Die Quadratur des hermeneutischen Kreises

Zu den (Un)Möglichkeiten digitaler Ansätze[1]

Evelyn Gius

Einleitung

Vielen Dank für die Einladung, hier zu sprechen. Ich habe mit Freude zugesagt, weil es für mich eine Ehre ist, hier zu sein, und ich danke Lina Franken und Sabina Mollenhauer für das Vertrauen, dass ich hier etwas sagen darf, was hoffentlich nicht nur mir etwas bringt, sondern auch Sie als Zuhörende und Lesende voranbringt. Ich bin keine Empirische Kulturwissenschaftlerin und auch keine Expertin der Ethnografie oder verwandter Bereiche – ich bin Literaturwissenschaftlerin mit einem Schwerpunkt in den Computational Literary Studies, einem Teilbereich der Digital Humanities. Ich werde im Folgenden versuchen, etwas zur Frage der Messbarkeit unserer Gegenstände zu sagen. Dabei werde ich – und ich hoffe, Sie sehen mir das nach – vor allem aus der Perspektive der Literaturwissenschaft sprechen. Nicht, um über Literaturwissenschaft an sich zu sprechen, sondern weil ich mich damit am besten auskenne und von dort aus Sie fragen möchte: Wie ist es denn bei Ihnen? Haben Sie ähnliche Probleme? Ich vermute, an einigen Stellen schon, an anderen wahrscheinlich weniger.

Ich glaube auch – das muss ich vorwegsagen –, dass die Literaturwissenschaft es ein bisschen einfacher hat als die Empirische Kulturwissenschaft, was das Digitale angeht. Denn wir sind weniger automatisch mit gesellschaftlichen und allgemeinen Fragen verstrickt, wenn wir digital forschen. Das heißt, viele ethische Fragen, die im Rahmen der Tagung diskutiert wurden, sind in der Literaturwissenschaft keine Vorbedingung, um arbeiten zu können.

Mit meinem Beitrag hoffe ich etwas zu dem zu sagen, was Katharina Kinder-Kurlanda so treffend unter Critical Proximity diskutiert (Kinder-Kurlanda in diesem Band). Was kann man tun, um ins digitale Forschen zu kommen und von dort aus vielleicht auch Kritik zu üben? Den Teil mit dem Kritiküben werde ich hier auf meine Heimatdisziplin, die Literaturwissenschaft, beschränken.

1 Bei diesem Text handelt es sich um ein für die Schriftfassung leicht gekürztes und für bessere Lesbarkeit überarbeitetes Transkript des mündlichen Vortrags.

Mein Beitrag ist wie folgt gegliedert: Ich werde erst drei Vorüberlegungen diskutieren, die ich mir mit dem Titel selbst eingebrockt habe. Das sind die Quadratur des Kreises, die Hermeneutik und schließlich das Messen – also das, was hinter dem Titel steckt. Danach werde ich versuchen herauszuarbeiten, wo diesbezüglich bei uns – also in der Literaturwissenschaft – Probleme bestehen, wenn man sich fragt: Darf man Literatur messen? Dabei wird es sowohl um die Literaturwissenschaft generell als auch um mein eigenes Feld gehen, die Computational Literary Studies, wobei ich einen Teil meines Arguments im Teilbereich der Computational Narratology entwickeln werde. Danach will ich zwei Lösungsansätze vorstellen. Der erste betrifft die Frage, was man tun kann, um die Komplexität, die solchen Ansätzen natürlich innewohnt, ein Stück weit handhabbar zu machen. Als zweiten Lösungsansatz möchte ich vorschlagen, dass wir etwas anders über Forschungsprozesse nachdenken – auch, um ein Interface zwischen qualitativen und quantitativen Methoden zu etablieren.

