



HAL
open science

Von Graphen zu Word Embeddings. Zur Entwicklung des mathematischen und visuellen Instrumentariums der Literaturwissenschaft

Anne Baillot, David Lassner

► To cite this version:

Anne Baillot, David Lassner. Von Graphen zu Word Embeddings. Zur Entwicklung des mathematischen und visuellen Instrumentariums der Literaturwissenschaft. *Germanica*, 2022, La carte et la frise: les “ images ” de l’histoire littéraire, entre visualisation et modélisation, 71, pp.191-203. 10.4000/germanica.19002 . hal-03687146

HAL Id: hal-03687146

<https://hal.science/hal-03687146>

Submitted on 3 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Von Graphen zu Word Embeddings – Zur Entwicklung des mathematischen und visuellen Instrumentariums der Literaturwissenschaft

Anne Baillot, Le Mans Universität & ICAR CNRS-UMR 5191 & Maison Française d'Oxford (CNRS-UAR 3129)

David Lassner, TU Berlin & BIFOLD

Visuelle Repräsentationen von Literaturphänomenen haben lange die Institutionalisierung von Literaturwissenschaft geprägt. Literarische Bewegungen wurden in Schulbüchern in Form von Zeitleisten dargestellt, dramatische Handlungen mit der Freytagschen Pyramide abgebildet, der Interpretationsprozess durch den sogenannten hermeneutischen Zirkel repräsentiert – und dies schon seit dem 19. Jahrhundert. Bereits diese Ansätze setzten sich zum Ziel, übergeordnete Strukturen anschaulich zu machen, die sich aus dem literarischen Text ableiten lassen. Die Idee des Distant Reading (Moretti) ist damit in diesen Formen der Textrepräsentation präsent, lange bevor die Digital Humanities sie zu einem Bestandteil der Analyse von Literatur und zur Integration quantitativer Methoden in den Literaturwissenschaften machten. Lange Zeit aber diente die systematisierte Erfassung von übergreifenden Phänomenen vorrangig einer Normierung des Wissens über Literatur, gegen die sich die Digital Humanities letztlich wehren.

Normative Intention hin oder her: Fakt ist, dass die Digital Humanities, und spezifischer die Computational Literary Studies (bzw. Digital Literary Studies, Kleymann 2021), auf Vorarbeiten zurückgreifen konnten, die eine solche Integration von visuellen Repräsentationen von Text in den wissenschaftlichen Explorationsprozess durchaus erleichterten. 1967 erarbeitete der Semiologe Jacques Bertin ein universelles Alphabet der Grundeinstellungen visueller Repräsentationen. Die Anwendung in der Geographie oder der Soziologie war in einem ersten Schritt das Nächstliegende, doch die Erweiterung auf literarische Phänomene konnte sich ebenfalls in einem weiteren Schritt darauf stützen. Der anschließende Ansatz von Isabelle Meirelles (etwa 2019) ist ein gutes Beispiel für die Einbettung von Arbeiten aus dem Bereich des visuellen Designs in andere wissenschaftliche Zusammenhänge – darunter der Literaturwissenschaft – unter dem Vorzeichen der Informationswissenschaften. In pädagogisch aufgezogenen Aufsätzen gelingt es Meirelles, das Grundvokabular des visuellen Informationsdesigns für die Literaturwissenschaft lesbar zu machen, d.h. die Grundprinzipien in hermeneutische Werte ummünzbar zu machen.



Ausgehend von diesen Beiträgen der visuellen Informationskommunikation möchten wir im Folgenden zunächst einen kleinen Umriss dessen präsentieren, wie in der Literaturwissenschaft versucht wurde, sich im Zusammenhang mit den Digital Humanities ein visuelles Instrumentarium zu schaffen. Im zweiten Teil gehen wir konkreter auf die mathematische Grundlage solcher Visualisierungsformen ein, um einen Eindruck dessen zu vermitteln, welche Erkenntnisse dadurch über literarische Texte gewonnen werden können. Als Leitfaden dient uns die Frage, was mit diesen visuellen Repräsentationen gezeigt werden soll, d.h. wie sie zum hermeneutischen Prozess bei der Arbeit am literarischen Text beitragen: Tragen sie zur Normierung von literaturwissenschaftlichen Diskursen bei, oder im Gegenteil zur Entstehung neuartiger Erkenntnisse? Wie ist der Beitrag des dadurch ermöglichten Skalenwechsels ins Quantitative einzuschätzen?

