***A Critical Review of Chapter 1 of The Dangerous Art of Text Mining: A Methodology for Digital History***

For my review, I have read Chapter 1 of The Dangerous Art of Text Mining: A Methodology for Digital History (2024) by Jo Guldi. When delving deeper into information about this text, I was unaware that it was so recently published. Chronologically, the text is mere weeks in circulation as of this writing and is unlikely to have yet been subject to many critical reviews.
 Digital libraries are often seen as equivalent to physical historical documents, at least in terms of the average person I have met. “Why go to a library when everything is on Google?” is a phrase I have often encountered during the process of obtaining my MLIS degree. I have struggled to answer this question quickly, particularly when I can’t use library jargon. Based on the length of this book (and in fact, each chapter), I suspect this question *can not* be answered promptly and arguably never will be, due to many factors related to the nature of studying human history overall.

While this book (at least in Chapter 1) provides some hope, clues, and answers to the pitfalls of modern data analysis, this chapter included no detailed solutions and was meandering and long-winded. I attempted to pair this reading of Chapter 1 with a reading of the Preface and Introduction sections. The result of this endeavor is that I lacked the time to read the entirety of what is ultimately a deeply illuminating book due to the sheer length of each section.

Guldi (2024) repeats throughout the chapter that a humanist approach is essential to simultaneously understanding and interpreting data and digital archives. However, this analysis seems incredibly general and overly judgemental of data analysts who do not share the author’s experience. Seemingly lucky enough to have the best of both worlds, Guidi almost smugly indicates that the humanist and historical approach is the superior one, and while I agree with many of the views postulated, it seems that this incredibly long book could have been shortened to an article or otherwise summarized in a pithier format. During this critical period of increasingly short attention spans, this book seems to meander hopelessly, waxing poetic on the virtues of humanism, and is entirely unfocused on immediate solutions.

Surprisingly, despite the information above, I agree with most of the points made in this chapter. It would be prudent, however, for this author to write a short crash course on how to make these concepts a reality, beyond mere speculation. However, the author may eventually state some (or all) of the things I am critical of being absent from this chapter.

One of the eye-opening things that I took away from the excerpt of this text is that historical context is key to understanding and analyzing data appropriately and that the conception of occlusion exists– that is to say, a blurring and misunderstanding of the true context and meaning of an event in favor of focusing on quantitative data alone. As a student and enthusiast of creating a positive user experience, occlusion seems to be a much-overlooked but vital factor in analyzing digital library data, archives, and even metadata.

A quick dictionary search defines occlusion as a “process whereby something is hidden or obscured from prominence or view.” In Guldi’s (2024) assessment, the lack of a humanist approach in data analysis is postulated to have led to a higher-than-desirable level of occlusion. Guldi (2024) states that while no amount of study or historical data, humanist or otherwise, can completely remove occlusion, the humanist approach greatly remedies this phenomenon. I’m not entirely sure I agree wholeheartedly with that assessment, but perhaps I should finish the book before making a final judgment.

***References***

Guldi, J. (2023). The dangerous art of text mining. In *The dangerous art of text mining* (pp. 27-56). Yale University Press.