Vorüberlegungen

Die Quadratur des Kreises

Beginnen wir also mit der Quadratur des Kreises. Die Quadratur des Kreises ist ein klassisches Problem der Geometrie. Die Aufgabe besteht darin, aus einem gegebenen Kreis mit Zirkel und unbeschriftetem Lineal in endlich vielen Schritten ein Quadrat mit demselben Flächeninhalt zu konstruieren. Die Herausforderung dabei – so kann ich Ihnen das als mathematische Laiin zumindest erklären – liegt darin, dass wir es bei der Zahl π (Pi), die wir, wie Sie wissen, für die Flächenberechnung benötigen, mit einer transzendenten Zahl zu tun haben. Das heißt: Es handelt sich um eine Zahl, die man nicht mit natürlichen Zahlen abbilden kann. Man kann sie also nicht durch endlich viele elementare algebraische Operationen mit Ganzzahlen – also durch Dividieren, Multiplizieren, Quadrieren und so weiter – zur Null machen. Das ist aber wichtig, wenn man etwas machen will, was man mit den uns irgendwie verständlichen Operationen messen kann. π können wir uns nämlich nicht vorstellen, weil es unendlich viele Nachkommastellen hat und wir dafür in unserer realen Welt keine Entsprechungen haben, die das abbilden. Zur Quadratur des Kreises gab es im Laufe der Zeit viele Versuche. Franco von Lüttich zum Beispiel hat im 11. Jahrhundert 22/7 als Ersatz für π genutzt und einen Kreisdurchmesser von 14 angenommen. Nachdem er erkannt hat, dass man die Kreisfläche als halbes Rechteck darstellen kann, dessen eine Seite der Umfang (also 22) und dessen andere Seite der Radius ist (also 7), hat er den Kreis mit Zirkel und Lineal in 44 Segmente geteilt. Davon hat er jeweils 22 abwechselnd mit den Spitzen nach oben und unten aufeinander und die beiden Flächen nebeneinander gelegt. So konnte er aus den 44 Kreisseg-

menten ein Rechteck mit den Seitenlängen 11 und 14 konstruieren. Damit wäre die Quadratur des Kreises potenziell gelöst. Das Problem ist allerdings, dass die Quadratwurzel aus 22/7, die Franco als Annäherung an π und außerdem vermeintlich konstruierbare Zahl genutzt hat, um diese Fläche zu berechnen, schon wieder eine transzendente Zahl ist. Das Problem wurde also doch nicht gelöst. Seit dem Ende des 19. Jahrhunderts wissen wir auch, dass es eine unlösbare Aufgabe ist, zumindest wurde von Ferdinand von Lindemann bewiesen, dass π eine transzendente Zahl und somit nicht konstruierbar ist, was Franco von Lüttich 800 Jahre vorher nicht wusste.

Soweit zur Mathematik in den Problemen. Die Frage für uns ist, haben wir so ein Problem, haben wir ein transzendente-Zahl-Problem in unseren Fächern, das sich nicht anhand von uns begreifbarer Mittel konstruieren lässt? Und wenn ja, wie sehen die Lösungen dann aus? Oder haben wir vielleicht auch umgekehrt Lösungen für Probleme, die keine transzendente-Zahl-Probleme sind, und erzeugen trotzdem transzendente Zahlen, also unnötig komplizierte Lösungen die wir nicht anhand der Grundoperationen unseres Fachs erfassen können? Letzteres ist als mögliches Problem auch nicht von der Hand zu weisen.

Hermeneutik

Kommen wir von der Quadratur des Kreises zum hermeneutischen Kreis oder, wie er inzwischen meist bezeichnet wird, zum hermeneutischen Zirkel. Das ist ein in der Literaturwissenschaft weit verbreitetes Modell, das darstellt, wie das Textverstehen funktioniert. Es gibt verschiedene Ansätze und Termini dafür, aber im Wesentlichen geht es um ein wiederholtes Hin- und Her-Schwenken zwischen dem eigenen Vorverständnis und dem Verständnis des Textes, durch welches wir schließlich zu einem Gesamtverständnis kommen können. Im Reallexikon der deutschen Literaturwissenschaft werden drei Verständnisse von Hermeneutik genannt. Diese möchte ich nicht im Detail erörtern, sondern nur kurz zitieren. Ich kann Ihnen schon vorweg sagen: Das erste ist das Verständnis, auf das ich mich hier beziehe: „Unter Hermeneutik versteht man derzeit, der dreierlei:" – schön ist auch, dass Klaus Weimar hier „derzeit" schreibt – „(1) Die Theorie des Lesens, Verstehens und der Interpretation von Texten; (2) eine philosophische Richtung, die einerseits den Textbegriff universalisiert, und andererseits (eben deshalb) das Verstehen als Teil des zu Verstehenden konzipiert (vgl. Hermeneutik$_2$)". Das ist dann ein von Gadamer und anderen geprägter Hermeneutikbegriff. Ich denke, es gibt auch in der Empirischen Kulturwissenschaft einiges, was unter dieses Verständnis fällt. Schließlich ist Hermeneutik laut Weimar „(3) eine Art der Interpretation, die in scheinbarer oder wirklicher Übereinstimmung mit der Hermeneutik$_2$ die Wiedergewinnung von vorgesehenem Sinn zum Ziel hat, im Gegensatz etwa zur Dekonstruktion", und so weiter, dann: „Die Verwendung von Hermeneutik in Bedeutung drei ist mindestens problematisch" (Weimar 2000). Was ich hier meine,

ist Hermeneutik im ersten genannten Verständnis, also Verstehensprozesse im weitesten Sinne.

