International Journal of Advanced Smart Convergence Vol.9 No.3 161-168 (2020)
http://dx.doi.org/10.7236/IJASC.2020.9.3.161

IJASC 20-3-20

The Future Past of Humanities Research : Musing Methodology in the Digital Convergence Era

Kim, Jiyun

Master(M.A.), Department of English Language and Literature, Yonsei University, Korea
jamesjoyce@yonsei.ac.kr

Abstract

Over the last half-century, computer science has revolutionarily changed the landscape of humanities research. This digital shift in research methodology has reached from the brainstorming process to preserving, constructing, collecting, visualizing, and even analyzing materials. Such transformation has brought about the birth of the new field of study: Digital Humanities (DH). DH undeniably has saved much of the physical chores and provided a new angle to interpret the text, thereby making its meteoric rise as a promising future of the humanities. Based on such innovation, electronic circuitry can seem to replace the imagination that detects relationships and significances of research data with ever-improving interfaces. However, despite hitherto technological development, the thousands-year-old essence of traditional liberal arts—human creativity—remains the heart of humanities research and always will. This paper starts by proving this proposition in the way of comparing the old and new liberal arts research methods, focusing on literary studies. Meanwhile, it thoroughly investigates how digitalized bibliographies, search engines, databases, and digital projects provide the most useful data preservation and virtual experience of browsing in the library, along with their limitations due to the intrinsic quality of humanities research data. Also, it probes the differences between traditional and digital data analysis in current methods of literary studies, ultimately presenting the ideal direction for humanities development in the era of digital convergence.

Keywords: *Digital Humanities, Humanities Research, Research Methodology, Literary scholarship, Computer-aided Research*

1. Introduction

> Icarus began to feel the joy
> Of beating wings in air and steered his course
> Beyond his father's lead. . . .
> Meanwhile, the heat of sun struck at his back.
> — Ovid, The Metamorphoses, VIII [1]

Since the initial databases appeared in the form of "punched cards with sequential file technology" in the

Manuscript Received: July. 18, 2020 / Revised: July. 24, 2020 / Accepted: August. 1, 2020
Corresponding Author: jamesjoyce@yonsei.ac.kr
Tel: +82-10-5106-5568, Fax: +82-02-392-0275
Master(M.A.), Department of English Language and Literature, Yonsei University, Korea

1940s and 1950s [2], computer science has made tremendous strides, granting wings to such humanities subjects as History, Philosophy, Linguistics, and Literature by introducing radical changes in research methodology. Its technology has allowed scholars to save enormous time and energy formerly spent on finding materials. Moreover, computer science has also proposed computer-assisted techniques for investigating materials, consequently begetting a new field of study: Digital Humanities (DH). While DH, the integration of Information and Communication Technologies (ICT) with the humanities, helps to visualize, interpret, analyze, and even construct the text, the thousands-year-old method of traditional liberal arts continues to go on strong, as reliable as ever. That is to say, digital technologies "serve, not replace, the imagination that detects relationships and significances that cannot be contained in electronic circuitry" [3]. No matter how much time passes, the essence of humanities—*human creativity*—should be the heart of its study. This paper will briefly compare the old and new liberal arts research methods focusing on literary studies and then investigate DH as Icarus's wings, a source of potential danger, and a set of tools that a researcher has to throw away after their utilization. Finally, it will present the ideal direction for humanities development in the era of digital convergence.

2. Discussion

2.1 Collecting Materials

Traditional literary research begins with a question in the form of a topic sentence, which encourages a researcher to define core ideas and concepts relevant to the research project. For instance, a researcher might come up with a sentence like 'I want to explore the feminist viewpoint on the 20th-century American Gothic novel.' The main concepts here are 'feminist literary criticism' and 'contemporary American Gothic literature.' This broad idea can be narrowed down to a more specific topic, such as the "identity crisis of minorities in *Interview with the Vampire* (1976) by Anne Rice." In this case, 'individual identity in society,' 'minority discourse,' or 'vampire as a minority' are the possible main concepts. After this process of brainstorming, the scholar will research the relevant materials in the library catalogs or archives.

