

“Ars longa, vita brevis” - a saying whose meaning is often debated by its interpreters - art is long, but life is short. Art has grown timelessly over the centuries, and too often our lives are not long enough to explore the towering forms and pieces of creativity out there. Dedicating one’s life to the relentless studying of great literature and other pieces of art is not an easy feat. This quandary has ushered some individuals to look for a new means of effectively and efficiently analyzing literature. With one tried and true method, close reading, to a newer unorthodox approach, distant reading, we will look at these two means of analyzing literature to better understand these concepts through the benefits and limitations of both.

Through Kathryn Schulz’s “What is Distant Reading?” and Alan Liu’s “The State of the Digital Humanities”, we learn about what distant reading is, along with the important and underlying issues, trade offs, and benefits of this powerful yet double-edged tool. Looking closer in Lisa Rhody’s “Topic Modeling and Figurative Language” and Matthew Wilken’s “Canons, Close Reading, and the Evolution of Method”, we get to explore distant reading in action and see what findings each author stumbles upon.

Distant reading is a newer form of literary analysis performed in a computer program by aggregating and analyzing text; the program identifies literary genres and envisions plots by mining through each word and phrase. Franco Moretti, who founded the Stanford Literary Lab and also coined the term “distant reading”, claims that simply reading texts is archaic and is insufficient for understanding the full scope of the material. He argues that one’s exposure would be limited to a small sample size of texts when hundreds in the same topic exist. Utilizing distant reading and text mining, this opens up endless possibilities of better understanding the scope and nature of numerous works and topics.

However, there are underlying concerns and limitations with this method of textual analysis, and others argue that close reading is still preferable. Assuming that there is a generic structure throughout all texts diminishes literature to a simple arrangement of words following a set of rules lacking individual uniqueness and profound themes. With close reading, one can interpret texts through detailed observation and reflection, develop new insights and ideas, and have a deeper appreciation for the humanities.

In Wilken’s “Canons, Close Reading, and the Evolution of Method”, he explored the benefits of fairly analyzing a significant amount of texts from respective topics instead of resorting to only “canon” works and common tropes. Utilizing Geodict, a distant reading app using geospatial analysis and text extraction, on the entirety of works in the Wright American Fiction Collection corpus, he visualized a map of locations mentioned in American fiction from before and after the Civil War. Through this, he found an interesting trend of texts mentioning other countries, implying America’s imaginative landscape to be more outward than inward, contrary to the notion of America at the time. In addition, he noticed growing mentions of Western US, shedding new light on American regionalism during this period. He concluded that compelling observations can be found with an expansive and balanced dataset, even potentially subverting biases established by “canonical” works. Wilken does make a critical point that his continual use of computational text mining made him unlikely to close read when it could have yielded useful findings, and that his reliance on distant reading risked degrading his close reading skills.

In “Topic Modeling and Figurative Language”, Rhody tested the LDA (Latent Dirichlet Allocation)’s capability of categorizing and formulating topics on a corpus of ekphrasis poems, recognized for its metaphors and figurative references of visual arts. She accentuated the algorithm’s accuracy and potency with non-fictional works, such as the *Science* magazine, due to the texts’ practicality and lack of underlying context. In contrast, Rhody revealed odd topics that the LDA modeled with the poems, such as OCR (optical character recognition; words spelled wrong or non-English), ‘chunk topics’ (words specialized for a specific poem), ‘evident’ topics (words that differ from their literal meaning and their manner of speech within the poem), and ‘opaque’ topics (words loosely related with one another). A clear defect with distant reading is shown with LDA’s ‘evident’ topics, illustrating the machine’s incapability of interpreting figurative meaning. However, Rhody concluded that such unreliable and seemingly inconsistent data, especially ‘opaque’ topics, may prove useful if one were to switch to close reading.

Distant reading gives us the capability to categorize, analyze, and digitize entire collections, allowing us to absorb information faster than ever before. However, while distant reading may prove beneficial for studying literature and yield more information than from close reading, we should not be entirely dependent on these computational tools as we may lose the vitality and deeper meaning of these texts, and at times could yield shallow and inaccurate results. Although close reading limits individuals’ time, energy, and scope, interpretations through close reading have led to profound insights and advancements in literature, unique styles and themes derived from texts, and personal connections to the works with which we engage.

Dr. Nan Z. Da’s “The Digital Humanities Debacle” highlights many of the issues with digital humanities. The issue of the computational methods’ inability to explain literary complexity is similar to the shortcomings of distant reading; there is a difficult tradeoff between precision and efficiency. She states, “Even if you *do* manage to train a computer to detect literary phenomena that come in entirely different forms and lengths, varying vastly in content, you still have too few of them for any kind of aggregate analysis.” It seems as though there is still no computationally effective way to analyze large amounts of text without sacrificing the important nuances in literature. Our stance, however, disagrees with Dr. Nan Z. Da’s opinion. As explored throughout the findings mentioned above, distant reading is a practical tool best used wisely; it is more helpful to be cognizant of its current limitations, not discard it, and continue using it to our advantage while also letting close reading cover its pitfalls.

## References

Liu, Alan. "The State of the Digital Humanities." *Arts and Humanities in Higher Education*, vol. 11, no. 1-2, 2012, pp. 8–41., doi:10.1177/1474022211427364.

Rhody, Lisa. "Topic Modeling and Figurative Language." *Journal of Digital Humanities*, 2012, <http://journalofdigitalhumanities.org/2-1/topic-modeling-and-figurative-language-by-lisa-m-rhody//>.

Schulz, Kathryn. "What Is Distant Reading?" *The New York Times*, The Mechanic Muse, 24 June 2011, [www.nytimes.com/2011/06/26/books/review/the-mechanic-muse-what-is-distant-reading.html](http://www.nytimes.com/2011/06/26/books/review/the-mechanic-muse-what-is-distant-reading.html).

Wilkins, Matthew. "Ch. 14: Canons, Close Reading, and the Evolution of Method." *Debates in the Digital Humanities*, edited by Matthew K. Gold, University of Minnesota Press, 2019, <http://dhdebates.gc.cuny.edu/read/untitled-88c11800-9446-469b-a3be-3fdb36bfbd1e/section/6c7cbaa1-5ff8-4439-9ffb-aecbc6d5734>.