2살이 늘어난 8살로 기록된다.) 그는 이곳에서 걸레에 쓰는 실을 짜는 일을 했다. 교구의 도제 등록부에 의하면, John은 1783년 8월 30일 Essex 지방의 Barking에 사는 어부 Morris Jones의 도제(徒弟)로 보내진다. 이 때 만들어진 고용 계약서에는 7주 후에 마스터인 Morris Jones가 2파운드를 받고, 3년간의 고용 기간이 만료되면 2파운드 2실링을 더 받는 것으로 되어 있다. John은 3년 동안 일한 후에 옷 한 벌을 받기로 하였다. 1785년 4월 18일에 John은 다시 St. Clement Danes 구빈원의 명부에 올랐다. 하지만 4월 22일에 열린 입원 자격 심사에서 그가 도제 생활을 한 지 9달 만에 도망쳤던 사실이 드러나 입원이 거부되었고, 다음날 그는 Barking으로 추방되었다. 1786년 3월 15일, John은 14살이라고 나이를 속이고 다시 구빈원에 들어오려고 했지만, 이번에도 심사를 통과하지 못했다. 3월 17일, 존은 다시 Barking으로 돌려보내졌다.²¹⁾

London Lives, 1690~1800: www.londonlives.org

21) London Lives: <http://www.londonlives.org/static/ConwayJohn1775-1786.jsp>

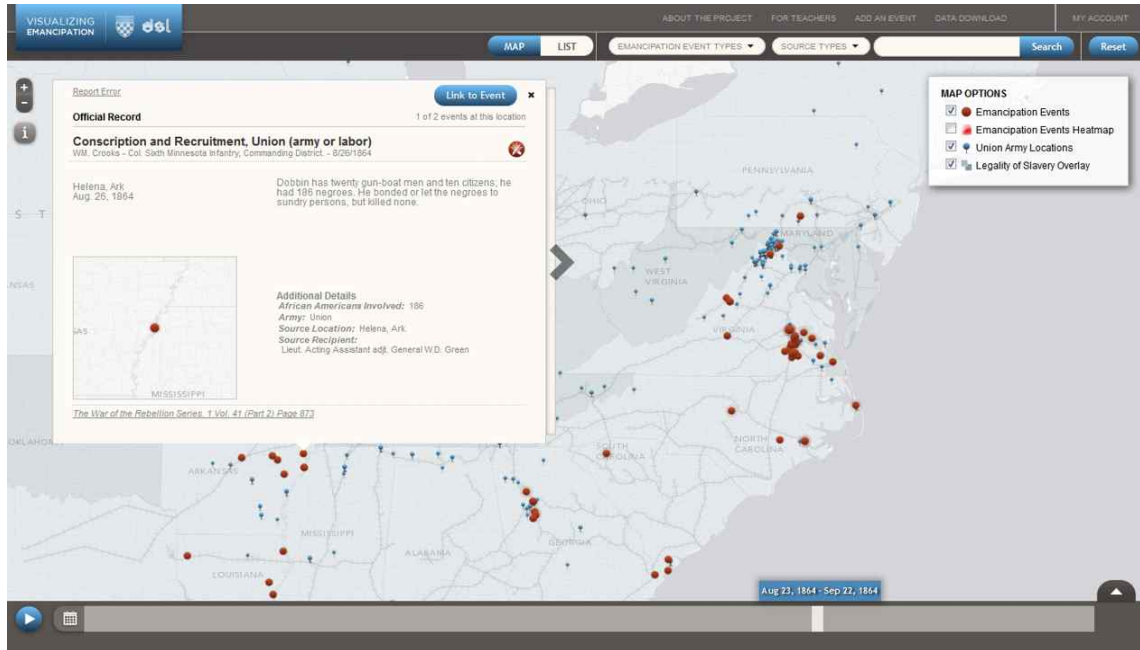
‘London Lives’ 연구는 ‘고문서’로 분류되는 옛 기록물의 데이터베이스가 단순히 문서의 보존을 위한 장치에 머물지 않고 새로운 지식을 발견하는 유용한 수단으로 활용될 수 있음을 잘 보여주고 있다. 그 유용한 활용은 디지털 기술의 도움만으로 가능한 것이 아니며, 미세한 데이터의 조각에서 의미 있는 사실의 실마리를 발견하는 인문학적인 식견과 통찰이 필요한 일이다.