1. Computer, Tools, Netzwerke: das goldene Zeitalter der Textrepräsentation?

Der 2005 von Franco Moretti veröffentlichte Band *Graphs Maps Trees* leitete eine Welle der Begeisterung für Formen der visuellen Repräsentation von Informationen aus literarischen Texten aus. Neben der Beleuchtung von unterschiedlichen Repräsentationsmustern (Graphen, Karten, Bäume) war es Morettis großes Verdienst, auf die Notwendigkeit hinzuweisen, visuelle Repräsentation nicht nur als Abbildung oder Illustration zu konzipieren, sondern diese an einer Vorarbeit am Text zu koppeln, welche die Informationsextraktion zu Visualisierungszwecken bewerkstelligen helfen würde. Dieser Zwischenschritt der Modellierung von Information ist nicht nur die Schaltstelle für den Einsatz von mathematischen Methoden in der Arbeit an literarischen Texten: Er leitet auch im Selbstverständnis der Literaturwissenschaft eine epistemologische Wende ein in dem Sinne, dass der hermeneutische Prozess von nun an nicht nur aus einer Spirale des sich vertiefenden Textverständnisses zusammensetzt, sondern explizit eine Phase der Strukturierung von Information mit einbettet. „Worin besteht die Information, die es zu interpretieren gilt?“ und „Wie sammle ich diese Information?“ werden von dann an legitime Fragen der digital agierenden Literaturwissenschaft – auch eine Herausforderung für die explizit nicht digitale.

Morettis drei Visualisierungsmethoden (Graphen, Karten, Bäume) ebneten zwar den Weg zu einer Vielfalt von Visualisierungsformen, dabei die Frage nach der Formalisierung von Informationen aufwerfend, doch es pendelten sich Anwendungsfälle für diese unterschiedlichen Methoden ein, die die Kreativität in deren Umsetzung einschränkte.

Diese Entwicklung lässt sich besonders gut am Beispiel von Netzwerken nachvollziehen. Weil sie darauf angelegt sind, Beziehungen gewichtet zu repräsentieren, wurden Netzwerkvisualisierungen bevorzugt eingesetzt, um



Personenkonstellationen abzubilden. Das macht Moretti selbst z.B. im Zusammenhang mit Dramenfiguren. Doch wer über einfache Illustrationen hinausgehen will (etwa mit der Struktur, bei der ein Pfeil=eine Beziehung), braucht Rechenmodelle, die zum Einen die Größe der Punkte (nodes) und zum Anderen die Breite der Kanten (edges) nach ihrer respektiven Wichtigkeit bestimmen (Baillot 2018). Die große Beliebtheit von Netzwerkvisualisierungen insbesondere im Kontext von historischen Forschungen, die verstärkt mit Informationen aus dem Bereich der automatischen Eigennamenerkennung arbeiten, führte zur Entwicklung einer Reihe von unterschiedlich leicht handhabbaren Tools, um diese darzustellen: Vistorian, Palladio, Nodegoat, Gephi.¹ Gephi ist ein relativ komplexes Instrument, das auch 3D-Visualisierungen generiert – jedoch ohne, dass dem Nutzer/der Nutzerin immer klar ist, nach welchen Regeln diese Netzwerkvisualisierungen generiert werden. Neuere Ansätze sind um Transparenz in den verwendeten Einstellungen bemüht, so etwa die Plattform DRACOR², die Theaterproduktionen gewidmet ist. Dort kann man beispielsweise für jedes Schiller-Stück die Figurenkonstellationen in einem Personennetzwerk repräsentiert finden, ebenso die Verwandtschaftsbeziehungen zwischen den Figuren und die Verteilung des Textes je nach Szene – die eingesetzten Einstellungen sind immer am Rande der Webseite mit einsehbar.³

Was in der Darstellung großer Informationsmenge an Klarheit gewonnen werden kann, wird aber oft an der Undurchsichtigkeit der Berechnungsmethodik verloren. Als Geisteswissenschaftler*in sieht man zwar etwas, verfügt aber nicht unbedingt über das Wissen, das es möglich machen würde, das Gesehene treffend zu interpretieren. Während sich Netzwerkvisualisierungen im Zusammenhang mit (realen oder fiktionalen) Personen als eine Form des literaturwissenschaftlichen Erkenntnisgewinns etablierten (Baillot 2017), wurde speziell für die Literaturwissenschaft ein Instrumentarium zur Verfügung gestellt, das die Vielfalt der Bedürfnisse adressieren wollte, welche literaturwissenschaftliche Fragestellungen mit sich bringen. Die Voyant-Tools⁴ stellen eine Breite Palette an Visualisierungsformen für die Exploration literarischer Texte zur Verfügung. Sie bieten u.a. auch Wordclouds, d.h. Schlagwortwolken, eine der einfachsten Formen der Repräsentation von Worten und ihrer Verteilung in Texten. In einer Wortwolke werden die vorkommenden Worte je nach Gewichtung im Ausgangskorpus unterschiedlich groß dargestellt: Je öfter ein Wort vorkommt, desto größer erscheint es in der Wortwolke – so dass durch die visuelle Repräsentation ein Eindruck der relativen Wichtigkeit der vorkommenden Worte vermittelt wird. Allerdings reicht es nicht aus, einfach den Text einzuspeisen und „Wortwolke generieren“ anzuklicken. Füllwörter müssen herausgefiltert werden, unterschiedliche Formen desselben Wortes (Singular/Plural, unterschiedliche Konjugationen eines Verbs, deklinierte Formen eines Nomens,..) zusammengeführt.

