Beyond the Boundaries of Individual Universities: Allegiance to Digital Humanities Education in Korea

Lee, Jae-Yon
bukuwom@gmail.com
Ulsan National Institute of Science and Technology, Korea, Republic of (South Korea)

Kim, Yongsoo
vadoropupille@gmail.com
Hallym University, Korea, Republic of (South Korea)

Ryu, Su-Rin
surinryu@gmail.com
Gachon University, Korea, Republic of (South Korea)

Mun, Soo-Hyun
munshyun@hanyang.ac.kr
Hanyang University, Korea, Republic of (South Korea)

Kim, Hyounghun
h.kim@unist.ac.kr
Ulsan National Institute of Science and Technology, Korea, Republic of (South Korea)

Panel Abstract: This panel is about a digital humanities (DH) education project that will be experimented within Korea for the next two years. The goal is to collaborate in developing DH classes that cross individual universities and open those courses to participating institutes. To this end, scholars in humanities and engineering from five universities in different regions will work together with support from the National Research Foundation. The aim of this panel is to present the overall design of this educational project, outline four subjects to be developed, and seek advice from overseas experts who are already running similar courses or are interested in DH education.

DH research is currently increasing both quantitatively and qualitatively in Korea. However, the education to support DH research is still at its nascent stage. In order to present a desirable case of DH education by reshaping the traditions of humanistic inquiries with the digital technology that develops day by day, our participants intend to develop a model called “networked digital humanities education.” The network expresses the relationship between people, things, ideas, and other components in the form of the web, which allows us to revisit the existing order, form, and structure by incorporating hitherto unrecognized elements. For our project, we use this concept of the network in two ways: first, an interdisciplinary bridge to renew humanistic inquiries with the methods of social science and engineering, and second, an institutional platform for educational collaboration beyond the constrained setting of individual universities.

Instructors from four fields in humanities (Korean literature, English literature, German linguistics, European history) will work with computer science experts and develop a course in each area. “A Literary Understanding of Multimodal Generation” in Korean literature, for instance, raises a question as to how to generate the image-linked narratives through the multimodal training of Artificial Intelligence and how to interpret them with literary tradition. “English Literature and Network Analysis” traces the complicated linguistic, social, and regional interactions through social network analysis and finds an effective way of using this social scientific method in the pedagogy of English literature. “Applying German Corpus-Based Dialog Analysis to Machine Learning” questions how we can apply linguistic knowledge to improving the latest language models and in what ways building a German language dataset can be useful for this agenda. Finally, “Reconstructing Mr. Seseman’s World” in European history focuses on Mr. Seseman, the father of Heidi's friend, Clara, in the story, *Heidis Lehr-und Wanderjahre*, published in 1881. It traces the worldview constructed by a 19th-century middle-aged man in the fiction by analyzing big data, such as trial records, newspaper articles, geospatial information through various DH methods while linking his *weltanschauung* to the general social, economic, and historical changes of the time. This panel may provide insight for those who would collaborate to design DH courses with help from outside their own institution and run an inter-university micro-degree program.

Paper 1

Title: A Literary Understanding of Multimodal Generation

Author: Jae-Yon Lee and Hyounghun Kim

Affiliations: School of Liberal Arts, Ulsan National Institute of Science and Technology, Korea; Graduate School of AI, Ulsan National Institute of Science and Technology, Korea

Abstract: A language model (Bengio et al. 2000) is an algorithm that assigns probabilities to text sequences based on large-scale corpus. It allows programmers to generate text, summarize it, and even create particular genres of writing, such as newspaper articles and poems. Nowadays, this model is applied to sound or image data or both in conjunction with texts, which is called multimodal training **(**Ngiam et al. 2011). Engineers use this cutting-edge technology to generate stories with image or sound prompts, or vice versa. However, critics have yet to seriously engage with the textual results generated by multimodal training.

This DH course focuses on both text generation and interpretation. Specifically, it seeks to extract simple texts from camera images and train them using literary works, including poems and essays, to create short stories following a certain narrative arc. Images are processed by RasNet or Vision Transformer, and text is generated by GPT-2 or GPT-3 (Generative Pretrained Transformer-2 or -3)

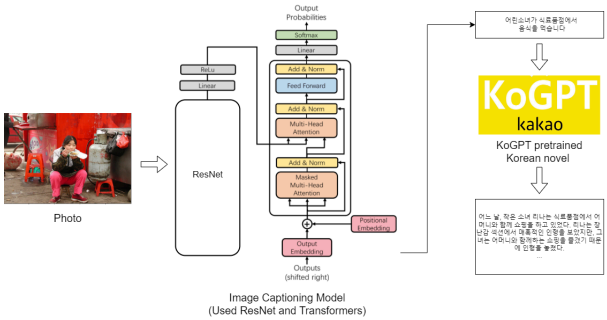

Image Captioning Model
(Used ResNet and Transformers)

Then, we will interpret the results from various critical perspectives, including Benjamin's flâneur as a lens to analyze capitalist modern culture (Benjamin 1999) and object-oriented ontology, which allows non-human beings to play a role as a social agent (Harman 2018). Students will understand the basic multimodal machine learning process in engineering that processes visual and linguistic data together. They will also consider what viewpoints to use when reading the unfamiliar expressions created by this artificial intelligence and what implications there may be in the act of reading.

