

문학 텍스트에서 플랫폼까지: 호이트 롱 교수와 함께 탐색하는 디지털 인문학의 지평 *

From Literary Texts to Platforms: Exploring the Horizons of Digital Humanities with Professor Hoyt Long

Hoyt Long **

 0000-0002-8562-5426

목차

1. 인터뷰 번역본
 - 1.1 디지털 인문학과 일본학
 - 1.2 디지털 미디어와 플랫폼 연구
 - 1.3 디지털 인문학과 교육
 - 1.4 디지털 인문학과 인공지능
 - 1.5 디지털 인문학 2026 국제학술대회
2. Interview in English
 - 2.1 Digital Humanities and Japanese Studies
 - 2.2 Digital Media and Platform Studies
 - 2.3 Digital Humanities and Education
 - 2.4 Digital Humanities and Artificial Intelligence
 - 2.5 Digital Humanities 2026 Conference

*이 인터뷰는 2024년 10월 21일, 시카고 대학교(University of Chicago) 동아시아 언어문명학과(Department of East Asian Languages and Civilizations)의 호이트 롱(Hoyt Long) 교수 연구실에서 진행되었으며 인터뷰 번역문과 원문을 순차로 게재하였다. 인터뷰어 김용수(Yongsoo Kim)는 대한민국 한림대학교 영어영문학과 및 디지털인문예술전공 교수이다. 『디지털인문학』(Korean Journal of Digital Humanities)의 편집위원장을 맡고 있으며, 미국 소설과 영화, 라캉 정신 분석 이론, 디지털 인문학을 중심으로 연구하고 있다. 저서로는 공동 저술한 *Digital Humanities and Scholarly Research Trends in the Asia-Pacific* (2019) 등이 있다. vadoropupille@gmail.com

The interview took place on October 21, 2024, in Professor Hoyt Long's office at the Department of East Asian Languages and Civilizations at the University of Chicago. The translated text and original text are presented sequentially, and the interviewer's information is as follows: Yongsoo Kim is professor of English and Digital Arts & Humanities at Hallym University in South Korea. He is editor-in-chief of the *Korean Journal of Digital Humanities*. His research primarily focuses on American novels and films, Lacanian psychoanalytic theory, and digital humanities. His publications include *Digital Humanities and Scholarly Research Trends in the Asia-Pacific* (co-authored, 2019). vadoropupille@gmail.com

**호이트 롱은 시카고 대학교 (동아시아언어문명학과) 소속으로 일본 문학과 디지털 연구를 가르치는 교수이다. 주요 연구 및 강의 분야는 문학, 미디어, 책의 역사, 플랫폼 연구, 문화 분석학(Cultural Analytics), 문화적 인공지능(Cultural AI) 등을 포함한다. 최근 저서로는 *The Values in Numbers: Reading Japanese Literature in a Global Information Age* (2021)이 있다. hoytlong@uchicago.edu

Hoyt Long is professor of Japanese Literature and Digital Studies at the University of Chicago (Department of East Asian Languages and Civilizations). His research and teaching interests include literature, media, book history, platform studies, cultural analytics, and cultural AI. He is the author, most recently, of *The Values in Numbers: Reading Japanese Literature in a Global Information Age* (2021). hoytlong@uchicago.edu

1. 인터뷰 번역본

1.1 디지털 인문학과 일본학

김용수:

디지털 인문학(Digital Humanities, DH)에 대한 관심은 어떻게 시작되었는가? 일본 문학에 대한 초기 연구가 어떻게 전산 문학 연구(Computational Literary Studies, CLS)로 발전하게 되었는가? 디지털적 접근은 일본학에 어떤 변화를 가져온다고 보는가?

호이트 롱:

대략 2010년 무렵이었다. 그 당시 나는 내 첫 번째 논문, 그러니까 미야자와 겐지(宮沢賢治)에 대한 전통적인 논문을 마무리하고 있었다. 그때 내 관심은, 생전에는 상대적으로 잘 알려지지 않았던 미야자와가 사후에 어떻게 인정받게 되었는가 하는 점이었다. 역사적 연구로 분명해진 것은, 그 인정을 가능하게 한 것이 다양한 지지자들의 네트워크였다는 점이다. 동료 시인들이나 지역 기반의 지지자들, 작품의 지지자들이 존재했고, 나는 그들이 어떤 방식으로 하나의 집단을 형성했는지를 포착하고 싶었다.

바로 그 무렵, 미국의 연방 자금 지원 기관인 미국국립인문기금(National Endowment for Humanities, NEH) 산하의 디지털 인문학 사무국(the Office for Digital Humanities)에서 다양한 주제로 여름 집중 과정을 지원하고 있었다. 그 해 주제는 ‘네트워크 분석’(network analysis)이었고, UCLA의 팀 탕겔리니(Tim Tangherlini) 교수가 주관했다. 그때 처음으로 ‘디지털 인문학’이라는 용어를 접했지만, 네트워크 분석이라는 방법이 내 질문에 실마리를 줄 수도 있다는 점은 바로 이해했다.

당시 나는 첫 교수직에 막 임용된 상태였고, 박사 과정에서 전산적 방법이라고는 전혀 사용하지 않았다. 사실 학부 때는 일본 문학(Japanese literature)과 컴퓨터과학(computer science)을 복수 전공했지만, 대학원에 들어오며 컴퓨터과학은 내려놓고 전통적인 문학 연구에만 몰두하고 있었다. 그런데 그 여름 프로그램은 전 세계에서, 짐작컨대 30~40명가량이 모여서 이 새로운 방법론을 함께 배우고 실험하는 자리였다. 그 자리에서 훗날 디지털 인문학 분야에서 활발하게 활동하는 인물이 될 사람들을 많이 만났다. 가령 하비에르 차(Javier Cha) 교수도 그곳에서 몰두하던 사람 중 한 명이었다.