¹ Vgl. <https://vistorian.net/>, <https://hdlab.stanford.edu/palladio/>, <https://nodegoat.net/>, <https://gephi.org/>.

² Vgl. <https://dracor.org/>.

³ Am Beispiel der *Räuber* hier einsehbar: <https://dracor.org/ger/schiller-die-raeuber#network> (durch Anklicken von „Relations“ bzw. „Speech distribution“ Zugriff auf weitere Informationen).

⁴ Vgl. <https://voyant-tools.org/>.



Die zahlreichen Kritiken an Wortwolken (Felix et al. 2018; Hearst et al. 2019; Boullier & Crépel 2009) haben dazu geführt, dass diese kaum noch als Methode des wissenschaftlichen Erkenntnisgewinns betrachtet werden, und vielmehr als visueller Aufhänger eingesetzt werden: ein gutes Beispiel für die Gratwanderung zwischen Erkenntnisgewinn und schönem Schmuck, die den Einsatz von visuellen Textrepräsentation in der literaturwissenschaftlichen Forschung oft noch charakterisiert.

In ihrem 2013 Artikel „Information visualization for Humanities scholars“ gehen die Entwickler der Voyant Tools auf die Geschichte der dynamischen Repräsentationsmöglichkeiten ein, die, angelehnt an die Entwicklung von Browsern im Laufe der Zeit, ihrem eigenen Darstellungsangebot vorausgingen. Es wird auf die notwendige Vorarbeit eingegangen, die geleistet werden soll, ehe von literarischen Texten Visualisierungen bewerkstelligt werden, denn, wenn diese es ermöglichen können, „den Wald“, der „vor lauter Bäumen“ (§ 47) nicht zu sehen war, vor Augen geführt zu bekommen, ist das Risiko aber auch groß, da einen Wald zu sehen, wo nur Bäume stehen, d.h. ein allgemeingültiges Muster erkennen zu wollen, wo de facto eine Idiosynkrasie (§ 49) einschlägig ist. Die Vielfalt der Visualisierungsformen – auch die in Voyant angebotenen – öffnet nicht nur die Möglichkeit eines an diesen angelehnten Interpretationsansatzes, sondern macht auch die Notwendigkeit klar, im Schritt vor der Generierung von Visualisierungen eine weitere Arbeitsetappe zu konsolidieren, an der Schnittstelle zwischen Korpuskonstitution und Informationsmodellierung. Je nachdem, welche Texte in Voyant (oder in einem anderweitigen Visualisierungsgenerierungssystem) eingespeist werden, und in welcher Form sie es werden, wird etwas anderes abgebildet.

An der Notwendigkeit, sich darüber Gedanken zu machen, welche Informationen repräsentiert werden sollen, setzt CATMA-fortext⁵ ein. CATMA hat zum Ziel, es Literaturwissenschaftler*innen bei der Annotation von Texten leichter zu machen. Eine solche Annotation verleiht den Informationen Struktur, die im Nachgang visualisiert werden können. Damit wird eine didaktische Anleitung an Formen des Skalenwechsels in der Literaturwissenschaft angeboten, d.h. eine Herangehensweise an die Zwischenschritte, die Literaturwissenschaftler*innen einlegen müssen, damit die für sie relevanten Texte durch die Maschine so prozessierbar sind, dass die Forscher*innen am Ende mit der Repräsentation des von Ihnen Annotierten in der Interpretation des Textes weiterkommen.

In diesem Fall wie in vielen anderen jedoch bleiben die gängigen Visualisierungsoptionen in den Literaturwissenschaften weit hinter dem, was dem Korpus gerecht werden würde. Zum Einen bleiben in der Regel die dahinterliegenden Rechenmethoden unklar (was ja daran liegt, dass sie für Literaturwissenschaftler*innen nicht unmittelbar verständlich sind und in den einschlägigen Arbeitsumgebungen selten prominent bzw. didaktisch genug aufgeführt werden). Zum Anderen würde ihre Komplexität für Repräsentationsformen plädieren,

⁵ Vgl. <https://fortext.net/tools/tools/catma>.



die weit über das hinausgehen, was das menschliche Auge in den von Bertin und Meirelles definierten Vokabularien leicht erfassen kann. Diese Vokabularien wurden neuerdings erweitert z.B. durch Kleymann und Stange 2021. Diese führen nun Visualisierungen für die Digital Humanities auf vier Postulate zurück, die in Anlehnung an die Arbeiten von Johanna Drucker definiert wurden.

