Introduction

The term 'distant reading' resonates across DH: It is played on in book titles (Distant Horizons, Underwood 2019) and adapted to new fields ('Distant Viewing', Arnold and Tilton 2019). It spurs alternative formulations ('Scalable Reading', Mueller 2012) and is present in mainstream media ("What is Distant Reading?", Schulz 2011). It is a popular and integrating term, but can take very specific meaning as well.¹

However, the semantic content carried over in each case of adoption or adaption is often unclear. Recent debates, like the special issue of PMLA (On Franco Moretti’s Distant Reading 2017) or the paper by Nan Z. Da (Da 2019) and the reactions to it, have challenged some of the assumptions of 'distant reading'. Also, the polysemy of the term may have contributed to misunderstandings in these debates.

Therefore, our aim is to recover the historicity of the term 'distant reading', first introduced by Franco Moretti (2000) in his discussion of world literature as a system, by delineating

¹ This contribution has emerged from work in the COST Action "Distant Reading for European Literary History" (CA16204, https://distant-reading.net), a pan-European, collaborative networking project launched in 2017.
how its meaning has changed over time and reconstructing some of the key theoretical assumptions it carries both as a term, a concept and a practice.

**Historical roots**

The pre-history to the concept now covered by the term 'distant reading' reaches back to the 15th century, when a rhetorical topos of "too many books" appeared (see Blair 2011). The solution was in excerpts and encyclopedias, based on the principles of compilation and summarization. The goal was to provide access to the essence of all relevant books instead of having to see them all at the same time. Of course, quantitative approaches to literary texts have appeared before the advent of computing (e.g. Mendenhall 1887) and computational approaches have diversified before the term 'distant reading' appeared (e.g. Ellegård 1962, Mosteller and Wallace 1963, Burrows 1987; see Hockey 2000).

**Conceptual Development**

When Franco Moretti first coined the term 'distant reading' in 2000, he used it with a meaning reminiscent of the compilatory origins of the concept, similar to "second-hand reading": using research literature, metadata or other short-cuts like titles and subtitles instead of reading the full text. From this starting point, and in parallel with more computational and more quantitative practices, Distant Reading has evolved to designate any computational, but especially quantitative, method of literary text analysis - so much so that the term now 'self-evidently implies computation' (Goldstone 2017, 637; see also Underwood 2017 and Bode 2017).

**Theoretical Assumptions**

A fundamental assumption of the earlier concept of 'distant reading' was that because metadata or secondary literature are created by humans who have read the full texts, they can stand in for the full text. Also, that the bird’s eye’s view provides insight into the longue durée and into literature as a system (Oberhelman 2015). A fundamental assumption of current Distant Reading research is that useful (even if imperfect) formal and quantifiable textual features can be used as indicators or proxies for relevant literary phenomena, hence the centrality of modeling (see McCarty 2005; Flanders and Jannidis 2019) in Distant Reading research practice. Finally, the idea that despite the broadening meaning of the term “literature” (decanonization), literary texts have a specific way of functioning that requires the adaptation of methods to this domain.
Conclusion

We hope that by more usefully contextualizing the development of the strategic term 'distant reading', we can help avoid misunderstandings in current debates about computational approaches in humanistic inquiry.

Bibliography


