

Quantifying Literature: It's Not All About the Digital

Marie-Christine Boucher
Justus-Liebig-Universität Gießen

marie.c.boucher@ggk.uni-giessen.de

Abstract:

Ted Underwood's *Distant Horizons: Digital Evidence and Literary Change* sets off for a pacifist mission to reconcile traditional literary studies and distant reading, by recentering the debate around statistical methods borrowed from social sciences, away from buzzwords and polemical statements. In doing so, he makes a case for new discoveries in the field of literary studies that – even though they have been made easier by recent advances in access to digital libraries and machine learning algorithms – are really mostly based on methods that have been around for at least half a century. Perhaps the biggest strength of Underwood's newest work is that it could convince even the fiercest traditionalists that quantitative analysis should have its place in literary studies.

How to cite:

Boucher, Marie-Christine: „ Quantifying Literature: It's Not All About the Digital [Review of: Underwood, Ted: *Distant Horizons. Digital Evidence and Literary Change*. Chicago/London: The University of Chicago Press, 2019.]“. In: KULT_online 61 (2020).

DOI: <https://doi.org/10.22029/ko.2020.1025>



Quantifying Literature: It's Not All About the Digital

Marie-Christine Boucher
Justus-Liebig-Universität Gießen

Underwood, Ted. *Distant Horizons: Digital Evidence and Literary Change*. Chicago/London: The University of Chicago Press, 2019. xxii, 206 pages, 27,50 USD. ISBN: 978-0-226-61297-3.

It does not seem like a coincidence that the cover of Ted Underwood's new monograph displays a pastel sunset. If *distant reading* and *digital humanities* tend to be associated with brand-new methods and technologies, as well as broad, polemical statements, this book adopts a much softer tone and seeks to achieve exactly the opposite. Underwood repeatedly refers to the controversies stirred by Franco Moretti's *Slaughterhouse of Literature* (MLQ: Modern Language Quarterly, 61:1, March 2000) at the beginning of the 21st century, and weaves his own argument around it as a positive antithesis: Rather than redefining his method as *not* distant reading, the author uses the term sparingly, and avoids technical jargon and references to technology as much as possible, in an effort to displace the current debate around the method to another level. Distant reading, Underwood argues, should not be debated with new technologies at the center of the argument: "[T]o frame a fair critique of distant reading, one might ask whether it simply moves further toward social science than most literary scholars are willing to go." (p. 161)

Underwood, who is a Professor of English and Information Sciences at the University of Illinois at Urbana-Champaign, offers a very balanced and careful argument for the use of distant reading, not as a revolution, not as something new that should come to replace what literary studies already does. The author is clearly aware that he is not preaching to the choir, and that his audience might very well be resistant to his argument, and he offers a 'positive defense' of what literary studies already does "very well" (p. 147). He does not argue that most literary scholars should (or ever will) become distant readers, or even that it would be possible to implement these techniques in the near future – his defense is rather of the cautious kind, as he proposes to open the field to distant reading while "limit[ing] the authority of numbers in the humanities" (p. 147).

The first and last chapter form parentheses around the three middle chapters, which are devoted to analyses. In the first chapter, Underwood offers a careful description of what it means to frame statistical models of a literary concept. His main argument in favor of those methods for the field of literary history is that a “single pair of eyes at ground level can’t grasp the curve of the horizon, and arguments limited by a single reader’s memory can’t reveal the largest patterns organizing literary history” (p. x). The last chapter is devoted to the risks inherent in the practice of distant reading: the risk of forgetting pleasure, of preoccupation with technology, and that it may exceed the structure of current university departments.

Chapters 2, 3, and 4 are dedicated to three major concepts of literary history: genres, prestige, and gender. Taking great care to address the very common argument against quantitative methods – ‘but, didn’t we know that already?’ –, Underwood argues for the most part that, indeed, literary scholars’ intuitions about long-lasting literary phenomena were right, and that the opposite would even be rather surprising. How could so many scholars have been wrong about so many things? What they might have been wrong about, though, is the Hegelian dialectical model of literary evolution that presupposes that every generational change comes as a reaction to its previous generation. What Underwood shows in his experiments is that the evolution is usually the result of a much longer arc of change.

In the last chapter, Underwood briefly addresses a key question, which he has also discussed in the public sphere: By refusing to let technology gain ground in literary studies, scholars of the field are taking for granted that literature ‘belongs’ to them – yet, who is to say that some computer scientists or biologists are not already making literature the object of their research? (*The Chronicle of Higher Education*, March 27, 2019, <https://www.chronicle.com/article/Dear-Humanists-Fear-Not-the/245987>) This points to a critical issue: In order to make a case against quantitative methods in literary analysis, one should maybe not simply refuse them from the get-go. The field needs scholars that are willing and able to argue for and against those numbers by speaking the same language.

In using an almost didactic tone, Underwood carefully lays the ground for further research in the field of literary history using distant reading and offers a monograph that should be accessible to novice and dilettante as well as more experienced distant readers. The former will be able to follow the author’s argument even without being familiar with statistical research

methods (which are explained in detail in the appendices A and B: *Data* and *Methods*), while the latter will be able to fully grasp the meaning of the data visualizations, and might be tempted to replicate or further some of the experiments themselves (the data and code that support *Distant Horizons* are available on *Zenodo*).

This is perhaps the biggest strength of Underwood's undertaking. In being the velvet glove to Moretti's iron fist, the author makes a case that could possibly convince even the most conservative literary scholars that, even if they are not yet ready to ride off together into the sunset with distant readers, this academic town *is* big enough for the both of them.

Die Quantifizierung der Literatur: Mehr als nur das Digitale

German Abstract:

Ted Underwoods *Distant Horizons: Digital Evidence and Literary Change* begibt sich in einen pazifistischen Prozess, traditionelle Literaturwissenschaft und *distant reading* in Einklang zu bringen. Um dies zu erreichen stellt Underwood die Debatte um statistische Methoden, die den Sozialwissenschaften entlehnt sind, abseits von Modewörtern und polemischen Aussagen neu in den Mittelpunkt. Dabei plädiert er für neue Entdeckungen im Bereich der Literaturwissenschaft, die - obwohl sie durch die jüngsten Fortschritte im Zugang zu digitalen Bibliotheken und Algorithmen des maschinellen Lernens erleichtert wurden - in Wirklichkeit größtenteils auf Methoden basieren, die es seit mindestens einem halben Jahrhundert gibt. Die vielleicht größte Stärke von Underwoods neuestem Werk besteht darin, dass sie selbst die härtesten Traditionalisten davon überzeugen könnte, dass quantitative Analysen ihren Platz in der Literaturwissenschaft haben sollten.