In ihren zahlreichen Arbeiten zur Frage nach der Visualisierung von Daten geht Johanna Drucker unter anderem auf zwei Punkte ein, die im Zusammenhang mit literaturwissenschaftlichen Fragestellungen von Interesse sind und in einem 2016 veröffentlichten Video-Vortrag besonders anschaulich gemacht werden. Dort heißt es zunächst, dass in der Visualisierung Argument und Beleg kollabieren, was das Auseinandernehmen beider Elemente schwieriger macht. Offensichtlich dient die Visualisierung in diesem Fall der Analyse, nicht der Exploration (und inwiefern sie dann zum hermeneutischen Erkenntnisgewinn beiträgt, ist nicht ganz klar). Eine der Fragen, die sich Literaturwissenschaftler*innen stellen sollten, ehe sie Visualisierungen einsetzen, ist also, ob diese als analytischer Baustein dienen sollen oder als anregende Illustration, die womöglich das Auge in eine neue Richtung lenkt. Der zweite von Drucker erwähnte, hier relevante Punkt weist auf eine weitere Beschränkung hin, die es ebenfalls in Betracht zu ziehen gilt. In ihrem Video bemüht Drucker unterschiedliche Arten und Weisen, denselben Datensatz zu visualisieren. Einmal wird das Bild um 90 Grad gedreht ("If you can turn it on the side, you can read it"), einmal wird eine dritte Dimension hinzugefügt, mit dem Ergebnis, dass plötzlich Punkte sich voneinander abheben und sichtbar werden, die in konzentrierteren Repräsentationen zu dicht beieinander standen, als dass ein Muster erkannt werden könnte. Diese Form der Umvisualisierung wirft die Frage nach der Willkür der gewählten Visualisierungsformate und ihres Einflusses auf die daraus abzuleitende Interpretation auf. Drucker erkundet in ihrem Vortrag die Grenze, die zwischen hochdimensionalen mathematischen Repräsentationen und visuellen Repräsentationen verläuft: Sicher kann noch eine dritte Dimension hinzugefügt werden, aber dann ist im Grunde für das menschliche Auge Schluss, während mathematische Modelle in viel mehr Dimensionen arbeiten als nur drei. Das, was berechnet werden kann, übersteigt in vielem das, was das Auge visuell erfassen kann. Wozu dann die schönen Visualisierungen, wenn sie das Berechnete nicht wiedergeben können?

Damit wird auch hinterfragt, worin die hermeneutische Leistung in der Literaturwissenschaft besteht und welche Medialitätsformen dabei ihren Platz haben. Die Kompetenz, mit Graphen zu arbeiten, wird in literaturwissenschaftlichen Curricula nicht beigebracht. Auch intertextuelle Interpretationen beruhen in der Regel auf einer linearen Zusammenstellung von Textpassagen, ja insgesamt ist immer noch Linearität die vorherrschende Form des Lesens, auch wenn digitale Formate es ermöglichen würden, diagonale Lektürewege zu entwickeln bzw. weitere Kombinationen von Textualität heranzuziehen. Nicht nur quantitativ, sondern auch qualitativ bieten digitale Methoden Wege, das Lesen und Interpretieren literarischer Texte zu erneuern. Aber inwiefern brauchen wir dabei mathematische Verfahren?



2. Textrepräsentationen des maschinellen Lernens: ein Zwischenschritt für Visualisierungen als Interpretationshilfe der Literaturwissenschaft