미국국립인문기금(NEH)에는 디지털 인문학 사무국이라는 부서가 있다. 연방 기관이긴 하지만, 특정 주제를 정해 2주짜리 워크숍을 열 수 있도록 지원해 주는 프로그램이 있고 지금도 신청 가능하다. 대학이 이 프로그램에 지원하면 컴퓨터공학, 네트워크 분석, 인문학 분야의 전문가들을 불러 모아 강의를 진행한다. 나는 그 워크숍에서 일주일 동안 네트워크 분석, 토픽 모델링, 텍스트-데이터 마이닝 방법을 배웠다. 그 경험이 내게는 아주 결정적인 전환점이 되었다고 생각한다. 그때를 계기로 디지털 인문학 연구(DH)에 본격적으로 뛰어들게 되었다. 그런 기회는 흔치 않지만, 엄청난 영향을 줄 수 있다.

인문학뿐 아니라 서로 다른 학문 분야의 사람들을 한데 모은다는 점이 무척 중요하다. 그 워크숍 직후 나는 시카고로 옮겨, 리처드 소(Richard So)와 함께 20세기 초 시회(詩會)를 네트워크 분석으로 연구하기 시작했다. 사회적 역학을 분석하고자 했지만, 텍스트 자체에 대한 분석이 결여되

면 핵심을 놓치는 느낌이 들었다. 그래서 통계와 자연어 처리(NLP) 관련 수업을 청강하기 시작했고, 실제로 그런 분석을 어떻게 수행할 수 있는지도 배웠다. 컴퓨터공학, 통계학 수업을 들었고, 텍스트·데이터 마이닝을 이미 하고 있던 사람들에게 기초부터 가르쳐 달라고 직접 부탁했다. 그렇게 하면서 네트워크 분석에 텍스트 분석을 본격적으로 결합하게 됐고, 그때부터 연구가 본궤도에 올랐다.

김용수:

이런 컴퓨터 기반 접근은 일본학 분야에서 전반적으로 어떻게 수용되고 있는가?

호이트 룡:

그 당시만 해도, 일본학에서는 이런 방식의 연구를 아무도 하지 않았다. 미국이나 북미 지역에서는 역사학이나 문학 분야에 이런 작업을 시도하는 사람들이 점차 늘고는 있었지만, 여전히 소수 집단이었다. 몇 년에 걸쳐, 나는 이 분야에서 일하는 사람들을 시카고로 초청해 지금 이 분야가 어디쯤 와 있는지, 그리고 이런 방법들을 우리 학문 분야에 어떻게 적용할 수 있을지 함께 이야기하는 자리를 만들었다. 지금까지 이런 주제로 워크숍을 세 번 정도 열었다.

일본사 연구자인 마크 라비나(Mark Ravina)도 일본어 텍스트 마이닝 교육을 위한 지원금을 따내서, 둘이서 함께 워크숍을 두 번 진행했다. 일본학을 전공하면서 이런 도구를 배워보고 싶은 사람들에게 “일주일 동안 같이 공부하면서 일본어 텍스트 마이닝의 기초를 배워 보자”고 공개적으로 제안한 거다.

나는 2014년에서 2015년쯤부터 본격적으로 텍스트 분석 프로젝트를 시작했다. 처음엔 영어 텍스트 중심으로 작업했는데, 곧 다시 일본어 자료로 눈을 돌리게 됐고, 일본 내 연구자들과도 접촉하기 시작했다. 그 당시만 해도 문학 연구자 중에 이런 방식에 관심 있는 사람은 거의 없었다. 대신 국립일본어학연구소(the National Institute for Japanese Linguistics) 같은 곳의 언어학자들 중에는 텍스트·데이터 마이닝을 선도적으로 시도하는 사람들이 있었다.

이건 부분적으로는 세대 차이의 문제이기도 하다. 일본에서는 문학 연구자들이 몇몇 작가나 아주 좁은 주제에 평생을 바치는 경우가 많다. 예컨대, 나쓰메 소세키만 연구해서 그의 작품도 다 읽고, 생애도 속속들이 알고 있다면, 이미 코퍼스를 완전히 꿰고 있는 상황이기 때문에 디지털 방법을 써야 할 이유를 못 느낀다. 여러 이유가 있겠지만, 아마 이런 이유 때문에 일본에서는 이 분야가 더디게 발전해 온 것 같다.

일본 외의 지역은 다르다. 나도 처음엔 특정 작가 연구로 시작했지만, 학계 내에서 다른 사람들과 소통하려면 내 연구가 더 큰 주제와 어떻게 연결되는지를 보여 줄 수 있어야 했고, 디지털적 접근이 그걸 가능하게 해줬다.

10년 전쯤만 해도 반발이 꽤 있었다. 그때는 이걸 하는 사람이 정말 한 줄이었다. 흥미롭긴 하지만, 전통적인 연구자들에게는 그 잠재적 이득이 잘 와닿지 않았던 거다. 그런데 도구들이 점점 더 쉬워지고, 사람들이 익숙해지면서 분위기가 바뀌기 시작했다. 지금은 다들 구글 검색 엔진도 쓰고, 엔그램 뷰어(Ngram Viewer)도 쓰고, 디지털 데이터베이스도 당연히 쓰고 있다. 한번 그 정보 생태계 안에 들어오면 그다지 난관이 못 되는 거다. 게다가 이런 도구가 뭘 할 수 있는지 궁금해하는 젊은 연구자들도 점점 늘고 있다.