4) 노예 해방의 시각화

‘노예 해방의 시각화’(Visualizing Emancipation)는 미국 남북전쟁 기간 동안 400만 명의 흑인 노예 해방의 시간적, 공간적 추이를 시각적으로 탐색할 수 있게 한 데이터베이스 편찬 프로젝트이다. 미국 리치몬드 대학 디지털 인문학 센터에서 미국 인문학재단(National Endowment for Humanities, NEH)의 디지털 인문학 프로그램의 지원을 받아 수행되었다.

미국 역사상 ‘가장 극적인 사회 변화’라고 일컬어지는 노예 해방은 어느 한 시점에 전국적으로 이루어진 사건이 아니다. 남북전쟁 기간 동안 정부의 정책, 북군과 남군의 지역별 전투 및 해당 지역의 점령 상황, 또한 각 지역의 노예제 농장과 그곳 흑인 노예들의 저항과 대립, 전투 참여 등에 따라 다양한 양상으로 전개되었는데, 이 연구는 노예 해방과 관련된 그 같은 다양한 사건들의 정보를 세밀하게 수집, 정리하여 전자적인 데이터베이스로 편찬하였고, 이를 통해 미국의 노예제 종식 과정에 어떠한 패턴이 있는지를 발견할 수 있게 하였다. 시각화된 데이터는 구글 어스(Google Earth)에서 열람할 수 있도록 KML(Keyhole Markup Language)²²⁾ 파일로 제공하고 있다. 이것은 연구자들은 위한 자료로서뿐 아니라, 학생들을 위한 교육 자료, 일반인들을 위한 교양 정보로도 활용되고 있다.

22) KML(Keyhole Markup Language): 구글 어스, 구글 지도에서 지리적 위치와 관련 정보를 표현할 수 있도록 하는 XML 기반의 전자지도 스크립트 언어



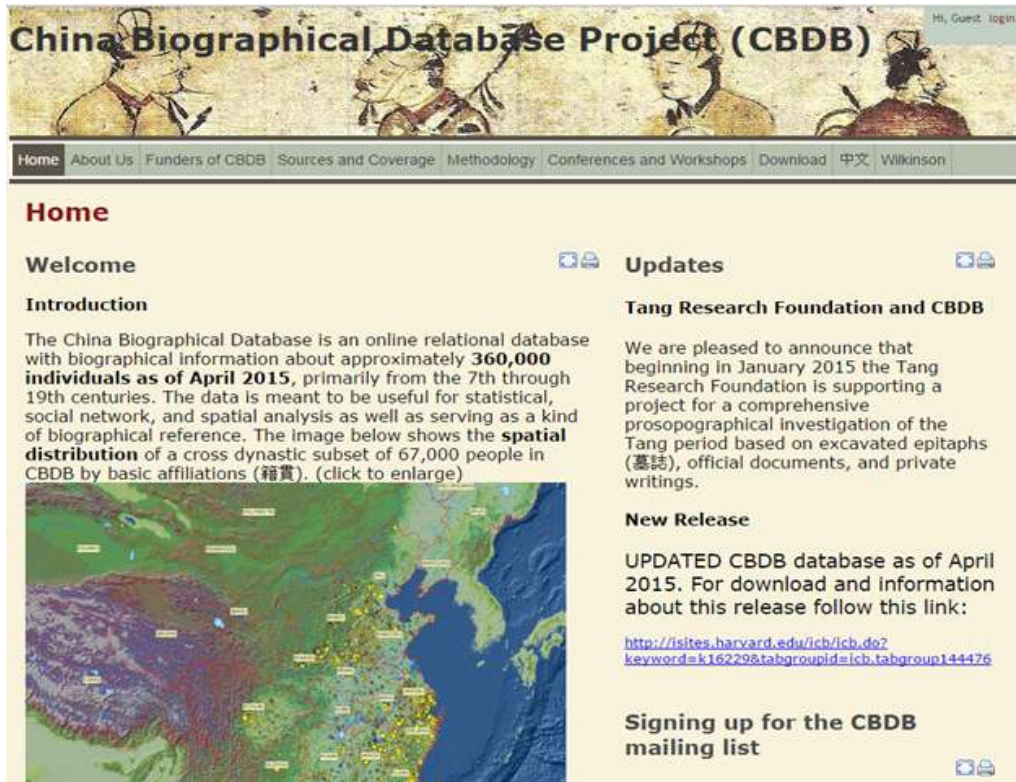
Visualizing Emancipation Project
(Map of Slavery's end during the American Civil War)²³⁾