Im vorangegangenen Abschnitt wurde schon darauf eingegangen, dass es hochdimensionale mathematische Repräsentationen gibt, die als Zwischenschritt erzeugt werden, um dann in niedrigerer Dimensionalität visualisiert zu werden. Dieses Verfahren wird bspw. bei Ramsay 2011 diskutiert, wenn er relative Worthäufigkeiten bestimmt, dann allerdings nur eine kleine Auswahl tabellarisch aufführt. Wichtiges Kriterium bei Ramsay ist das Algorithmische ('Toward an Algorithmic Criticism'), also eine Abfolge von Befehlen, die bspw. aus einem Textkorpus eine hochdimensionale Repräsentation für die enthaltenen Dokumente erstellt, dann paarweise Ähnlichkeiten berechnet und diese dann visualisiert. Dafür werden z.B. Bag-of-Words verwendet, also eine Repräsentation, die für jedes Wort im Gesamtkorpus (Vokabular) zählt, wie oft es in jedem Dokument enthalten ist. Die Repräsentation eines Dokuments ist dann also ein Vektor, dessen Anzahl an Einträgen der Länge des Vokabulars entspricht. Durch die Berechnung der paarweisen Ähnlichkeiten der Dokumente fallen die vielen Dimensionen zusammen und lassen sich visualisieren. Ein Beispiel für eine Anwendung in den Computational Literary Studies die der gerade Beschriebenen sehr ähnlich ist, wäre Stilometrie mithilfe von z.B. Burrows' Delta (Burrows 2002). Die Stilometrie versucht anhand der Distanzen zwischen diesen Bag-of-Words Vektoren Texte unterschiedlicher Autor*innen zu unterscheiden (Jannidis et al. 2014 für eine Anwendung im Bereich der deutschen Literatur). Um aber beim Beispiel von Schillers *Räuber* zu bleiben, könnten die Personen im Stück stilometrisch verglichen werden, indem die Reden jeder Person zusammengefasst werden und jeweils eine Bag-of-Words erstellt werden könnte. Die Distanzen zwischen den Vektoren der verschiedenen Personen können in einem Dendrogramm visualisiert werden. Die Visualisierung kann dann von einem Menschen interpretiert werden und somit ist es ein 'Criticism' des Menschen mithilfe eines Algorithmus, denn die konzeptionelle Arbeitsetappe (die Wahl der Art der Repräsentation, die der Fragestellung gerecht wird) hat vorher stattgefunden. Offensichtlich ist nicht jede Repräsentation bzw. Visualisierung für jedes Korpus und jede Fragestellung passend, und die Entscheidung darüber kann auch nicht von einem Algorithmus getroffen werden. So könnte man in unserem hypothetischen Fall der *Räuber* von Unterschieden in den Reden der verschiedenen Personen ausgehen, die mithilfe von Stilometrie gefunden werden könnten.

Wir behaupten nicht, dass sich, durch die Anwendung von Methoden des maschinellen Lernens, plötzlich Literaturwissenschaftler*innen durch statistische



Modelle ersetzen ließen, dennoch verschieben sich die klaren Grenzen zwischen der regelbasierten Lesemaschine und dem urteilenden Menschen ein wenig, wenn statistische Repräsentationen verwendet werden.

Anstelle einer Bag-of-Words Repräsentation ließen sich auch Word Embeddings einsetzen. Mit Word Embeddings (Mikolov et al. 2013, Pennington et al. 2014) sind Repräsentationen gemeint, bei denen jedem Wort im Korpus ein Vektor zugeordnet ist. Die Repräsentationen der Worte im Vektorraum sind hierbei so angeordnet, dass Worte, die sich syntaktisch oder semantisch ähnlich sind, auch im repräsentierten Raum nah beieinander liegen. Diese Word Embeddings lassen sich beispielsweise auf einem anderen, als dem zu analysierenden Korpus trainieren (wir könnten hier das zu analysierende Korpus *Subjektkorpus* nennen und das Korpus, auf dem die Word Embeddings trainiert wurden als *Referenzkorpus* bezeichnen) und enthalten damit bestimmte syntaktische und semantische Informationen des Referenzkorpus. Word Embeddings werden dafür verwendet Varianten von Named Entities zu identifizieren (Rusinek et al. 2021). In Schillers *Räuber* könnten beispielsweise auch Worte mit bestimmten syntaktischen Eigenschaften identifiziert werden, etwa Imperativformen. Eine Visualisierung der Ergebnisse basierend auf diesen Repräsentationen wären dann nicht mehr rein regelbasiert, sie wären auch nicht nur basierend auf dem Subjektkorpus, sondern eine selektive Aggregation der dem Referenzkorpus extrahierten Informationen. Kategorisierend könnte man von einer Entwicklung von 'regelbasiert' (Algorithmus), über 'nur basierend auf dem Subjektkorpus' (klassische Statistik) hin zu 'basierend auf Referenzkorpus' (Maschinelles Lernen) sprechen. Man könnte diese Interaktion zwischen großem Referenzkorpus und zu untersuchendem Subjektkorpus in Analogie zu der Interaktion zwischen Leser*innen eines Textes (Subjektkorpus) vor dem Hintergrund ihres gesamten Lesehorizontes (Referenzkorpus) sehen.