While a conventional researcher might make a list of keywords by herself to help visualize the brainstorming process, her colleague in the new generation will carry out this step more straightforwardly and practically, using several search engines. Today's search engines connected with cloud databases are certainly made to provide not only a service but also management of data [4]. They automatically generate recommended keywords and curate relevant journals, books, media, and other sources in Online Public Access Catalogs (OPACs) and databases that may contribute to the research. Although Thomas Mann criticized machine sources for the way they were "generated only in response to specific words being typed in," using "rigid verbal criteria of exclusion" [5], the new relevance logic of the digital search has since made phenomenal progress. Now search engines such as ProQuest, JSTOR, and EBSCOhost minimize the influence of so-called 'stop words' like articles and prepositions. Moreover, they provide advanced search functions with Boolean operators (*e.g.*, "OR," "AND," "NOT," asterisks, quotation marks, and parentheses), and other identifiers such as source type, study fields, content subjects, publication date, length, citation index, and author information, sorting the most relevant of the search terms to the beginning. In addition to limiting and modifying the results, the system algorithm recommends documents on similar topics based on the scholar's search history. In this way, modern search engines have overcome the early models' syntactic and semantic failings and helped to make serendipity play a more meaningful role in the bibliographic search. Other than that, some Libraries are also have been "providing with social services and plays a role like a social media" [6]. With such enhanced accuracy, utility, and convenience, online searching presents a genuinely virtual experience of browsing in the library.

This advance in technologies has undoubtedly decreased the need to visit several libraries, allowing researchers to reference online databases quickly and effortlessly around the clock. A few examples of such databases would include Web of Science (WoS), Scopus, The Modern Language Association International Bibliography (MLAIB), WorldCat, and Annual Bibliography of English Language and Literature (ABELL). These collections store multiple types of microform material, including electronic books, scholarly journals, periodicals containing original illustrations, newspaper articles, and even manuscripts. As an enormous number of bibliographies and other source materials are now being produced, databases reserve them in electronic form, enabling scholars to identify suitable data from a massive pile of materials that one might otherwise never master *durante vita*. Such digitized storages serve a dual purpose: "they [preserve] rare and fragile materials from damage incurred through excessive use while providing wider access to them" [7]. Since digitization projects have become a reliable method of preserving items, microform bibliographies have become more readily available online, allowing the Web to play a pivotal role in archival scholarship. There already has been an explosion of interest in e-reading and e-book devices like the Kindle, the iPad, and the Nook with extensive text digitization projects being undertaken, the most well-known of course being Google Books (books.google.co.kr) and Project Gutenberg (www.gutenberg.org). Also, scholars like Franco Moretti have taken up "data mining and visualization to perform 'distance readings' of hundreds, thousands, or even millions of books at a time" [8]. Not to mention that there are a myriad of valuable sites created by government authorities (*e.g.*, Library of Congress, United States Census Bureau), recognized organizations (*e.g.*, the American Literature Association, the American Philosophical Association), educational institutions (*e.g.*, universities, libraries, research centers, academic departments, scholarly projects), and individual scholars. Some of these websites function as 'meta-pages' or 'gateways' that provide links to other relevant sites.

Nevertheless, "despite technological advances, the nature of archival materials themselves will not change; a literary manuscript will always be a manuscript" [7]. In other words, the essential quality of all archived material, which is fundamentally analog, will remain the same even though the advance in technology allows its transformation into electronic form. Of course, from around the 1980s, when Stephen King wrote *Word Processor of the Gods* (1983) with his Wang word processor, at least some archive material has been produced in digital form. However, although archives are progressively digitizing their material and making this available for researchers to consult online or to download, many of them, including Archives Hub (www.archiveshub.ac.uk), National Archives (www.nationalarchives.gov.uk), and Access to Archives (A2A; www.nationalarchives.gov.uk/a2a), remain in a preliminary phase; they reproduce only a small portion of their materials in digital form and utilize fundamental web tools to increase the researcher's ability to figure out and navigate relevant materials. Thus, scholars still have to plan a trip for most pre-21st century archival research with patience and preparedness, although new technologies have removed much of the legwork. Besides, other traditional stages of research, such as consulting with experts in the field or participating in research communities, also still hold key positions in contemporary research. Therefore, the principal merit of computer science in the first step of research, data collection, remains confined to helping the traditional process be more convenient and time-saving; it has not created a new work routine entirely.