5) 중국 역대 인물 데이터베이스

중국 역대 인물 데이터베이스(中國歷代人物傳記資料庫, China Biographical Database, CBDB)는 미국 하버드 페어뱅크 중국학 연구센터, 타이완 중앙연구원 역사언어연구소, 중국 북경대학교 중국고대사 연구센터가 공동으로 편찬하였다.

1995년, 펜실베이니아 대학 교수를 역임한 미국의 중국학자 Robert M. Hartwell이 작고하였을 때, 그가 남긴 유산에는 중국의 역대 인물 25,000명의 신상과 가계 정보 디지털 데이터가 포함되어 있었다. CBDB는 하트웰의 데이터를 토대 자료로 삼아서 편찬된 인물 정보 중심의 중국학 연구 데이터베이스이다.

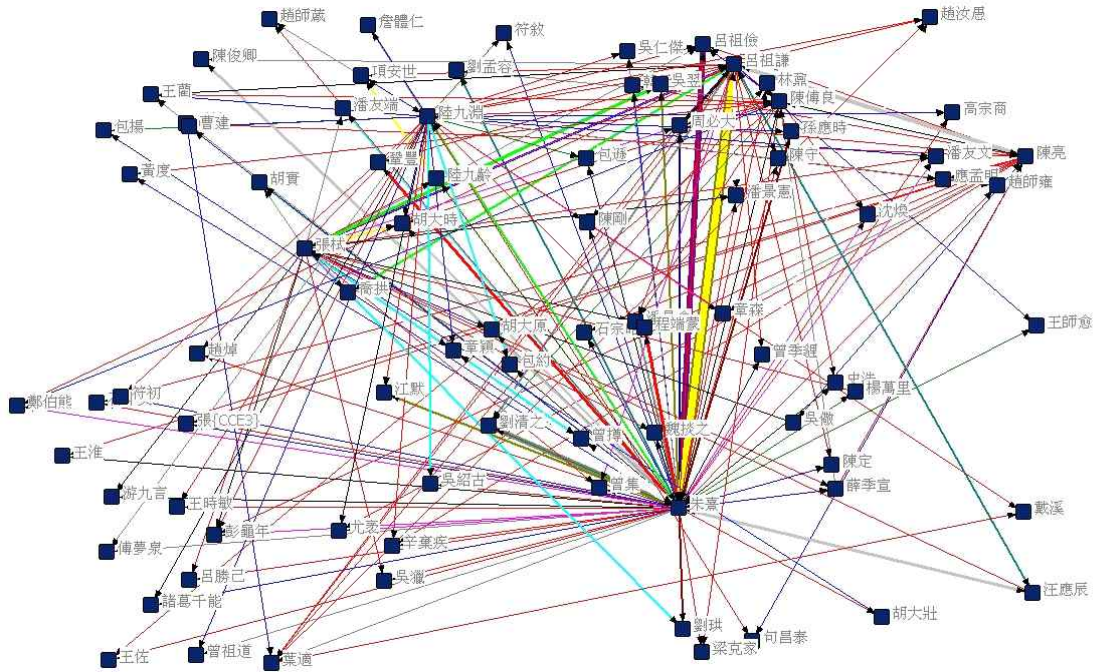
23) Visualization of Emancipation: <http://dsl.richmond.edu/emancipation/>