Die konzeptionelle Arbeitsetappe wird bei diesem Ansatz weiterhin benötigt, nur müssen andere Überlegungen angestellt werden, z.B. 'Wie konstruiere ich mein Referenzkorpus?' 'Welche Eigenschaften sollten in der Repräsentation enthalten sein, - sollte ich bspw. nur Lemmata des Referenzkorpus verwenden?'. Es bleibt also wie bisher ein Großteil der Methoden- und Parameterauswahl in den Händen der Forschenden. Außerdem besteht bis heute trotz intensiver Forschung der Machine Learning Community das Problem der Repräsentation von längeren Textsequenzen. Auch wenn die Embeddings schon lange nicht mehr wie Word Embeddings auf einzelne Worte beschränkt sind, können viele Modelle (sogenannte Language Models, Devlin et al. 2019), die die verbreitete Transformerarchitektur (Waswani et al 2017) verwenden, Sequenzen von maximal ein paar hundert Worten Länge verarbeiten. Mit Modell ist hier ein Machine Learning Modell gemeint, also eine Instanz einer bestimmten Architektur, die mithilfe eines Trainingsalgorithmus und eines Datensatzes 'trainiert' werden kann, um auf einem anderen Datensatz Vorhersagen



zu treffen. Das Modell beinhaltet die Architektur sowie die zu trainierenden bzw. schon trainierten Parameter. Datensatz sowie Trainingsalgorithmus sind üblicherweise nicht Teil des Modells.

In dem oben vorgeschlagenen 'Lesehorizont' steckt für viele Anwendungen allerdings auch eine Gefahr: Wenn das Referenzkorpus schlecht dokumentiert ist bzw. so groß ist, dass es gar nicht sauber dokumentiert werden kann, ist es ebenso schlecht auszumachen, welche Biases von dem trainierten Modell übernommen wurden. Diese Kritik wurde von Bender et al. 2021 auf den Punkt gebracht. Gefahren der Fortsetzung bzw. Verstärkung dieser Biases bestehen insbesondere bei Nutzer*innen moderner Natural Language Processing(NLP)-Anwendungen, bei denen oft gar nicht öffentlich ist, worauf diese trainiert wurden. Forschende sollten sich diesen Gefahren in ihren Arbeiten bewusst sein. Es gibt allerdings auch Gelegenheiten zum Erkenntnisgewinn: In Zusammenhang mit einfachen Word Embeddings wurde schon gezeigt, dass die Analyse von Biases in Teilen der verwendeten Referenzkorpora wichtige Erkenntnisse über gesellschaftliche Stereotype bestimmter Zeiten beleuchten können (Garg et al. 2018).

In diese Richtung argumentiert auch Ted Underwood in seinem kürzlich verfassten Beitrag (Underwood 2021), indem er formuliert, dass "The immediate value of these models [d.h. Language Models] is often not to mimic individual language understanding, but to represent specific cultural practices (like styles or expository templates) so they can be studied and creatively remixed." Und Underwood fährt fort, dass "[w]hen research is organized by this sort of comparative purpose, the biases in data are not usually a reason to refrain from modeling – but a reason to create more corpora and train models that reflect a wider range of biases." Die Idee wäre also beispielsweise ein generelles Language Model auf einem großen Korpus der heutigen Zeit zu trainieren (pre training) und es dann auf andere historische Kontexte zu fokussieren (fine tuning). Problematisch kann daran sein, dass damit nicht gut dokumentiert bleibt, wieviel aus dem Korpus der heutigen Zeit mittransportiert wird, das dann nicht dem historischen Anwendungskontext entspricht. Gleichzeitig gibt es wohl heute nur wenige historische Kontexte, für die genügend große Textkorpora existieren, die ein Training ohne pre training erlauben.

Ein möglicher Mittelweg wäre hier, Word Embeddings auf einzelnen historischen Kontexten (sub-corpora) zu trainieren, die aber während des Trainings in Beziehung zueinander gesetzt werden. So kann eine weniger datenhungrige Methode verwendet werden, die keine undokumentierbaren großen Trainingskorpora braucht, die hingegen trotzdem ermöglicht, die Repräsentation der verschiedenen Subkorpora zu vergleichen (Lassner and Brandl et al., forthcoming).



Schluss

Die Tiefe der mathematischen Kenntnisse, die zur Beherrschung des visuellen Instrumentariums notwendig ist, um neuartige literaturwissenschaftliche Erkenntnisse zu produzieren, ist mit der Ausbildung als Literaturwissenschaftler*in kaum vereinbar. Wer kann schon sowohl Machine Learning Methoden vergleichen, auswählen und anwenden, als auch die generierten Visualisierungen einordnen, analysieren, und in literaturwissenschaftliche Interpretation umwandeln? Heutzutage sicherlich nur eine handvoll Personen. Dennoch sind die hier dargelegten Methoden und Probleme nicht abschreckend gemeint. Vielmehr ging es uns darum, auf die Komplementarität unterschiedlicher Forschungsfelder und -methoden hinzuweisen, die herangezogen werden müssen, um aus dem visuellen Repertoire inhaltlich Wertvolles schöpfen zu können, und die Literaturgeschichtsschreibung zu erneuern.