One may object to this statement by arguing that since new digital projects will continue to appear with ever-improving interfaces, a completely new data collection method may appear as a game-changer and that eventually, this will generate a whole different way of research. There indeed are many respectable digital projects; one example of them is the Rossetti Archive (www.rossettiarchive.org) which renders high-quality digital images of Dante Gabriel Rossetti's manuscripts and paintings along with detailed physical descriptions, histories, provenance, literary contexts, and bibliographies. By doing so, it allows researchers to see at a glance the eclectic mix of information about a particular material. It also plans to provide customized online exhibition

in the future. Another example is the Australian Web Archive (trove.nla.gov.au). This archive brings together the diverse set of Australian digital collections, including the PANDORA Archive and the Australian Government Web Archive, through a single interface. This immense database provides sound, video, manuscripts, and other kinds of extensive data sets in collaboration with multiple libraries and cultural collecting organizations. However, although these two cases embody the innovative model of digitization projects and even more revolutionary ones will continue to appear, scholars still have to decide what to type in the search box and do some post tasks such as modifying the search results. That is, the skills needed for gathering data effectively and efficiently, which are the researcher's judgment, creative ideas, readiness, and perseverance, remain as before; these still require human intelligence. Thus, no matter how technology develops, gathering research materials remains intrinsically the same activity as it was in earlier periods of scholarship.

2.2 Analyzing the Data

A similar story may be told about hypothesis-formation. This procedure covers note-taking, analyzing information, and generating original thoughts. While many conservative scholars prefer taking notes by hand on memo pads, sheets of paper, or index cards, there are nowadays many kinds of digital note-taking tools and reference managers such as Zotero and Endnote. These programs help researchers to store, manage, and cite bibliographic references, with their customized recording functions. If the researcher is digital-friendly and familiar with their operations, they are indeed powerful tools for collecting and organizing research information. Otherwise, the old ways are still sufficient, proving that digital technology is merely a handy option for researchers.

In terms of current methods of literary analysis, there are mainly two types: qualitative and quantitative. While a classic researcher concentrates on the former, a contemporary colleague who is familiar with digital technologies—and thus probably a member of the DH field—will give attention to the latter as well. If a traditional researcher wants to investigate how James Joyce portrayed the "paralyzed" state of Dublin at the turn of the century [9], for example, a close reading of the characters' utterances, behavior, and optokinetic descriptions in *Dubliners* (1914) from a psychoanalytic and historical view might be a useful starting point. A DH scholar, on the other hand, might generate a 'volume plot graph,' which shows the density of imagery using Voyant (voyant-tools.org) and compare the number of image words with their occurrences in the short story collection, visualizing a large-scale landscape of Joyce's Dublin. Besides, he can employ Franco Moretti's network theory by making visible "specific 'regions' [of character] within the plot" [10]. As seen in this case, classic literary research concentrates more on the innate qualitative traits of the literary elements and the humanistic insights regarding the relation of images and context; in contrast, DH works on "the revelation of patterns and structures which would be impossible to discern with the naked eye" [11]. The latter represents a quantitative method.