Paper 2

Title: English Literature and Network Analysis: Toward Networked Digital Humanities Pedagogy in the South Korean Higher Education System

Author: Yongsoo Kim

Affiliations: Department of English Language and Literature, Hallym University, Korea

Abstract: My digital humanities course titled “English Literature and Network Analysis” is part of the larger project of developing a networked digital humanities educational model at the undergraduate level in South Korea. This course is designed to provide English major students with an opportunity to study digital humanities by focusing on social network analysis methodology. Literature is filled with relationships between words, texts, characters, authors, genres, and institutions, among others. These relationships can be effectively explored through network analysis, especially with the help of computer algorithms. Network analysis of literature involves natural language processing, data science, software literacy, data visualization, as well as a deep understanding of literary texts and conventions.

This course will be launched in the spring semester of 2023 in the Department of English Language and Literature at Hallym University. Students will have no prior experience of digital humanities, and so the course will be carefully designed to lead them to understand the relationship between literature and digital humanities, to study network analysis methodology, to apply the method to literary works, and to finally perform a DH project. My presentation will cover the course development process, the actual implementation of the course, the course evaluation, and the future plan to develop a networked undergraduate course of “English Literature and Network Analysis” between the four major South Korean universities. Student projects will be exhibited and archived online on a course project website. This website will evolve into a platform that connects students from different universities in the coming years.

Paper 3

Title: Applying German Corpus-Based Dialog Analysis to Machine Learning

Author: Su-Rin Ryu

Affiliations: Liberal Arts College, Gachon University, Korea

Abstract: We present a lecture model designed for Korean undergraduate students majoring in German language and literature, aimed at experiencing the application of linguistic knowledge to enhance the performance of the latest deep learning language models. Firstly, we construct a dataset by extracting dialogue sentences from the spoken German corpus (DGD) and analyze linguistic features focused on speech acts. By training several pre-trained language models with this dataset, we will experiment with performance measures such as conversation sequence prediction. Students will understand how useful linguistic analysis methods are for extracting various information from language resources and how they can be applied to the latest language application technologies while learning word embedding, tokenization, and other techniques using the PyTorch library. Furthermore, it will be possible to expand the subject of the lectures by using various language corpora such as Korean and English.

Paper 4

Title: Reconstructing Mr. Seseman’s World

Author: Soo-Hyun Mun

Affiliations: History Department, Hanyang University, Seoul, Korea

Abstract: Mr. Seseman, the father of Heidi's friend, Clara, in *Heidis Lehr-und Wanderjahre* published in 1881, is a white bourgeois male merchant from one European country, who symbolizes the main protagonist group in the modern Western world. In *Heidi*, Seseman is described as a man who usually stays in Paris on business, comes back home with many beautiful things, drinks black coffee, carries a revolver gun, and interacts with his family through letters. He is entirely estranged by Johanna Spyri, a female writer.

The goal of this course is to reconstruct Mr. Seseman’s world through DH approaches. The fact that the world of the 19th century European male citizen is one of the most popular topics in the history discipline makes it easy and suitable to focus on them for DH undergraduate course. Through DH methods of analyzing big data such as trial records and newspaper articles, etc., and visualizing it in various ways, it will be possible to clearly recognize the critical edge of the DH methods.

This course is comprised of three different parts. First, by analyzing statistical data accumulated in social history research, which shows economic expansion, population growth, urbanization, and social transformation, the students can find the overall characteristics of the 19th century more vividly (Mitchell 2003; Blaney et al. 2021). Secondly, we will catch a glimpse into his “Alltag” and professional world and through text mining approaches to the newspaper articles on the basis of “Zeitungsportal”, the British Newspaper Archive, and the Gallica-newspapers of France. Lastly, through QGIS, we will place the major historical events of the 19th century within a geospatial context.

Bibliography

Bengio, Yoshua / **Ducharme, Réjean** / **Vincent, Pascal** (2000): “A Neural Probabilistic Language Model”, in *Advances in Neural Information Processing Systems* , 13: 1137-1155.

Benjamin, Walter (1999) *Arcades Project.* Cambridge: Belknap Press of Harvard University Press.

Blaney, Jonathan / **Winters, Jane** / **Milligan, Sarah** / **Steer, Martin** (2021) *Doing Digital History: A Beginner's Guide to Working with Text as Data.* Manchester: Manchester University Press.

Harman, Graham (2018) Object-Oriented Ontology: A New Theory of Everything. New Orleans: Pelican.

Mitchell, B.R. (2003) *International Historical Statistics: Europe, 1750-2000.* London: Palgrave Macmillan.

Ngiam, Jiquan / **Khosla, Aditya** / **Kim, Mingyu** / **Nam, Juhan** / **Lee, Honglak** / **Ng Andrew Y.** (2011): “Multimodal deep learning”, in *Proceedings of the 28th international conference on machine learning (ICML-11)* : 689-696.