Festzuhalten bleibt, dass selbst die komplexesten Machine Learning-Methoden ohne menschliches Zutun ('algorithmic criticism') wenig ertragreich sind, und dass die Qualität der Forschung auf eine Mischung von close und distant reading ankommt. Es geht nicht darum, die Maschine von sich aus lesen zu lassen.

Die vielleicht wichtigste Verschiebung, die den verbreiteten Einsatz von Visualisierungsmethoden mit sich bringen, ist ein besseres Verständnis einerseits für Fragen der Textrepräsentation, andererseits für Methoden der Modellierung von Information. Im digitalen Kontext können zahlreiche Repräsentationen einer Textinstanz vorkommen: der Scan einer Buchseite, ihr Inhalt im Fließtextformat, eine annotierte Fassung in XML-TEI, ein Netzwerk, das Beziehungen abbildet, eine Tabelle, die Berechnungen zusammenfasst... all das sind Textrepräsentationen. Damit wird das Primat der Printseite ein Stück infrage gestellt. Desweiteren haben wir hier Modellierung auf zwei Ebenen geschildert: Modellierung von Informationen im Text (beispielsweise die Kennzeichnung von Vorkommnissen von Eigennamen), und Modellierung der Art und Weise, wie diese Informationen behandelt werden (welche Vorkommnisse werden als besonders wichtig betrachtet, und wie das berechnet wird). Denn eins ist sicher: Für die quantitative Literaturwissenschaft ist Informationsmodellierung unabdingbar. Das heißt nicht, dass wir Germanistik-Studierende zu Informatiker*innen ausbilden müssen. Das heißt aber wohl, dass ihnen die Prozessualität der Schritte im Laufe des Erkenntnisgewinns viel bewusster sein müsste, als es ein klassisches literaturwissenschaftliches Curriculum ermöglicht. Aber damit werden neue Horizonte eröffnet. Die Mühe lohnt sich allemal.

Literaturverzeichnis

BAILLOT, Anne, "Die Krux mit dem Netz Verknüpfung und Visualisierung bei digitalen Briefeditionen", in: *Quantitative Ansätze in den Literatur- und Geisteswissenschaften. Systematische und historische Perspektiven*, hg. v. Andrea Albrecht, Toni Bernhard, Marcle Lepper, Sandra Richter, Marcus Willand, De Gruyter, 2018, S. 355-370.



- BAILLOT Anne, "Das Netzwerk als Kunstwerk", in: *Die Kommunikations-, Wissens-, und Handlungsräume der Henriette Herz (1764-1847)*, hg. v. Hannah Lotte Lund, Ulrike Schneider, Ulrike Wels, Vandenhoeck & Ruprecht, 2017, S. 45-53.
- BENDER, Emily M. und Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜 In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). Association for Computing Machinery, New York, NY, USA, 610–623. <https://doi.org/10.1145/3442188.3445922>
- BERTIN, Jacques, *Sémiologie graphique. Les diagrammes, les réseaux, les cartes*, 1967 (englische Übersetzung 1983, Madison, Univ. of Wisconsin Press)
- BOULLIER Dominique und Maxime Crepel, « La raison du nuage de tags : format graphique pour le régime de l'exploration ? », in: *Communication & langages*, Vol. 160, 2009, issn : 0336-1500, doi : 10.4074/S0336150009002099, S. 119
- BURROWS, John. 'Delta': a measure of stylistic difference and a guide to likely authorship. *Literary and linguistic computing* 17.3 (2002): 267-287.
- CATMA: <http://catma.de> (link is external); Gius, Evelyn, Jan Christoph Meister, Malte Meister, Marco Petris, Christian Bruck, Janina Jacke, Mareike Schumacher, Dominik Gerstorfer, Marie Flüh und Jan Horstmann (2020): CATMA 6.2 (Version 6.2). Zenodo. DOI: [10.5281/zenodo.1470118](https://doi.org/10.5281/zenodo.1470118) (link is external).
- DEVLIN, Jacob, Ming-Wei CHANG, Kenton LEE, Kristina TOUTANOVA. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. NAACL. 2019. <https://aclanthology.org/N19-1423>
- DRUCKER, Johanna (2016) "What do digital humanists want from visualization?" <https://www.youtube.com/watch?v=kdey5H2NI9w>
- DRUCKER, Johanna (2018). "Non-representational approaches to modeling interpretation in a graphical environment", in *Digital Scholarship in the Humanities*, 33(2), pp. 248-263.
- FELIX, Cristian, FRANCONERI, Steven L., BERTINI, Enrico, « Taking Word Clouds Apart : An Empirical Investigation of the Design Space for Keyword Summaries », in *IEEE Transactions on Visualization and Computer Graphics*, Vol. 24, 2018, doi : 10.1109/TVCG.2017.2746018
- JÄNICKE, Stefan, Greta FRANZINI, Muhammad F. CHEEMA und Gerik SCHEUERMANN, 2015. On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. In: Rita BORGIO, Fabio GANOVELLI und Ivan VIOLA (Hrsg.), *Eurographics Conference on Visualization (EuroVis)*. Zürich: The Eurographics Association. Mai 2015. Verfügbar unter: <https://www.informatik.uni-leipzig.de/~stjaenicke/Survey.pdf>.
- JANNIDIS, Fotis und Gerhard Lauer: Burrows Delta and its Use in German Literary History. In: Matt Erlin, Lynne Tatlock (Hg.): *Distant Readings. Topologies of German Culture in the Long Nineteenth Century*. Rochester: Camden House 2014, 29-54.