Nun stellt sich die Frage: Wie funktioniert das in der Literaturwissenschaft, was sind die Arbeitsschritte in der hermeneutisch-literaturwissenschaftlichen Textanalyse? Das mag nach einer banalen Frage klingen, aber es ist gar nicht so klar. Eine der annähernd expliziten Beschreibungen, die allerdings nicht spezifisch hermeneutisch, sondern genereller ist, haben Tillmann Köppe und Simone Winko (2013) vorgelegt. Demnach geht es bei der literaturwissenschaftlichen Textanalyse grob gesagt um drei Schritte: Erstens um die lesende Aufnahme des Textes, man rezipiert also den Gegenstand. Zweitens erfolgt die begriffliche Fixierung von Textphänomenen, das heißt, man schaut auf Textphänomene und auf Textstrukturen – das wird oft auch als Analyse im engeren Sinne bezeichnet. Drittens geht es darum, die Textbedeutung nachvollziehen, also das, was im Reallexikon unter Hermeneutik in der ersten Bedeutung firmiert. Das heißt, dass im dritten Schritt auf Basis der vorhergehenden Lektüre und Analyse die umfassende Textbedeutung, die Wirkungsgeschichte des Textes, die Autor:innenintention oder anderes nachvollzogen werden. Die Möglichkeiten in der Literaturwissenschaft sind hier vielfältig, da wir einen sogenannten Methodenpluralismus haben – der allerdings ein anderer ist, als der, den die Empirischen Kulturwissenschaft hat. Bei uns bedeutet Methodenpluralismus, dass die Methoden relativ unabhängig voneinander existieren und angewendet werden. Bei Ihnen ist es, glaube ich, hingegen so, dass Sie mehrere Methoden haben, um Vergleichbares zu erreichen, und diese in einem Zusammenspiel angewendet werden.

Aber zurück zur literaturwissenschaftlichen Textanalyse. Man kann sie zwar annäherungsweise beschreiben, wie Sie eben gesehen haben, aber im Grunde handelt es sich eher um eine Blackbox. Das ist gar nicht als Provokation gemeint. In meiner Wahrnehmung haben wir in unseren Disziplinen etablierte Verfahren, die trotzdem Blackboxes sind. So wie die Hermeneutik, die ein Verfahren ist, das durchaus beschrieben, aber im Detail trotzdem oft unklar ist. Das Ziel ist jedenfalls, Texte zu verstehen.

Messen

Damit komme ich zur dritten und letzten Vorüberlegung: Messen. Was ist eigentlich Messen? Ich finde es sehr interessant, dass viele – ich auch – beim Messen erst einmal an physikalische Dinge denken. Zum Beispiel denken wir ans Temperaturmessen, ein klassisches Messverfahren, das ein Thermometer nutzt. Dieses hängen wir irgendwo hin und können dann die Temperatur als Gradzahl ablesen. Tatsächlich sind Messverfahren ursprünglich in den Naturwissenschaften entwickelt und eingesetzt worden. Das erscheint uns vielleicht auch naheliegend, schließlich kann man, so glauben wir zumindest, physikalische Größen gut erfassen. Der Witz ist

nun aber, dass viele physikalische Größen gar nicht direkt gemessen werden können. Das heißt, dass man Messinstrumente entwickeln muss, die andere Größen messen, die mit dem, was man eigentlich messen will, systematisch zusammenhängen. Diese sogenannten Näherungsgrößen muss man erst einmal identifizieren. Hier müsste es bei uns allen klingeln, denn das Problem mit den Näherungsgrößen kennen wir, weil wir an unsere Gegenstände oft nicht direkt herankommen. Interessant ist, dass das in der Physik zum Teil gar nicht anders ist. Wenn Sie zum Beispiel eine Waage nehmen, dann misst diese die Masse eines Körpers nicht direkt, sondern die Waage misst die Gewichtskraft des Körpers im Schwerefeld der Erde. Wenn wir uns auf dem Mond auf eine Waage stellen, ändert sich deshalb zwar deren Anzeige, aber es ändert sich natürlich nicht unser Gewicht. Wir wiegen immer noch gleich viel, aber das, was auf der Waage steht, ändert sich. Das ist ein Zeichen dafür, dass die Waage nicht das Gewicht misst oder zumindest als Messinstrument auf Faktoren reagiert, die die Vergleichbarkeit des Gewichts auf dem Mond und der Erde stören. Bei der Temperatur ist es ähnlich. Da kommt noch die historische Anekdote hinzu, dass man früher gedacht hat, Temperatur könne man nicht messen, weil Temperatur eine subjektive Empfindung sei. Aus heutiger Sicht ist das schwer vorstellbar, weil Thermometer in unserer Wahrnehmung mit zu den paradigmatischen Messinstrumenten gehören.

