

1. Title of the Project

Negotiating Legibility: Bridging the Gap between Close and Distant Reading

2. Coördinators

Inge van de Ven (Culture Studies)

Menno van Zaanen (Communication and Information Sciences)

3. Project Summary

For years, the uses and misuses of Digital Humanities (DH) have been fiercely debated. This year, Daniel Allington, Sarah Brouillette and David Golumbia published a much-discussed piece in *LA Review of Books*, critiquing Digital Humanities advocates for their alignment with the “neoliberal takeover” of universities, provoking Alan Liu to defend DH on Twitter.¹

A recurring point of polarization in this debate is the valuation of “distant reading” versus the defense of “close reading”. Since 2000, many have followed Franco Moretti's provocative call for distant reading. Moretti deemed close reading “a theological exercise” and urged humanists to “read less”. Others, like Michael Manderino (2015) and Antoine Compagnon (2014), attempt to rehabilitate close reading, arguing that we need its associated skills and strategies more than ever in our media-saturated age.

The issue is to a considerable extent a crisis of attention. Close reading is a strategy that entails devoting attention to minute details. Some have argued that such strategies stand to be regained and revalued in a time when we encounter vast bodies of information through multiple platforms. Yet as Moretti has noted (2000, p. 57), close reading necessarily implies a select canon, or a small slice out of the available data. Acts of selection are losing currency as big data theorists today deem sampling “an artifact of a period of information scarcity, a product of the natural constraints on interacting with information in an analog era” (Mayer-Schönberger and Cukier 2013: 16-7), and as companies like Google strive to collect and organize the world's information (Vaidhyanathan 2011: 2). A close-up perspective pertains to the small; distance allows us to see the bigger picture, and the latter is currently privileged.

But is a defense of the old humanist strategy of close reading the only way out of this polemic? Are quantified, big-scale methodologies and meticulously attentive readings mutually exclusive? If not, how do we bridge the gap and unite the most valuable properties of both approaches to textuality? This study seeks to reflect on and develop analytical instruments that combine classical-humanist attention to the singular object with methods applicable to variable scales of textuality.

How do we engage with literary/media objects (either minimalist or maximalist, digital or analog) that resist this binary between close or distant reading, and demand a variation between scales? Qualitative, traditional humanities methods of textual analysis fall short of analyzing thousand- or million-page works, micronarratives, or Twitterbot poetry: such objects demand new ways of

¹ “Neoliberal Tools (and Archives): A Political History of Digital Humanities.” *LA Review of Books*, 1-05-2016; for Liu's response see <<https://storify.com/ayliu/on-digital-humanities-and-critique>>

reading. Yet quantitative DH methods like distant and algorithmic reading (See Franco Moretti, *Distant Reading*. London: Verso, 2013; Anne Burdick, Johanna Drucker and Peter Lunenfeld et al. *Digital Humanities*. Cambridge: MIT P, 2012.) often tend to undervalue specific features in favor of larger trends and patterns.

We will analyze a corpus of literary objects from the minimalist to the maximalist that solicit readings which zoom in and out between part and whole, micro and macro, surface and depth. On these basis of these, we propose reading strategies that move beyond the dichotomy and allow us to oscillate between the close and the distant, small and large-scale, minimalist and the maximalist, deep and hyper attention.

In this project, we aim to add grey scales to the originally black/white distinction between close and distant reading. Can we develop tools that allow us to identify large patterns to indicate larger trends within or between documents, while at the same time identify outliers or documents that may propose alternative views on the topics? This investigation requires both the development of computational tools that deal with large amounts of documents (e.g., topic identification, summarization), but in order to evaluate the computational analysis, in depth, qualitative analyses of the performance of the computational analyses is essential. Effectively, this combines close and distant reading, enabling close reading of “interesting” documents that either describe common trends in large document collections or provide views that counter this common trend.

4. Project timeline

Tasks

- Collection of suitable corpora (text collections) on a range of topics and a variety of scale. For example, corpora dealing with topics in which most texts deal with one view and some other texts present a different view on the topic. These may be found in the form of online forum posts (large scale, short texts) or academic articles on topics that have attracted much discussion (small scale, long texts).
- Topic analysis of texts in the corpora. This clusters documents based on the content. Techniques as Latent Dirichlet Allocation (LDA) might be useful here, but alternative techniques may need to be found if the texts discuss very similar topics. Additionally, systems that identify sentences that summarize the texts may be applied, illustrating the main viewpoint of the text, which allows for contrasts with the corpora.
- Evaluation (through in-depth analysis, close reading) of the identification of the computationally assigned topics. This shows how reliable distant reading is with respect to several tasks, such as getting the gist of the texts in the corpora, or understanding the subtle discussions and arguments made between the texts on a shared topic.

Milestones

- Collection of corpora on a range of topics that allows for the evaluation of close reading versus distant reading. These corpora will be identified by the student assistant with a cultural background (making sure a difference in close and distant reading can be found) as well as the student assistant with a computational background to make sure the collection of the texts is practically feasible.
- Annotated version of the corpora where the annotations indicate different viewpoints on a particular issue. This annotation is based on close reading to serve as a gold standard for alternative, computational approaches.

- Article on the experiences of the differences between close reading (performed by the cultural student assistant) and distant reading (performed by the computational student assistant).

5. Research Trainee Profile

To perform the research proposed in this project, both quantitative and computational analyses are required. Even though the coordinators are keen on working together on this project, they will need to cross the boundaries of their own expertise area to properly combine the techniques that are required as this project relies on the combination and interaction of computational and qualitative research. The student assistants will help in bridging these boundaries. At the same time, the students get experience in performing research (in their own field), while at the same time recognizing that alternative research methodologies (performed by the other student) exist and lead to useful results.

Because the proposed research deals with both cultural, communication sciences, and computational topics, students from a wide range of backgrounds may be interested. In particular, students with a background in the following tracks may be suitable:

- * Culture studies: Online culture
- * CIW: Human Aspects of Information Technology
- * CIW: Data Journalism
- * CIW: Text and Communication/Communication Design\

However, students with other backgrounds might be suitable candidates for the positions as long as they have a background and interest in either cultural/qualitative or computational/quantitative research.

The educational level of the student assistants (Ba, Ma, ReMa) is not particularly important. They should, however, be motivated to learn about a range of methodologies, which may be different from the methodologies they have learned about so far. Since the project relies on the interaction between the in-depth analysis of texts as well as automatic (computational) analysis of texts, we would prefer to find a student with a background in culture or text studies and another student who is knowledgeable in computational, big data techniques.

How to apply

Send a curriculum vitae as well as a brief motivation letter to
 Inge van de Ven (I.G.M.vdVen@uvt.nl)
 Ad Backus (a.m.backus@tilburguniversity.edu)
 Menno van Zaanen (mvzaanen@uvt.nl)