The most common practice of digital literary analysis is to classify the most frequently used and shared words in a visual graph. Such digital analysis has the appearance of the microscopic examination of the cellular structure of the text. Two examples of such graphs made by the Voyant tool are figures 1 and 2 below. Figure 1 is a 'word-cloud' of this paper, which shows the most frequently occurring words. Each word's size indicates its stake in the whole corpus, providing a convenient overview of the material. Another specimen is figure 2, a 'context graph' of this paper, which displays the network of higher frequency terms in the document. It helps to figure out the relationship between key concepts and their collocates. Any Voyant user can make this word-cloud instantly by just inserting the text to the tool. Other than calculating the number of specific keywords and the types of semantic categories in a novel/genre/group of authors, there are numerous other means of data

Figure 1. A Word-cloud of this paper

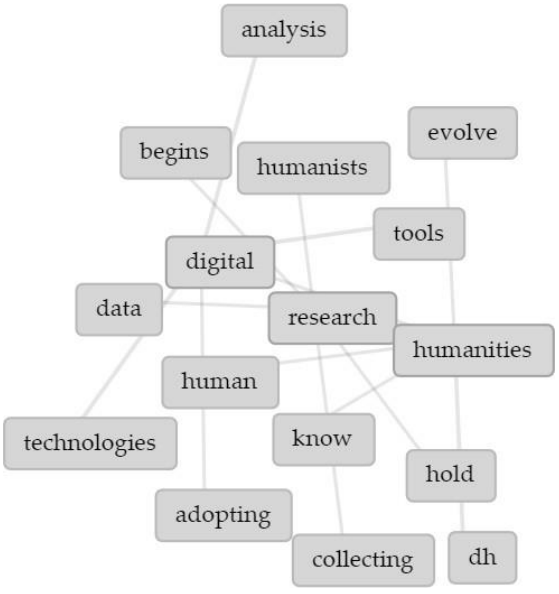

Figure 2. A Context Graph of this paper

mining, all based on the researcher's needs. For instance, a researcher may analyze authorship through computational elements of texts" [12] or investigate the proportion of authors who feature in the top 50 for academic and newspaper criticism [13]. Furthermore, identifying the most critical passages by creating a section-by-section map indexed to concepts of interest with Tableau Public (public.tableau.com) is also possible [14]. In this way, DH draws on a wealth of visualization with multi-variant statistical techniques and large-scale analysis.

2.3 The Final Stage

The analysis itself, however, is not the destination of the research. Instead, a researcher should excavate new knowledge from these distinct dispositions of information and images, using evidence from accredited methods of interpretation. For discovering hidden implications, digital analysis has several virtues. Before anything else, it provides a new perspective on previously unknown aspects of the material. This digital epiphany is significant because one of the humanities' roles is to breathe new life into past research. Thanks to the humanists shedding new light on previous work, literature can revive itself, passing on the lamp to new scholars. That is to say, computer-aided analysis suggests a valuable new way to look at the old objects, such as *Beowulf* (*ca*. 700-1000 AD) or Shakespeare's *Hamlet* (*ca*. 1599).

Furthermore, by adopting direct visualization, digital analysis creates new possibilities for liberal arts to communicate with scholars from the disciplines of the natural and social sciences, scholars whom humanists do not usually encounter at conferences. For a long time, "because of its highly individualistic, sometimes idiosyncratic nature" [15], the humanities have been regarded as having crucially different attributes from the scientific disciplines. Now, at last, by employing computerized research, humanists can convey their theories in the language of computer science. This digital shift is meaningful since interdisciplinary conversations widen the discourse and strengthen the viability of the discipline; the humanities can now contribute both to its own ongoing legacy and the broader goal of a holistic human intelligence.

These advantages of computerized analysis make DH seem irresistibly attractive to researchers, along with

other benefits of collecting and storing research data, which is why it is not an exaggeration to call computer science the 'wings' of humanities. Nevertheless, DH also has some significant weaknesses. These weaknesses sometimes make it seem as though these wings are, like those of Icarus, made of wax. In consequence, these are central matters that require attention. One of them is that in digitized research, the agents who shape the ideas, build the methods, and interpret the results are often different persons. This division of labor is practically unavoidable because "humanities scholars know little about programming and software engineering, and programmers know little about humanities scholarship. Going it alone is an option only for the few who have time to master both" [16]. Such collaborative work in DH minimizes the place of solitary meditation, which has long accompanied creative accomplishment in the humanities and often interrupts the flow of research, making it a thorny group project with individual risks. As René Descartes describes in his *Discourse on Method* (1637), "there is often not so much perfection in works composed of many pieces and made by the hands of various master craftsmen as there is in those works on which but a single individual has worked" [17].