Anhand der Erfindung des Thermometers kann man übrigens auch gut erklären, wie ein wissenschaftlicher Prozess zur Etablierung einer Messprozedur ablaufen kann. Grob gesagt, hat jemand festgestellt, dass Körper – aus physikalisch näher zu beschreibenden Gründen – Wärme oder auch Kälte abstrahlen und dass die Wärme bzw. Kälte wiederum Quecksilber-Verhalten beeinflusst, denn es dehnt sich bei Wärme aus bzw. zieht sich bei Kälte zusammen. Das sieht man noch besser, wenn man Quecksilber in einem sehr schmalen Röhrchen hat. Eine weitere Eigenschaft von Quecksilber, die anders ist als bei vielen anderen Flüssigkeiten, besteht darin, dass es sich gleichmäßig ausdehnt und zusammenzieht. Man kann deshalb an diesem Röhrchen ablesen, wie warm oder kalt es ist. Dafür nutzt man die Striche auf dem Thermometer, wobei auch die nicht sozusagen naturgegeben sind. Man kann nicht einfach irgendwelche Striche aufzeichnen, sondern muss eine Skala entwickeln, indem man das Messinstrument nutzt und eine Kalibrierung entwickelt, die schließlich eine Skala informieren kann. Wie Hasok Chang so schön herausgearbeitet hat, bildeten am Anfang der Geschichte der Temperaturmessung Indikatoren wie der erste Frost oder eine Kerzenflamme die Orientierungspunkte der Temperaturskala (Chang 2004). Irgendwann ist man dann auf Gefrier- und Siedepunkt von Wasser unter bestimmten Bedingungen gekommen, inzwischen sind es komplexere, aber auch stabilere thermodynamische Bedingungen, anhand derer Temperatur im Internationalen Einheitensystem (SI) festgelegt ist. Für uns ist hierbei vor allem interessant, wie ein Messverfahren entwickelt werden kann. Bei der Temperatur

sind die verschiedenen Thermometer die Messinstrumente, die so entwickelt wurden.

Nun haben wir es gerade in unserer Forschung oft mit der Situation zu tun, dass wir keine eigenen Instrumente oder Messverfahren haben. Und es ist sehr wichtig, auf diese Weise über unsere Forschung zu denken, also sich zu fragen: Weiß ich überhaupt, wie ich messen soll? Und was messe ich denn da eigentlich? Das sind, glaube ich, wichtige Fragen, die wir öfter im Fokus haben sollten. Dann ist es natürlich auch noch so, dass ich, wenn ich ein Instrument habe, damit natürlich auch sehr Verschiedenes messen kann. Es ist nicht das Instrument, das uns sagt, was ich damit tun kann, und auch die Bedeutung des Messergebnisses ist nicht direkt gegeben. Zum Beispiel gibt es wesentliche Unterschiede in der Einschätzung von Gradzahlen, je nachdem, ob man die Temperatur eines Babys, eines Hochofens oder eines Kühlschranks misst. Da möchte man gerne verschiedene Temperaturen jeweils erreichen bzw. vermeiden.

Probleme des Messens in der Literaturwissenschaft

Darf man Literatur messen? Von Äpfeln und Birnen (und Bananen)

Soweit also zu den Vorüberlegungen. Ich komme jetzt zum Hauptproblem, für das ich Vorschläge entwickeln möchte: Darf man Literatur messen? Die Computational Literary Studies machen das bereits. Das ist an der Darstellung der Ergebnisse besonders offensichtlich, die sich insofern von typischen literaturwissenschaftlichen Publikationen unterscheiden, dass sie zumeist in Zahlen und quantifizierenden Abbildungen erfolgen. So werden geordnete Listen von nach Distinktivität – also nach besonderer Relevanz – gewichteten Wörtern genutzt, um Aussagen über Unterschiede zwischen zwei Textkorpora zu treffen, etwa um die Detektivromane von Doyle von seinen anderen Texten zu unterscheiden. Oder die Abbildung von Texten als Blätter in einem Baumdiagramm, einem sogenannten Dendogramm, erfolgt anhand von über Wortfrequenzen gemessener Textähnlichkeit, und ermöglicht es Autorschaftsfragen für Text mit unklarer oder strittiger Autorschaft zu betrachten (vgl. zum Beispiel die Abbildungen in Schöch 2017).