CBDB: China Biographical Database Project²⁴⁾

CBDB 데이터는 7~12세기에 중국에서 활동한 인물 360,000명에 관한 신상 정보(이름, 생/몰년, 별명, 임관, 출생/거주지, 저술 등)를 기본 정보로 수록하고 있으며, 여기에 더하여 그 인물들 사이의 친속관계와 사회관계를 기술하는 정보를 제공하고 있다. 특히 인물들 사이의 사회적 관계는 학문적 관계, 정치적 관계, 경제적 관계를 포함하는 500가지 이상의 관계성으로 상세하게 정보화함으로써 그 시대 중국사회 주요 인물들의 사회적 관계망을 시각화하고 다양한 관점에서 분석할 수 있는 길을 열어 주고 있다.

24) China Biographical Database Project: <http://isites.harvard.edu/icb/icb.do?keyword=k16229>



CBDB 데이터를 이용한 중국 인물의 사회적 관계망 시각화 예시

CBDB는 이 데이터를 중국학 연구에 관심을 갖는 연구자들이 응용할 수 있도록 다양한 접근 경로를 열어 두고 있다. 중국학에 관한 다른 연구 프로젝트나 일반 사용자들을 위한 정보 서비스 목적으로, 온라인상에서 CBDB 데이터를 끌어 쓸 수 있게 하는 API(Application Programming Interface)²⁵⁾를 제공할 뿐 아니라 이 데이터베이스의 전체 데이터를 다운로드 받을 수 있도록 허용하고 있다. 현재 미국, 유럽, 타이완의 여러 대학과 연구 기관에서 이 데이터를 이용한 2차적 연구 성과를 월드와이드웹 상에서 제공하고 있다.²⁶⁾

6) 조선시대 표류 기록의 시각화

조선시대 표류 노드 시각망은 조선시대에 발생한 표류 사건의 기록을 토대로, 표류인들의 해상·육상 이동 경로와 그들이 겪은 일에 대한 정보를 멀티미디어 데이터베이스로 편찬한 연구 결과물이다. 이 프로젝트는 우리나라 연구재단에서 2014년에 처음으로 시행한 디지털 인문학 연구 지원 사업의 일환으로 수행되었다.²⁷⁾

25) API(Application Programming Interface): 자신의 프로그램이나 데이터베이스 조작을 온라인상의 외부 시스템에서 할 수 있도록 제공하는 프로그래밍 인터페이스

26) CBDB API: <http://isites.harvard.edu/icb/icb.do?keyword=k16229&pageid=icb.page649231>

이 연구는 『조선왕조실록(朝鮮王朝實錄)』, 『변례집요(邊例集要)』, 『표인영래등록(漂人領來謄錄)』, 『제주계록(濟州啓錄)』, 『표해록(漂海錄)』 등 조선시대의 표류 사건에 관한 기사를 담고 있는 문헌을 검토하여, ‘표류 인물’, ‘출해 목적’, ‘표류 원인’, ‘표류 기간’, ‘표류 경로’, ‘표착 지점’, ‘송환 경로’, ‘송환 시기’ 등 의미 있는 정보 요소를 추출하고, 그 결과를 시각적 데이터 관계망 및 이동 경로가 표시된 전자지도로 표현한 것이다.