However, perhaps the most vulnerable point of DH is that it tends to focus more on the methodology than on its potential for humanistic insights. Rather than discussing the implications of the evidence, DH research tends to spend much of its energy in demonstrating the effectiveness of its statistical techniques and digitized analysis. Such 'putting the cart before the horse' is inevitable for DH to some degree; DH researchers have to face the problem of convincing their audience of the legitimacy of their computer-based analysis. An example that displays such a difficulty is Étienne Brunet's "L'exploitation des grand corpus: Le bestiaire de la littérature française" (1989). In the paper, he aims to "compter les mots, au lieu de les peser! Compter les bêtes, au lieu de les caresser!" ("count words rather than weigh them! Count the beasts rather than caress them!"; [18]), focusing on computational analysis and its description. However, Brunet's studies "did not have the hoped-for effect," although he anticipated that "computer-assisted analysis would be an effective critical method when supported by visualization" [16]. His paper was not successful because the eventual conclusion did not possess sufficient gravity, regardless of the usefulness of its digital tools. Under all circumstances, the creative insights of the human mind derived from the analysis should be the main point of the research. In the same vein, humanists generally avoid explaining their research methods in long discussions since it can distract the audience from the ultimate goal of literary interpretation. As in Brunet's case, computer technology cannot be the deputy for human intelligence in the final phase of the research. Researchers should be wary of such challenges given by computer science rather than blindly accepting it; just because it is new does not mean it is ideal.

Even though computer-aided analysis poses a threat to the humanities as in 'too-many-cooks-spoiling-the-broth' or a 'tail wagging the dog,' there is still an apparent motive for humanists to keep an eye on it. Leaving aside the discussed advantages of DH—providing an innovative angle, enabling large-scale analysis, and opening the door for interdisciplinary exchanges—, the humanities must evolve in line with the main branches of knowledge research to understand the current human condition better. Willard McCarty says about the humanities' conservative resistance to DH that "although better tools are possible, the humanist's perspective on tools problematizes them" [19]. That means traditional humanists ought to overcome their inherent wariness of machinery and embrace the technical improvement as devoted 'tools,' paradoxically, in order to continue their humanistic tradition. Simultaneously, to form an ideal partnership with advanced technology, humanist researchers must remember that the core value of the humanities lies in human intelligence, not in computers. They should utilize software as a ladder, a means for getting faster to knowledge, and discard it once they achieve their goal, as Ludwig Wittgenstein manifested in *Tractatus Logico-Philosophicus* (1921): "He must so to speak throw away the ladder, after he has climbed up on it" [20]. Computer science is lending a hand to humanities, with a broader vision and the connection to today's world. In terms of analyzing the materials, digital instruments can add a quantitative point of view to the classic qualitative inference. However,

when adopting digital analytics, a researcher must be careful since it is up to him or her whether the result will demonstrate the art of Daedalus or the folly of Icarus.

3. Conclusion

Beyond every challenge and opportunity before us, the lesson from the last decades is evident. That is, the role of human intelligence and creativity as a director of the project is irreplaceable. This verdict makes the conservative views of Altick and Fenstermaker still valid even after almost thirty years; digital technologies can only help initiate or facilitate humanities research, and after this, human intelligence must take over. However computer develops, it cannot derive the final insight for humanities by proxy of scholars. Nonetheless, it is also apparent that humanities should continuously evolve, accepting digital research tools where they are needed. In order not to be buried in the past, humanities should take technological changes on board. Eventually, humanist research ought to find a guide into tomorrow based on its own tradition; in this way, it can rediscover the essential value of humanities by means of the future past.