Nun halten nicht alle Forschenden solche Verfahren für geeignet. Für Kritiker:innen der computationellen Verfahren ist der Weg von solchen Ansätzen zu dem, was der Professor in der Akademie von Lagado in Jonathan Swifts „Gullivers Reisen" macht, höchstens ein kurzer. In der Akademie von Lagado passiert folgendes: Gulliver wird von einem Professor herumgeführt und dieser teilt seine Erkenntnisse und seine „Forschungsstrategie" mit ihm: „Ein jeder wisse, wieviel Mühe die gewöhnliche Erlernung der Künste und Wissenschaften bei den Menschen erfordere; er [der Professor] sei überzeugt, durch seine Erfindung werde

die ungebildetste Person bei mäßigen Kosten und bei geringer körperlicher Anstrengung Bücher über Philosophie, Poesie, Mathematik und Theologie ohne die geringste Hilfe des Genies oder von Studien schreiben können" (Swift o. J.: 199). Das klingt interessant, aber wie funktioniert das? Die beiden gehen zu einem Gerät aus einem hölzernen Gestell mit einem Gitter aus drehbaren Rahmen, an denen Tafeln mit Wörtern und Silben angebracht sind. Dann heißt es: „Jeder Zögling nahm auf seinen Befehl einen eisernen Griff zur Hand, von denen vierzig am Rande befestigt waren. Durch eine plötzliche Wendung wurde die ganze Anordnung verändert. Dann befahl er sechzehn Knaben, die verschiedenen Zeilen langsam zu lesen, und wenn sie drei oder vier Worte herausgefunden hatten, die einen Satz bilden konnten, diktierten sie diese vier anderen Knaben, welche sie niederschrieben. Diese Arbeit wurde drei- oder viermal wiederholt" (ebd.: 200).

Swift nimmt mit dieser Szene die Mechanisierung von Erkenntnis aufs Korn. Es handelt sich um Satire und wenn man das Gerät zu Ende denkt, dann sieht man auch, dass man so natürlich nicht schreiben kann, weil die potenziell möglichen Kombinationen an Wörtern zu Texten so zahlreich sind, dass man sie nicht alle innerhalb eines Menschenlebens prüfen kann. Trotzdem ist es ungefähr ein solches mechanistisches, unterkomplexes Vorgehen, das manche Kritiker:innen automatischen Verfahren unterstellen. Das ist nicht nur oft falsch, etwa wenn es um Vorwürfe gegen Verfahren der Computational Literary Studies geht, die nicht per se unterkomplex sind, wie ich später noch versuchen werde zu zeigen. Das ist auch gefährlich, da neuere Entwicklungen wie etwa Large Language Models (LLMs) damit im Zweifelsfall unterschätzt werden, anstatt dass man ihnen fundierte Kritik entgegenbringt.

Neben dem Mechanismus-Vorwurf, der die Digital Humanities generell betrifft, gibt es auch noch einen anderen beliebten, ich möchte fast sagen, Topos: Digital Humanities als strukturalistisches und ergo schlechtes Unterfangen. Was ist der Vorwurf? Zum Beispiel schreibt James Dobson über die Digital Humanities und vor allem über die computationell arbeitenden Literaturwissenschaften: „Yet all these different approaches can be understood to be part of a retrograde movement, that nostalgically seeks to return literary criticism to the structuralist era. To a moment characterized by belief in systems and structure and the transparency of language" (Dobson 2019: 57). Das heißt, auch hier geht es im Endeffekt um den Vorwurf eines unterkomplexen Verfahrens, das aus reduzierenden strukturalistischen Zugängen entsteht. Wenn man sich die Kritiken anderer anschaut – und Grundsatzdebatten zur Frage des Neoliberalismus oder anderer, aus meiner Sicht nicht direkt die Forschung betreffender Themen ausschließt, – sind es eine Reihe vor allem methodologischer Probleme, die den Computational Literary Studies vorgeworfen werden. Es herrsche eine Beliebigkeit der Auswahl der Methoden, es gäbe Schwierigkeiten bei der Begründung der ausgewählten Muster und Befunde –

sowohl dabei, was man untersucht, als auch dabei, was die Analyseergebnisse sind – und überhaupt sei das Verhältnis zwischen Daten und Interpretation schwierig.