그리고 최근 2년 사이 대형 언어 모델(Large Language Model, LLM)이 나오면서 상황이 완전히 달라졌다. 이제는 “AI가 우리가 하는 모든 작업에 영향을 주는데, AI에 대해 뭔가는 말해야 하

지 않나” 하는 분위기가 자연스럽게 생겼다. 지금은 디지털 방법론 이야기를 안 하고는 지나가기 어려운 시대가 된 셈이다.

1.2 디지털 미디어와 플랫폼 연구

김용수:

자기소개 글에서 “기계 번역 및 플랫폼이 문화 생산과 전지구적 유통에 미치는 영향에 대한 연구를 포함하여, 현대 디지털 미디어 연구로 관심을 옮겨가고 있다”고 쓴 바 있다. 디지털 미디어 연구와 플랫폼 연구로 전환하게 된 특별한 계기가 있었는가? 그리고 이러한 연구들이 디지털 인문학(DH)과는 어떤 관계를 맺고 있는가?

호이트 룡:

초기에는 문학사의 문제에 주력했다. 20세기 초 문학 텍스트를 바탕으로 장르, 문체, 인종적 재현의 진화 양상을 디지털 방법으로 분석하는 작업이었다. 이때 가장 큰 난점의 하나는, 전산적 방법이 인간 독자처럼 문학의 미묘한 요소를 포착할 수 없다고 보는 시선이었던. 디지털 분석, 특히 ‘멀리서 읽기’(distant reading)는 ‘가까이 읽기’(close reading)에 비해 너무 이질적으로 느껴졌기 때문에, 전통적인 문학 연구의 대체물이 될 수 없다고 보는 분위기가 지배적이었던. 많은 학자들이 특정 분야를 평생 연구해 왔다. 그런 사람들에게는 디지털 방법이 무엇을 더해 줄 수 있는지를 알아보는 것이 어렵고, 오히려 전산적 접근이 무엇을 놓치고 있는지를 쉽게 알아차릴 수 있다. 그만큼 자료를 너무 잘 알고 있기 때문이다.

하지만 동시대에 디지털로 태어난 미디어나 인터넷, 플랫폼 문화 연구로 넘어오면 디지털 도구를 사용하는 데 대한 저항이 훨씬 적다. 왜냐하면 이진 정전(canonical) 자료가 아니라 대중적인 자료이기 때문이다. 제임스 조이스나 셰익스피어처럼 “미학적” 가치가 크다고 간주되는 텍스트가 아니다. 그래서 대중문화를 다루기 시작하면 디지털 방법론을 쓰는 데 훨씬 덜 저항감을 느끼게 되고, 디지털 미디어를 디지털 방법 없이 연구하는 건 점점 불가능해지고 있다.

The Values in Numbers (2021)를 출간하면서, 문학사를 분석하는 데 이런 도구들을 어떻게 쓸 수 있는지에 대해 내가 말하고 싶었던 것을 어느 정도 다 말했다고 느꼈다. 그래서 이제는 새로운 방향, 예를 들어 플랫폼 스토리 같은 것을 탐색하고 있다. 지금 하고 있는 프로젝트 중 하나는 넷플릭스를 플랫폼으로 보고, ‘오징어 게임’을 사례로 삼아 넷플릭스가 텔레비전 프로그램의 유통과 전 세계 수용 방식을 어떻게 바꿨는지를 이해하려는 것이다. 우리는 영어, 일본어, 한국어 등 여러 언어권의 소셜미디어 반응 데이터를 대량으로 수집해서, 사람들이 이 드라마에 어떻게 반응했는지, 그리고 그 반응이 언어적·문화적 공동체에 따라 어떻게 달랐는지를 살펴보고 있다.

우리가 알아낸 것 중 하나는, 일본 시청자들은 오징어 게임을 익숙한 장르의 연장으로 받아들였다는 점이다. 고전 애니메이션과 만화에서 봐 왔던 ‘데스 게임’(death game) 장르 말이다. 그 익숙함 때문에 일본에서 오징어 게임은 다른 지역만큼 인기를 끌지 못했다. 그래서 탐색하기 시작한 질문 중 하나가, 이 드라마의 최초 수용 지점으로 일본 대중문화가 얼마나 작용하고 있는가 하는 것이었다. 일본 대중문화가 전 세계적으로 인기가 있다는 것은 이미 알고 있다. 그렇다면 그 인기 덕분에 오징어 게임도 크게 히트할 수 있었던 건 아닐까? 우리가 알아낸 바에 따르면, 그 성공은 아마도 전 세계적으로 이미 퍼져 있던 일본 대중문화 덕분일 가능성이 크다. 영어권 소셜

미디어에서도 일부 사용자들이 오징어 게임을 일본 대중문화의 관점에서 해석하고 있었고, 이것이 관심을 끌어올리는 데 도움이 되었다. 일본 대중문화는 오징어 게임이 세계적으로 더 쉽게 받아들여질 수 있도록 하는 참조점(reference point) 역할을 했다.