References

[1] Nina Kossman, *Gods and Mortals: Modern Poems on Classical Myths*. Oxford UP, p. 148, 2001.

[2] Charles W. Bachman, "The Programmer as Navigator," *Communications of the ACM (CACM)*, vol.16, no.11, pp. 653–658, Jan. 1973.
DOI: https://doi.org/10.1145/355611.362534.

[3] Richard Altick and Jon J. Fenstermaker, *The Art of Literary Research*, 4th ed., W.W. Norton & Company, pp. 181, 1993.

[4] Sungbin Park, Jong-Yong Lee, and Kye-Dong Jung, "The Design of Library System using the Cloud Environment Based on the Raspberry pi," ," *International Journal of Advanced Smart Convergence (IJASC)*, vol.4, no.1, pp.31-34, 2015.
DOI: https://doi.org/10.7236/IJASC.2015.4.1.31

[5] Thomas Mann, *Library Research Methods*. Oxford UP, pp.73, 1987.

[6] Toshiro Minami and Kensuke Baba, "A study on Finding Potential Group of Patrons from Library's Loan Records," *International Journal of Advanced Smart Convergence (IJASC)*, vol.2, no.2, pp.23-26, 2013.
DOI: http://dx.doi.org/10.7236/IJASC.2013.2.2.23

[7] J. Greg Matthews, *Literary Research and Irish Literature: Strategies and Sources*. Scarecrow Press, pp.123-155, 2009.

[8] Matthew G. Kirschenbaum, "What Is Digital Humanities and What's It Doing in English Departments?" *Defining Digital Humanities: A Reader*, Ed. M. Terras, J. Nyhan, and E. Vanhoutte. Ashgate, pp.195-204, 2013.

[9] James Joyce, *Dubliners: Authoritative Text, Contexts, Criticism*, Ed. Margot Norris, W. W. Norton, p. ix, 2006.

[10] Franco Moretti, *Distant Reading*. Verso, pp.217, 2013.

[11] Leighton Evans and Sian Rees, "An interpretation of Digital Humanities," *Understanding Digital Humanities*, Ed. D.M. Berry, Palgrave Macmillan, pp. 21-41, 2012.

[12] Martin Paul Eve, *Close Reading with Computers: Textual Scholarship, Computational Formalism, and David Mitchell's Cloud Atlas*, Stanford UP, pp.69, 2019.

[13] Katherine Bode and Tara Murphy, "Methods and Canons: An Interdisciplinary Excursion," *Advancing Digital Humanities: Research, Methods, Theories*, Ed. P.L. Arthur and K. Bode, Palgrave Macmillan, pp. 175-193, 2014.

[14] Mark Alfano, "Digital Humanities for History of Philosophy: A Case Study on Nietzsche," *Research Methods for the Digital Humanities*, Ed. I. Ievenberg, T. Neilson, and D. Rheams, Palgrave Macmillan, pp. 85-101, 2018.

[15] Charles L. Gilreath, *Computerized Literature Searching: Research Strategies and Databases*, Westview Press, pp.143, 1984.

[16] Geoffrey Rockwell and Stéfan Sinclair, *Hermeneutica: Computer-Assisted Interpretation in the Humanities*, The

MIT Press, pp.7-95, 2016.

[17] René Descartes, *Discourse on Method and Meditations on First Philosophy*, Tr. D. A. Cress, 4th ed., Hacket Publishing Company, Inc., pp.7, 1998.

[18] Étienne Brunet, "L'exploitation des grands corpus: Le bestiaire de la litterature francaise," *Literary and Linguistic Computing (L&LC)*, vol.4, no.2, pp. 121-134, 1989.
DOI: https://doi.org/10.1093/llc/4.2.121.

[19] Willard McCarty, "Beyond the Word: Modeling Literary Context," 2007.
http://www.mccarty.org.uk/essays/McCarty,%20Beyond%20the%20word.pdf.

[20] Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*. Ed. Bertrand Russell, The Humanities Press, pp.189, 1951.