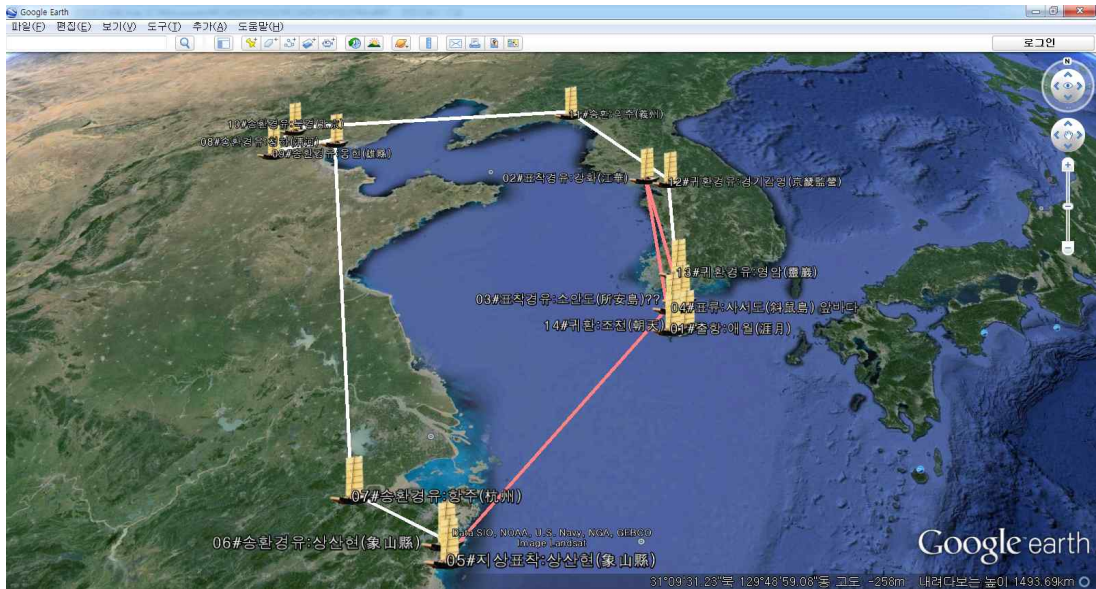
문헌 자료에서 추출한 표류 사건 기록은 위키(Wiki) 문서로 작성하여 웹상에서 열람할 수 있게 하였으며, 하나의 표류 안에서 발생한 작은 사건들, 표류 및 송환 과정에서 거쳐 간 지역들, 표류 선원들과 그들이 만난 인물들, 표류 선원을 취조한 조선 및 외국의 기관 등을 정보 개체로 설정하고 이것들 사이의 상관관계를 네트워크 그래프로 시각화 하였다. 아울러 표류 및 송환 과정에서의 이동 경로를 구글 어스 상에서 보여주는 전자지도, 제주도과 대마도의 해안 지역 중 표류 사건의 출해지 및 표착지로 추정되는 곳의 모습을 보여 주는 3D 영상 자료 등이 이 연구의 결과물 데이터베이스에 포함되어 있다.

이 연구는 종래 텍스트 자료의 조사 정리 형태로만 이루어지던 고문헌 기록 연구의 결과를 다양한 형태로 시각화할 수 있는 가능성과 함께 이를 통해 연구 결과에 대한 이해와 활용을 증진시킬 수 있음을 보이는 사례이다.



조선시대 표류 노드 시각망 (표류 사건 관계망)²⁸⁾

27) 이 연구를 위한 연구팀은 연세대학교, 선문대학교, 한국학중앙연구원의 연구진이 공동으로 구성하였다.



조선시대 표류 노드 시각망 (표류 및 송환 경로 전자지도)