이 데이터 전체를 대상으로 토픽 모델링(topic modeling)을 수행해서 어떤 공통 화제가 떠오르는지 살펴볼 수 있다. 다양한 플랫폼의 간 대화를 비교했더니, 언어와 플랫폼에 따라 떠오르는 주제가 달랐다. 예를 들어, 자본주의(capitalism)는 영어권과 한국어권 플랫폼 논의에서 핵심 주제였지만, 일본에서는 그렇지 않았고, 대신 장르적 요소에 더 집중했다.

또 우리는 제3의 데이터 분석 업체와 협력해서 오징어 게임이 다른 나라에서 얼마나 인기가 있었는지를 파악했다. 20개국에 걸친 수용 데이터(reception data)를 5년 동안 살펴봤는데, 우리가 알아낸 것은 오징어 게임이 예외적인 작품이라는 것이다. 다른 K-드라마들은 대중적 반응 폭이 매우 좁다. 아시아에서는 큰 주목을 받지만, 오징어 게임처럼 전 세계적으로 반응이 나온 경우는 없었다. 그래서 다음 질문이 자연스럽게 떠올랐다. 과연 어떤 다른 드라마가 전 세계적인 영향력을 갖는가? 그리고 오징어 게임은 왜 그 범주에 들어가는가? 이 드라마가 그런 성공을 거둔 사례라는 점에서 우리는 무엇을 배울 수 있는가? 20개국의 관객에게 동시에 어필할 수 있는 드라마는 어떤 특성을 갖고 있는가?

1.3 디지털 인문학과 교육

김용수:

디지털 인문학(DH)은 대학 교육 현장에, 특히 인문학 교육 현장에 어떤 변화를 가져왔다고 보는가? 시카고 대학교에서 디지털 인문학을 가르칠 때 겪는 구체적인 어려움은 무엇이며, 그것을 어떻게 해결하고 있는가?

호이트 롱:

시카고 대학교에서는 디지털 연구 석사 과정(MA program in digital studies)을 2년제로 운영하고 있다. 이 과정으로 디지털 인문학 강의를 담당할 교수진과 실무 인력을 확보할 수 있었고, 덕분에 박사과정 학생뿐 아니라 학부생에게도 다양한 강의를 제공할 수 있게 되었다. 한 명의 교수가 독립적으로 이끌어 가는 것이 아니라, 여러 분야에 걸친 다양한 교수진이 참여하고 있다는 점에서 이 프로그램은 특별한 의미를 갖는다. 물론 학부 수준에서는 여전히 그 향방이 불분명한 감이 있다. 내 생각에 학부생들은 디지털 인문학이라는 분야 자체를 잘 모른다. 잘 모르다가 그 존재를 우연히 접하거나 특별한 계기가 있어야 관심을 갖게 되는 경우가 많다. 그래서 내가 디지털 인문학과 관련해 주로 가르치는 수업은 대부분 석사과정 학생들을 대상으로 하고 있다.

프로그램 안에는 파이썬 기초 수업이 있고, 데이터 분석 기초, 기초 통계, 데이터베이스 등 여러 기초 과정이 포함되어 있어, 모든 대학원생이 수강할 수 있다. 또 디지털 인문학 학위 과정도 마련되어 있다. 이런 류의 방법론을 배우고 싶은 학생들은 언제든 와서 프로그램을 통해 관련 방법론을 배울 수 있다.

한편, 학부생들은 사정이 조금 복잡하다. 대부분 프로그래밍이나 통계에 대한 배경지식이 부족하고, 그만큼 디지털 인문학에도 익숙하지 않다. 학부생들에게 DH의 가능성을 효과적으로 소개할 방법을 아직 찾지 못한 상태다. 컴퓨터과학을 전공하면서 동시에 인문학 수업을 좋아하는 학생들

은 많고, 꽤 자주 보기도 하지만, 그 학생들이 이 두 가지를 어떻게 함께 할 수 있을지 알아낸 것은 아니다.

그런 점에서 미디어 예술 및 디자인(Media Arts and Design) 프로그램은 바람직한 방향의 시도라고 생각한다. 이 프로그램은 비디오 게임, 일러스트레이션, 시각 디자인 등도 포함하며, 학생들이 다양한 소프트웨어나 플랫폼 기술을 실제 창작 활동과 접목시켜 활용하기 쉽게 구성되어 있다. 창의적 재료를 다루면서 동시에 프로그래밍에도 깊이 관여하는 학생들이 이 프로그램에 참여하고 있다. 인간-컴퓨터 상호작용(Human-Computer Interaction, HCI) 분야나 창의적 작업을 지원하기 위한 디지털 도구의 인문학적 재설계 등에도 많은 가능성이 열려 있다. 이처럼 디지털 인문학이 큰 영향을 미칠 수 있는 흥미로운 지점이 존재한다. 시카고 대학교는 디지털 미디어와 디자인 연구를 모두 아우를 수 있는 좋은 환경을 갖추고 있고, 이를 ‘미디어 예술 및 디자인(Media Arts and Design)’이라고 부른다. 이 전공은 현재 인문학 분야에서 가장 빠르게 성장하고 있다.

1.4 디지털 인문학과 인공지능

김용수:

최근 디지털 인문학(DH) 연구에서 가장 주목하거나 흥미롭게 보고 있는 새로운 방향이나 흐름은 무엇인가? 아직 충분히 연구되지 않았지만 가능성이 크다고 생각하는 분야가 있다면?