- KOOLEN, Marijn, Jasmijn VAN GORP und Jacco VAN OSSENBRUGGEN, 2018. Toward a model for digital tool criticism: Reflection as integrative practice. *Digital Scholarship in the Humanities*. 2018. DOI [10.1093/llc/fqy048](https://doi.org/10.1093/llc/fqy048).
- GARG, Nikhil, et al. Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, 2018, 115. Jg., Nr. 16, S. E3635-E3644.
- HEARST Marti et al., « An Evaluation of Semantically Grouped Word Cloud Designs », in: *IEEE Transactions on Visualization and Computer Graphics*, 2019, doi : 10.1109/TVCG.2019.2904683
- HORSTMANN, Jan und Jan-Erik STANGE, 2018. Textvisualisierung. *forTEXT. Literatur digital erforschen*. 2018. Verfügbar unter: <https://fortext.net/routinen/methoden/textvisualisierung>.
- KLEYMANN, Rabea & STANGE, Jan-Erik “Towards Hermeneutic Visualization in Digital Literary Studies”
<http://www.digitalhumanities.org/dhq/vol/15/2/000547/000547.html> (§25, §35
- MEIRELLES, Isabelle (2019). “Visualizing information”, in Flanders, J., Jannidis, F. (eds.) *The Shape of Data in the Digital Humanities. Modeling Texts and Text-based Resources*. London, New York: Routledge, pp. S. 167-177 -> distinction between descriptive visualization (direct, unambiguous) and argumentative visualization (more speculative)
- MIKOLOV, Tomas, Ilya SUTSKEVER, Kai CHEN, Greg S CORRADO, and Jeff DEAN. *Distributed representations of words and phrases and their compositionality*. *Advances in neural information processing systems*, pages 3111–3119. 2013.
- MORETTI, Franco (2005). *Graphs, Maps, Trees. Abstract Models for a Literary History*. London/New York.
- MORETTI, Franco 2017: <https://litlab.stanford.edu/LiteraryLabPamphlet15.pdf> (Figuren+teil 4)
- PENNINGTON, Jeffrey, Richard SOCHER, and Christopher D MANNING. *Glove: Global vectors for word representation*. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.
- RAMSAY, S. (2011). *Reading Machines: Toward an Algorithmic Criticism*. University of Illinois Press.
- ROCKWELL, Geoffrey, “What is Text Analysis, Really?”, *Literary and Linguistic Computing*, Vol. 18, No. 2, 2003, p.209-219.
<https://geoffreyrockwell.com/publications/WhatIsTAnalysis.pdf>
- RUSINEK, Sinai und Nitzan Gado. 2021. Feeding a Gazetteer: Leveraging word embeddings for toponym mining. In *Proceedings of the 5th ACM SIGSPATIAL International Workshop on Geospatial Humanities (GeoHumanities'21)*. Association for Computing Machinery, New York, NY, USA, 28–35.
<https://doi.org/10.1145/3486187.3490202>
- SINCLAIR, S., Ruecker, S., Radzikowska, M. (2013). “Information Visualization for Humanities Scholars”, in Price, K. M., Siemens, R. (eds.). *Literary studies in the digital age: An evolving anthology*. New York: Modern Language Association.



Available at: DOI: 10.1632/llda.2013.6

(<https://dlsanthology.mla.hcommons.org/information-visualization-for-humanities-scholars/>)

UNDERWOOD, Ted. (2021) "Mapping the Latent Spaces of Culture", Preprint, may subsequently appear in Startwords. doi.org/10.17613/faaa-1r21

VASWANI, Ashish, Noam SHAZEER, Niki PARMAR, Jakob USZKOREIT, Llion JONES, Aidan N. GOMEZ, Łukasz KAISER, Illia POLOSUKHIN. *Attention is All you Need*. NeurIPS. 2017.