28) 조선시대 표류 노드 시각망: http://www.digerati.kr/mediawiki/index.php/표류_노드_시각망

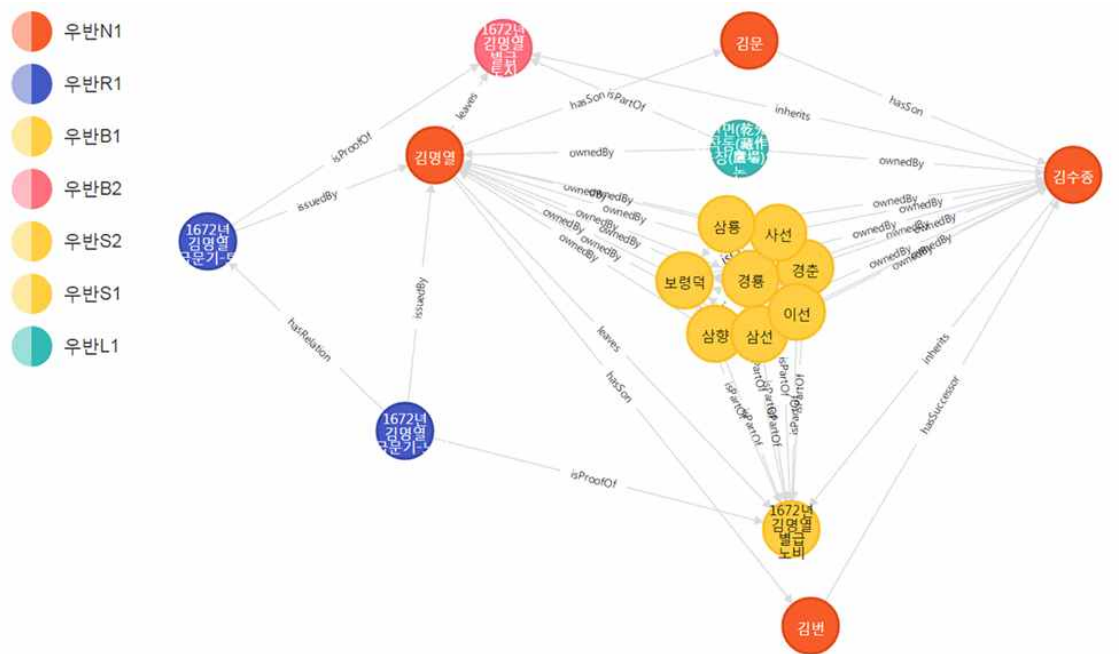
7) 문중 고문서 시맨틱 아카이브

문중 고문서 시맨틱 아카이브 구축 연구는 한국학중앙연구원 한국학대학원 인문학교실에서 수행하고 있는 조선시대 고문서 데이터 활용에 관한 연구이다. 이 기관의 고문헌 아카이브 장서각(藏書閣)에 소장된 조선시대 문중 고문서를 대상으로 하는 이 연구는 ‘고문서’라는 실물을 아카이빙하는 방법이 아니라, 그 문서 안에 기술된 내용을 데이터로 다루는 방법을 제시하기 위한 것이다. 디지털 환경에서 디지털 기술로, 고문서에 적힌 글의 내용을 분석하고, 새로운 의미를 찾아 그것을 재구성하는 방법을 찾고자 한 것이다.²⁹⁾

이 연구는 한국학중앙연구원 한국고문서자료관에서 디지털 사본을 서비스하는 65개 문중 고문서 중에서 재산 상속 내용을 담은 분재기류 문서, 재산 상속 사실을 증빙할 목적으로 작성된 소지, 입안류의 문서, 토지와 노비의 매매를 증빙하는 명문, 현금을 차용하고 발급한 차용증 등을 분석 대상으로 삼고 있다. 이와 함께 디지털적 연구 틀의 다른 한 축으로 삼는 것은 그 문중의 가계 계승 기록인 족보이다. 문중 고문서에 기록된 다양한 사건의 주체는 대부분 그 문중의 구성원들이다. 따라서 고문서에 나타난 인물들의 혈연관계를 보여주는 족보 데이터가 고문서의 내용 정보와 연계됨으로써 그들이 살아온 모습을 더 선명하게 드러낼 수 있게 된다.

이 프로젝트의 연구자들은 연구 대상으로 삼은 고문서와 족보에서 인물, 문서, 내용, 토지, 노비 등의 지식 요소들을 추출하여 범주화하고, 각 범주에 속하는 개체들이 서로에 대해 어떠한 관계를 맺고 있는지 조사하여 그 결과를 데이터베이스 속의 데이터로 기술하였다. 이러한 방법을 통해 독립적으로 존재 하던 개개의 문중 고문서들은 그 내용 요소들 사이의 관계성으로 인해 커다란 의미적 관계망으로 묶이게 된다.

29) 김현, 류인태, 「관계의 발견을 위한 디지털 스토리텔링 데이터 모델」, 『한국고문서 정서·역주 및 스토리텔링 연구 - 2015년 연구보고서』 (한국학중앙연구원 장서각, 2015); 김하영, 『門中古文書 디지털 아카이브 구현 연구』 (한국학대학원 석사학위 청구논문, 2015)



그래프 데이터베이스를 활용한 고문서 내용 분석 및 시각화³⁰⁾

30) <http://digerati.aks.ac.kr/paper/paper2015-3/2015-12-부안김씨.htm>