호이트 롱:

최근 가장 흥미로운 방향의 하나는 대형 언어 모델(Large Language Models, LLM)이다. 지난 1년 동안 이 모델들을 활용한 여러 프로젝트에 참여하게 되었고, 이러한 기술이 인문학 연구에 어떤 함의를 지니는지를 고민해 왔다. 해 볼 만한 도전 과제 중 하나는, 이 모델들이 문화사에 대해 무엇을 알고 있는지, 그리고 그것이 우리가 평소에 접근하기 어려운 영역을 어떻게 드러낼 수 있는지 이해하는 일이다.

나는 인문학자들이 이 모델들의 문화적 지식을 평가하고 테스트하는 데 중요한 역할을 할 수 있다고 본다. 컴퓨터과학자들은 이런 도구들을 평가하는 데 매우 능숙하다. 그들의 평가 방법론을 빌리되, 인문학적 질문에 맞게 적용해야 한다고 생각한다. 기존과는 다른 형태의 연구지만, 우리가 기여할 수 있는 분야다. 우리는 대규모 언어 모델(LLM)이 시(poetry), 어조(voices), 문체(styles), 그리고 다양한 문화적 맥락을 얼마나 잘 이해하고 수용할 수 있는지를 평가할 필요가 있다.

이 모델들이 무엇을 할 수 있는지 더 잘 이해하게 되면, 특정 문학 데이터(literary data)를 학습시켜 인문학 연구에 더 유용하게 쓸 수 있을 것이다. 예를 들어, 20세기 소설을 모델에 학습시킨 다음, 등장인물, 행위, 관계를 자동으로 식별하게 한다면, ‘멀리서 읽기’를 수행하는 데 매우 유용한 도구가 될 수 있다.

가령 어떤 인물의 성별이나 행동을 아주 잘 식별해 내는 모델이 있다고 해 보자. 이제 그 모델을 10,000권의 소설에 적용할 수 있게 되는 것이다. 마음에 둔 구체적인 질문 때문에 갑자기 ‘멀리서 읽기’로 돌아오게 되더라도, 이번에는 더 복잡한 질문에도 답할 수 있는 모델을 갖고 있는 셈이다. 즉, 더 이상 언어학적 구조나 전통적인 자연어 처리(NLP) 기법에만 의존하는 것이 아니라, 언어 모델이 알고 있는 것과 인식할 수 있는 능력을 활용해 ‘멀리서 읽기’를 수행하는 것이

다.

이미 일부 연구자들은 이러한 가능성을 적극적으로 탐색하고 있다. 전통적인 머신러닝 기반 분류(classification) 대신, GPT 모델을 활용한 분류를 시험해 보는 학자들도 봤다. 예컨대 하이쿠 텍스트 100편과 하이쿠가 아닌 텍스트 100편이 있을 때 LLM에게 둘의 차이를 알려 달라고 물어볼 수도 있다.

그리고 어떤 경우에는 전통적인 머신러닝 기법으로 우리가 해 왔던 작업을 꽤 잘 재현해 내기도 한다. 그게 가능해지면 더 복잡한 질문을 던질 수도 있으니, 그것이 어디에, 어떻게 유익하게 작용할 수 있는지도 아마 알 수 있을 것이다. 머신러닝을 다룰 줄 모르는 인문학자라도 언어 모델(language model)과 대화하는 방법은 알고 있으므로, 오히려 더 쉽게 다가갈 수 있는 가능성도 열리게 된다. 우리가 좋은 모델을 갖추게 된다면, 그 모델은 실제 인문학 연구에 적용할 수 있는 무언가가 되는 것이다.

아직 그 단계에 이르지지는 않았지만, 향후 5년 안에는 가능해질 수 있다고 본다. 테드 언더우드(Ted Underwood), 데이비드 배먼(David Bamman), 매튜 윌킨스(Matthew Wilkens), 멜라니 월시(Melanie Walsh) 같은 연구자들은 이미 대형 언어 모델(LLM)을 인문학 연구에 활용하기 위한 탐구를 진행하고 있다. 이 분야는 우리 인문학자들이 실질적 영향력을 발휘할 수 있는 영역이다.

나는 이러한 도구들을 활용해 가까이 읽기(close reading)와 디지털 방법론을 결합한 집중적 사례 연구(focused case study)가 가능하다고 생각한다. 모델의 성능을 평가할 때는, 특정 텍스트 단락에 대해 모델이 무엇을 보고 있는지, 인간 독자와 어떻게 다른 방식으로 해석하는지를 함께 비교할 필요가 있다. 항상 검증 단계가 있어서, 보다 폭넓은 분석을 목표로 하더라도 검증 단계에는 세심한 ‘가까이 읽기’가 깊이 관여하게 된다.

김용수:

인공지능(AI), 특히 생성형 AI(generative AI)이 디지털 인문학이나 더 넓은 인문학 교육에 어떤 영향을 미치고 있다고 보는가? 교육 현장에 이런 기술들이 부상하면서 동료들은 어떤 반응을 보이고 있고, 교육에 이런 도구들을 효과적이고 윤리적으로 통합하기 위한 방안은 무엇이라고 생각하는가?

호이트 룡:

교육의 측면에서 생성형 AI는 아주 큰 영향을 주고 있다. 우리는 이 도구들을 어떻게 책임 있게 사용할 것인지 고민할 필요가 있다. 만약 글쓰기 수업을 맡고 있다면, 학생들이 생성형 AI를 그대로 쓰도록 허용하는 건 말도 안 되는 일이다. 하지만 창의적인 방식으로 교실 안에 이 도구들을 들여올 수는 있다. 학생들이 이미 사용하고 있는 마당에 눈을 감고 무시할 수는 없다. 대신 이 도구들이 무엇을 할 수 있는지, 그에 대한 문해력(literacy)을 기를 수 있게끔 도와야 한다.