Ich würde das anders darstellen. Die Probleme bestehen zwar, aber erstens sind das keine spezifisch strukturalistischen Probleme. Es sind, wenn schon, Probleme jeder computationellen Textanalyse. Und zweitens ist auch das Computationelle zweitrangig, denn es sind vor allem Probleme der Textanalyse. Das heißt, das sind Probleme, die wir in der Literaturwissenschaft bei der Textanalyse immer haben. Mit einem weiten Datenbegriff, wenn Literatur als Daten gesehen wird, dann haben Sie diese Probleme auch, wenn Sie Literatur selbst lesen und interpretieren. Das habe ich versucht ein Stück weit herauszustellen, als ich vorhin über die Hermeneutik gesprochen habe und sagte, dass es auch dort einen gewissen Explizierungsbedarf gibt. Es ist nicht so klar, wie man Methoden, Daten oder Literatur auswählt, wie man eben zwischen Text und Interpretationen hin- und herpendelt und wie man Befunde dann narrativiert, um eine sinnhafte Gesamterzählung bzw. Interpretation zu erzeugen. Bei Problemen des Verfahrens sind die computationellen und die nicht computationellen Zugänge in meiner Sicht gar nicht so weit auseinander.

Angenommen – und das nehme ich jetzt einfach an, Sie können ja nicht so einfach widersprechen – Sie folgen meiner Argumentation in Teilen, sehen also die Vorwürfe als wenig stichhaltig. Dann geht es weniger um die Unterschiede zwischen computationeller und nicht-computationeller Literaturwissenschaft oder um die Frage der Unterkomplexität als vielmehr um die Frage der Passung der computationellen Methoden zum Gegenstand. Man könnte es auch so sagen: Vergleichen wir hier nicht Äpfel mit Birnen, wenn wir die in den Naturwissenschaften etablierten Messverfahren in den Sozial- und Geisteswissenschaften nutzen?

Darauf gibt es viele mögliche Antworten. Man könnte zum Beispiel entgegnen, dass es sehr wohl Bereiche unserer Disziplinen gibt, in denen Messverfahren angewendet werden, etwa in der Psychologie oder auch in der Soziologie. Ich beschränke mich hier darauf, eine mögliche Ursache für diese Skepsis aus der literaturwissenschaftlichen Perspektive zu beleuchten. Meine Argumentation wird damit enden, dass ich vorschlage, dass Äpfel und Birnen nicht die Frage sind, sondern es um Bananen bzw. Bananenhaftigkeit geht.

Was hinter dieser Mess-Skepsis stehen könnte, ist der vermeintliche Gegensatz zwischen Erklären und Verstehen. Diese Idee geht zurück auf Dilthey, der gesagt hat, dass Naturwissenschaften erklären und Geisteswissenschaften verstehen wollen. Es gibt also einen grundlegenden Unterschied in den Zielen und dieser ist nach der Auffassung einiger – wobei ich glaube, dass Dilthey selber gar nicht dazu gehört – ein unüberwindbarer Graben. Es gibt aber auch andere Haltungen dazu, die in meinen Augen plausibler sind. Im bereits erwähnten Beitrag schreiben Köppe und Winko: „Die traditionell bedeutsame Unterscheidung zwischen Erklären und Verstehen ist im Rahmen interpretationstheoretischer Überlegungen kaum aussagekräftig. Jede Interpretation ist in der einen oder anderen Weise damit befasst,

bestimmte Textbefunde zu erklären, und sie zielt darauf, ein besseres Verständnis des Werkes zu befördern" (Köppe und Winko 2013: 285). Das heißt, sowohl Erklären also auch Verstehen ist das, was wir in der Literaturwissenschaft – und ich vermute: in den meisten Wissenschaften – anvisieren.

Aber nochmal anders gefragt: Worum geht es denn in der Literaturwissenschaft? Hier folgt eine wichtige Beobachtung, die für alle Wissenschaften gilt. Nämlich, dass nicht die Gegenstände das wissenschaftliche Feld konturieren, sondern ihre Erfassung oder eben ihre Phänomene. Remigius Bunia hat das schön für die Literaturwissenschaft dargestellt:

> „[d]as, was ich unter Phänomen verstehe, darf nicht mit dem Gegenstand der Untersuchung verwechselt werden. Gegenstände der literaturwissenschaftlichen Untersuchung lassen sich nennen: Diese sind beispielsweise Romane, Gedichte, Text schlechthin, das Narrative an und für sich (…), Kunst und das Ästhetische. Leider sind jedoch Gegenstände der Untersuchung keineswegs etwas, das die Disziplinen zu konturieren hülfe."