내 수업에서는 창의적인 접근을 시도해 왔다. 예를 들어 신입생 글쓰기 수업에서는 ChatGPT로 에세이를 쓰는 것을 금지하지만, 그것이 학생들이 더 나은 독자가 되는 데 도움을 줄 수 있을지에 대해서는 고민하고 있다. 대부분의 신입생들은 우리가 읽는 텍스트에 대해 깊은 배경지식을 가지고 있지 않다. ChatGPT는 역사적·문화적 인식을 확장하는 하나의 매개체, 혹은 최소한 첫걸음이 될 수 있을지도 모른다. 우리는 종종 학생들에게 어떤 구절을 읽고 자신의 경험과 배경을 바탕으로 성찰해 보기를 기대한다. 하지만 만약 ChatGPT에게 어떤 역사적 인물의 역할을 맡기거나, 아주 다른 관점을 지닌 인물의 입장에서 같은 구절을 읽게 한다면 어떻게 될 것인가? 이는

교실에서 우리가 다루는 목소리의 범위와 읽기 및 해석의 유형을 확장하는 데 도움이 될 수 있을 것이다.

1.5 디지털 인문학 2026 국제학술대회

김용수:

마지막 질문이다. ‘디지털 인문학 2026 국제학술대회’(The Digital Humanities 2026 Conference)가 한국에서 개최될 예정이다. 이 학술대회에 대한 기대가 있다면 무엇인가? 동아시아 연구뿐만 아니라 전 세계 디지털 인문학 분야에 어떤 영향을 미칠 것으로 전망하는가?

호이트 룡:

‘디지털 인문학 2026 국제학술대회’이 동아시아의 주제를 연구하는 동아시아 및 북미 지역의 연구자들이 한자리에 모이는 것이다. 동시에, 동아시아 전문가가 아닌 다른 지역의 연구자들에게도 동아시아 연구에서 이루어지고 있는 연구의 수준과 가능성을 보여 주는 자리가 되기를 바란다. 그리고 이 행사가 한국 내 디지털 인문학의 성장을 촉진하는 계기가 된다면 정말 좋을 것이다. 그런 일들이 실제로 가능할지는 결국 얼마나 다양한 분야의 연구자들이 참여하는지에 달려 있다. 만약 그들에게 디지털 인문학이라는 분야의 잠재력을 실감할 수 있는 기회를 만들어 줄 수 있다면, 아주 환상적일 것이다.

2. Interview in English

2.1 Digital Humanities and Japanese Studies

Yongsoo Kim:

What first sparked your interest in digital humanities (DH), and how did your early work in Japanese literature evolve into computational literary studies? What difference do you think the digital approach to Japanese studies makes?

Hoyt Long:

It was roughly 2010. At that time, I was finishing up my first monograph, which was a traditional monograph on the writer Miyazawa Kenji. One thing I was interested in at that time was understanding how he, as a poet and a writer, was brought to recognition after his death because he was relatively unknown during his lifetime. It was clear from historical research that it required a network of supporters—other poets, local proponents, and supporters of his work. I was interested in how we could capture the various groups that gathered around him.

At that same time, the National Endowment for Humanities (the Office for Digital Humanities), which is a government funding agency, was sponsoring summer institutes on various topics. The topic that year was network analysis, organized at UCLA by Tim Tangherlini. That was the first time I'd really heard about digital humanities, but I grasped that this method—network analysis—could potentially help my research questions.

I was already teaching my first job at that time. I wasn't doing anything computational throughout my PhD. I double majored in Japanese literature and computer science, but I had set all the computer science stuff aside and was just focused on traditional literary studies. But this institute brought together, I guess, probably 30 to 40 people who wanted to learn this new thing of network analysis. Many of the people who were there eventually became very active figures in the digital humanities world. That's where I met Javier Cha, for instance, who was also there trying to figure things out.

The National Endowment for Humanities has an office called the Office for Digital Humanities. It's a federal agency, but they had this funding program, and it's still available. Schools can apply to teach a two-week workshop on a specific topic. So, they brought in experts in computer science, network analysis, and humanists. Over the course of a week, we were meant to learn the method and hear from people at the forefront of their fields. I learned about network analysis, topic modeling, and text and data mining during that particular week. I think it's been a transformative experience. Since then, I've been involved in DH work. Opportunities like that are rare, but they can have a tremendous impact.

Bringing together people from different disciplines—not just the humanities—is really valuable. Right after that workshop, I came to Chicago and began doing projects with Richard So, applying network analysis to poetry circles in the early 20th century. We were interested in understanding the social dynamics, but without the underlying text, we felt like we were missing something. So, we began auditing courses in statistics and NLP and learned how to do that kind of work. We took courses in computer science, statistics, and from people already doing text and data mining. We asked them to tutor us on some basics, and that's where we shifted to doing more text analysis combined with network analysis. Since that point, it's really taken off.

YK:

How is this computational approach accepted in Japanese studies in general?

HL:

At the time, nobody was doing this kind of work in Japanese studies. In the U.S. and North American context, there are increasingly more people doing that work in history and literature, but it's still a small group. Over the years, I've brought the community of people working on these topics to Chicago to talk about the state of the field and how we can apply these methods to our discipline. We've organized about three workshops on the topic.

Another scholar in Japanese history, Mark Ravina, also got funding to teach text mining for Japanese workshops. We actually did it twice, putting out an open call to the field, saying if you're in Japanese studies and want to learn how to use these tools, come spend a week with us, and we'll teach you the basics of text mining for the Japanese language.

I started doing text analysis projects in 2014 or 2015. Initially, these projects revolved around English-language sources, but then I began turning my attention back to Japanese language sources and started reaching out to scholars in Japan. At that time, I found very few people in literature who were interested, although I did find some in linguistics, particularly at the National Institute for Japanese Linguistics, who were pioneering the use of text and data mining.

It's partly a generational thing. In Japan, literary studies scholars tend to focus on a few authors or a very narrow topic. If you only study Natsume Soseki and you've read everything by him and know everything about his life, you might not see the value in digital methods because you already know your corpus. That's probably one reason why this kind of research has been slow to develop in Japan.

Outside of Japan, it's different. My initial research was on a particular writer, but in communicating with others in academia, you need to show how your work speaks to larger themes, and that's where digital approaches are helpful.

There was a lot of pushback initially, about ten years ago. Back then, there were only a handful of people doing this work. It seemed exciting, but traditional scholars had trouble understanding the potential benefits. That began to change as the tools became easier to use, and people became more familiar with them. Everybody now uses the Google search engine and the Ngram Viewer. We're all using digital databases and so it's less of a hurdle once you're already in that information ecosystem. You also had younger scholars who were interested in what these tools could do.

The use of large language models over the last two years has really changed things. Now, there's a sense that we need to have some kind of response to AI because it affects everything we do. Now it's hard to avoid the topic of digital methods.

2.2 Digital Media and Platform Studies

YK:

In your bio, you said you “have shifted attention to the study of contemporary digital media, pursuing projects on machine translation and the impact of platforms on both the production and global circulation of culture.” Do you have any particular reason to turn to digital media studies and platform studies? How are they related to DH?

HL:

Initially, my work was focused on literary history—using texts from the early 20th century and applying digital methods to answer questions about the evolution of genres, styles, and racial representation in literature. I think part of the challenge was that nobody believed that these computational methods could capture the subtleties of literature in the way that human readers do. Distant reading felt so alien to close reading, and because it was so alien, there was a sense you couldn't substitute it for the methods of traditional literary studies. A lot of scholars have spent their careers studying a particular field. It is harder for them to recognize the added value of digital methods, or else they can easily recognize what the computational approach misses because they know the

material so well.

But as we get into contemporary born digital media and studies of the internet or platform culture, there's far less resistance to using digital tools in part because this is not canonical material but popular material. It doesn't have as much "aesthetic" value as James Joyce or Shakespeare. So once we move to the popular culture, I think there's far less resistance to using digital methods, and it becomes increasingly impossible to study digital media without digital methods.

I published *The Values in Numbers* (2021) and felt like I had said what I wanted to say about the use of these tools for analyzing literary history. Now, I'm exploring new directions—platform studies, for instance. One project I'm working on is about Netflix as a platform, using *Squid Game* as a case study to understand how Netflix changed the circulation and global reception of television shows. We're looking at large collections of data of social media responses across different languages—English, Japanese, and Korean—to see how people reacted to the show and how those responses differ across linguistic and cultural communities.

We've found that Japanese viewers recognized *Squid Game* as an extension of a familiar genre—the death game genre they've seen in classic anime and manga. That familiarity meant the show wasn't as popular in Japan as in other places where it seemed more novel. One of the things we began to explore then is how much Japanese popular culture is acting as the first point of reception for this show since we know that Japanese popular culture has global appeal. So is there something about that appeal that allowed *Squid Game* to become a huge hit? What we find is that its success likely depended on the prior spread of Japanese pop culture across the world. We found in English-language social media that some users also viewed *Squid Game* through the lens of Japanese pop culture, which helped drive interest. Japanese pop culture acts as a point of reference that made *Squid Game* more accessible globally.

We can do topic modeling on all of this data and think about what shared topics emerge. We compared the conversations across different platforms, and the topics that emerged were different depending on the language and platform. For example, capitalism was a dominant topic in English and Korean discussions but less so in Japan, where people focused more on its genre elements.

We also worked with data from a third-party analysis firm to get a sense of how popular *Squid Game* was in different countries. We basically looked at reception data across 20 different countries over a five year period. What we found is that *Squid Game* was exceptional. It wasn't like other K dramas which have a very narrow band of popular response. They get high levels of attention in Asia, but *Squid Game* was the only one that had global coverage. And so then the question becomes: What other shows have global reach? And why is *Squid Game* in that same category? What can we learn by thinking about *Squid Game* as a show that achieved that level of success? What are the kinds of shows that appeal to audiences across 20 different countries at the same time?

2.3 Digital Humanities and Education

YK:

What do you think the digital humanities has brought to the college education landscape, especially in the humanities? What are the specific challenges in teaching DH at the University of Chicago and how do you deal with them?

HL:

We've developed a two-year MA program in digital studies at the University of Chicago, and that's allowed us to hire staff and faculty to teach the courses to support that. Through that MA program, we're able to offer courses to PhD students as well as undergraduates. That gives us a unique opportunity in that it's not just one professor in a department who has these skills but we have a broad range of faculty across many different disciplines. At the

undergraduate level, it's still a little unclear what's going to happen. I think undergraduates tend to know less about the digital humanities. They don't understand it as a field and so oftentimes it's through chance or serendipity that they come to learn that this is a thing they can do. So most of the courses I teach in digital humanities are for the MA program.