Und jetzt kommt endlich die angekündigte Banane:

> „Untersucht eine Biologin als Biologin Bananen, verfolgt sie weder die Marktpreise über die letzten Jahrzehnte noch fragt sie, zu welchen Reisgerichten Bananen passen könnten. Was an einer biologischen Untersuchung der Banane biologisch ist, definiert die Biologie, nicht die Banane." (Bunia 2011: 151)

Das klingt ein bisschen lustig, aber der wichtige Punkt ist: Was ist eigentlich aus disziplinärer Sicht die Bananenhaftigkeit eines Gegenstandes, also welche seiner Eigenschaften sind relevant? Das bestimmt, wenn wir Bunia folgen, nicht der Gegenstand, das bestimmt die ihn betrachtende Disziplin. Das heißt, was ist zum Beispiel die Bananenhaftigkeit des Gegenstandes in den Computational Literary Studies, in denen literaturwissenschaftliche Methoden computationell umgesetzt werden? Wie wird Bananenhaftigkeit gemessen? Darum wird es gleich gehen. Davor noch kurz zur Bedeutung der Analyse in der Literaturwissenschaft. Es gibt zwei Kandidaten für Bananenhaftigkeit, die wir betrachten können. Die werden unterschiedlich benannt, ich zitiere wieder Bunia, der folgende Benennung und Unterscheidung vorschlägt: „Die Fragen zielen auf zwei Phänomenkomplexe ab: einesteils auf die ‚Webart' eines Textes, andernteils auf seine Bedeutung. Ersterer Schwerpunkt lässt sich als Rhetorikanalyse, letzterer als Textdeutung bezeichnen" (ebd.: 152). Die Rhetorikanalyse beschäftigt sich mit der möglichen, generell beschreibbaren Form von Texten, während die Textdeutung einzelne Texte im Blick hat. Entsprechend ist Ersteres sehr generisch und auf „ein weites, prinzipiell unbe-

grenztes Korpus an Texten" (ebd.: 153) ausgerichtet und Zweiteres eben auf ein sehr kleines Korpus bzw. Einzeltexte.

Bei den generischen Betrachtungen der Rhetorikanalyse geht es entsprechend vor allem darum, Verfahren zu entwickeln, um Texte zu analysieren. Einzeltexte dienen nur als Beispiele. Das ist auch das, was im Zentrum der Narratologie steht, weshalb die wiederum sehr gut für die Computational Literary Studies geeignet ist, die ihre Stärke auch vor allem in der Methodenentwicklung und der Analyse von vielen Texten haben. Bei Textdeutung steht wiederum die Interpretation des Textes im Zentrum, des Einzeltextes als Einheit, könnte man auch sagen. Das heißt, es geht um die Spezifität dieses einen Textes oder dieses einen kleinen Korpus. Das ist eine Interpretation im traditionellen Sinne.

Bunia verweist darauf, dass beide Schwerpunkte gleichermaßen wichtig sind. Für das Literatur-Messen-Problem ist allerdings der generische Fokus der produktivere. Denn es geht nicht darum, und jetzt sage ich zum letzten Mal Banane, die einzelne Banane mit bestehenden Kriterien zu analysieren, sondern es geht eher darum zu sagen, was ist eigentlich das Interessante an den Bananen?

Computational Literary Studies

Ich bin bis jetzt eine Bestimmung der Computational Literary Studies schuldig geblieben und will das nun nachholen. Zuerst in Blick auf das Selbstverständnis des Felds. Schauen wir auf die Homepage des DFG Schwerpunktprogramms Computational Literary Studies, das seit 2019 läuft und an dem ich beteiligt bin (Jannidis 2021). Schwerpunktprogramme bei der DFG sind dafür da, sogenannte „emerging fields" zu fördern. Ziel dieses spezifischen Programms ist es, die Computational Literary Studies zu fördern, wobei als Ziel das „developing and establishing innovative computational methods in the field of literary studies" (ebd.) angegeben wird. Es geht also um die Verwendung von formalen Modellen in der Literaturwissenschaft, worunter im Zweifelsfall quantifizierende, auf jeden Fall aber computationelle Modelle verstanden werden. Das bedeutet auch, dass es darum geht, quantitativ-empirische Forschung mit qualitativ-hermeneutischer Forschung zu verbinden. Was ich hier außerdem hervorheben möchte, ist, dass in den Computational Literary Studies durchaus ein Bewusstsein dafür besteht, dass es auch problematische, zu reflektierende Aspekte gibt. Das sieht man etwa im Selbstverständnis des Journal of Computational Literary Studies, das 2022 im Kontext des Schwerpunktprogramms gegründet wurde und dessen Mitherausgeberin ich bin. Da steht unter anderem: „JCLS also acknowledges the debatability of the core concepts of Computational Literary Studies" (Journal of Computational Literary Studies), und das gilt sowohl in Bezug auf Computationalität als auch auf Literarizität, also sowohl Methoden als auch Gegenstände sind potenziell verhandelbar bzw. müssen immer wieder diskutiert werden.