We have courses in a basic introduction to Python. We have courses in the basics of data analysis, basic statistics, and database courses, and so any graduate student can take those courses. And we have a PhD certificate. If there's a student who wants to learn those methods, they can come here and they can do that through the MA program.

For undergraduate students, it's less straightforward. Most of them don't have a background in programming or statistics, so they're less familiar with digital humanities. We need to do a better job introducing undergraduates to the possibilities of DH. We just haven't found a good way to do that yet. We have a lot of students who are majoring in computer science and love humanity's courses and I see those students quite a lot, but they aren't at a point where they can figure out how to do those two things together.

The new media arts and design program is a step in the right direction, as it covers things like video games, illustration and visual design, and there it's much easier for students to be working with software and various platform technologies to do creative work. And so it's right there that we have students who are working both with creative materials and also deeply involved with programming. There's also a lot of potential in HCI (Human-Computer Interaction) and redesigning digital tools using humanistic knowledge to support creative work. It's an interesting space where DH can make a big impact. The University of Chicago is in a good position to have both digital media and design studies. It's called Media Arts and Design. And it is the fastest growing major in the humanities.

2.4 Digital Humanities and Artificial Intelligence

YK:

What new directions or trends in DH research are you most excited about? Are there areas of research you feel are currently underexplored but have great potential?

HL:

Large language models are one of the most exciting new directions. Over the last year, I've been pulled into projects using these models and trying to understand their implications for humanistic research. One challenge is understanding what these models know about cultural history and how they can help us see things we wouldn't normally access.

I believe humanists have a role to play in developing tests and evaluations to assess these models' cultural knowledge. Computer scientists are really good at figuring out how to evaluate these tools. I think we should be borrowing some of their methods but apply them to humanistic questions. It's a different type of research, but it's where we can add a lot of value. We need to assess these large language models' understanding of poetry, voices, styles, and their ability to adopt different cultural contexts.

Once we understand what these models can do, we could train them on specific literary data to make them more useful for humanistic research. For example, training a model on 20th-century novels and asking it to identify characters, actions, and relationships could be very valuable for distant reading.

Let's say you have a model that is really good at identifying characters, their gender and actions. Now you can apply that to 10,000 novels. Maybe you have a very specific question in mind and now all of a sudden you're back to distant reading, but you have a model that can answer more complex questions. So you're not relying on the linguistic structure or NLP methods, but you're using what the language model knows and what it can recognize

to do distant reading.

There are already scholars who are exploring that potential. So rather than a traditional machine learning classification approach, I've seen some scholars who are testing GPT to do classification. For instance, if we have 100 haiku texts and another 100 texts that are not haiku, we may ask LLM to tell the difference.

And in some cases it can do a pretty good job of replicating what we were able to do with traditional machine learning techniques. If that becomes possible, you could see where and how it might be advantageous because you get to ask more complex questions. It potentially then becomes easier for a humanist who doesn't know how to do machine learning as they know how to talk to a language model. After we have a good model, then we have something to apply to real humanistic research.

We're not quite there yet, but within five years, it might be possible. Scholars like Ted Underwood, David Bamman, Matthew Wilkens, and Melanie Walsh are already exploring the potential of large language models for humanistic research. It's an area where we can have a real impact.

I think there's a possibility of combining close reading with digital methods for more focused case studies. In assessing these models, we need to understand what they see in a particular passage compared to what we, as human readers, see. There's always a validation step, and close reading is deeply involved in this step, even if the goal is to get to a more broad-based analysis.

YK:

What impact do you see AI, particularly generative AI, having on digital humanities and humanities education more broadly? How have your colleagues responded to the rise of generative AI technologies in educational contexts, and what do you believe are the most effective and ethical ways to integrate these tools into teaching?

HL:

Regarding education, generative AI is having a huge impact. We need to think about how to use these tools responsibly. If you're teaching a writing course, it doesn't make sense to allow students to use these tools. I think there are creative ways we can bring the tools into our classroom. You can't just shut your eyes because students are using them, and we should help them develop the literacy around what these tools do.

In my own courses, I've tried to be creative. For example, in a freshman writing course, I don't allow students to use ChatGPT to write their essays, but I've wondered if it might help them become better readers. Most freshmen don't have a deep knowledge of the texts we're reading. ChatGPT could be a vehicle or at least a first step to help them expand their historical and cultural awareness. Oftentimes we just expect students to read a passage and reflect on it using their own history and background. But what if we could ask ChatGPT to impersonate a historical figure or ask it to play the role of somebody in a very different perspective and to read the same passage from that perspective. It may help expand the range of voices and types of reading and interpretation that we have in the classroom.

2.5 Digital Humanities 2026 Conference

YK:

This is the final question. The Digital Humanities 2026 Conference is set to take place in South Korea. What are your expectations for the conference? What kind of impact do you think it will have on the field of Digital Humanities, both in East Asian studies and on a broader, global scale?

HL:

Regarding DH 2026 in South Korea, I hope it brings together scholars from East Asia and North America who

work on East Asian topics. I also hope it shows the quality of work happening in East Asian studies to scholars from other regions who are not East Asian specialists. And it would be great if it could help grow the field in Korea as well. That just depends on how much participation you have from scholars who are outside the field. If you're able to create opportunities so that people not in the field can recognize its potential, that would be fantastic.