Kommen wir nun zum konkreten Vorgehen der Computational Literary Studies. Ich werde es anhand von drei Übersichten vorstellen. Die erste ist eine Übersicht über Handbücher, und zwar sechs an der Zahl, die im Kontext des EU-Projektes „Computational Literary Studies Infrastructure" entstanden sind, das international durchgeführt wird (CLS Infra). Die Beteiligten haben sechs kleine Einführungen in die Computational Literary Studies geschrieben, die als digital verfügbare Handbücher konzipiert sind (Schöch et al. 2023). Worum geht es da? Es gibt ein generelles Handbuch und dann Einführungen in fünf spezifische Gegenstandsbereiche, die als relevant angesehen werden: Autorschaft, Genre, Literaturgeschichte, Gender und Kanonizität. Die Handbücher haben alle den gleichen Aufbau: Es gibt einen einleitenden Teil, der inhaltliche Fokus wird vorgestellt, die geeignete Textauswahl sowie die Frage nach zusätzlich benötigten Informationen für die Analyse, also Metadaten, wird diskutiert, die konkreten Analyse-Methoden werden vorgestellt und schließlich die Evaluation dieser Methode. Hier sieht man schon, dass der Forschungsprozess etwas anders organisiert ist als in der nicht-computationellen Literaturwissenschaft. Wobei aus meiner Sicht der wesentliche Unterschied weniger im Computationellen, als vielmehr in der Tatsache liegt, dass die Methode immer evaluiert wird.

Einen weiteren, etwas anderen Blick auf die Methoden der Computational Literary Studies ermöglicht eine Übersichtsstudie, die über 200 im Jahr 2022 in vier größeren einschlägigen Publikationsorten erschienene Beiträge untersucht hat, die man als Computational Literary Studies bezeichnen könnte (Hatzel et al. 2023). Von diesen Beiträgen basieren 40 auf computationellen Analysen und wenn man sich die dort jeweils genutzten Methoden anschaut, sieht man sozusagen einen Ritt durch die Methoden der Natural Language Processing Community der letzten Jahrzehnte. Das reicht von alten Methoden, wie Entscheidungsbäumen, die seit den 1960er Jahren für Klassifikationsaufgaben genutzt werden, bis zu sehr aktuellen Methoden wie Transformer-Modellen, die die Grundlage für LLMs bilden. Diese Bandbreite zeigt, dass in den Computational Literary Studies alles Mögliche genutzt wird. Daran sieht man, würde ich sagen, dass der Innovationsdruck bezüglich der Methoden in den Computational Literary Studies deutlich geringer ist als in informatischen Fächern und damit, so denke ich, auch ein Stückweit die Passung von Methoden zu Gegenständen und Fragestellungen mehr im Blick ist.

Die dritte Übersicht zu den Computational Literary Studies kommt aus dem bereits erwähnten Schwerpunktprogramm. Für diese wurden alle elf Projekte der ersten Phase zu ihrer Methodennutzung befragt (Helling et al. 2022). Dabei ging es, anders als in der eben besprochenen Studie, nicht um die genutzten Algorithmen, sondern generell um Verfahren. Interessant ist, dass trotzdem nur zwei der erwähnten Methoden qualitative Verfahren sind. Es wird zum einen eine qualitative Analyse ohne weitere Spezifizierung aufgeführt und das andere ist die Analyse von manuellen Annotationen.

Soweit ein kurzer Einblick in das, was die Computational Literary Studies betreiben. Man kann sagen, dass zahlreiche computationelle Methoden mit einigen – wenigen – qualitativen Ansätzen gemischt werden.

Was heißt das für das Forschungsfeld? Das ist nicht ganz klar. Ich möchte zwei nicht ganz zusammenpassende Bestimmungen von Computational Literary Studies geben, um das zu verdeutlichen. Der Witz ist, dass ich an beiden beteiligt bin, man kann hier also nicht von einer Kontroverse sprechen, sondern, wenn überhaupt, eher von einer inneren Widersprüchlichkeit. Zuerst noch einmal zum Schwerpunktprogramm. In dem Antrag haben wir geschrieben: „CLS are firmly grounded in literary studies". Das heißt, die Computational Literary Studies gehören zu den literaturwissenschaftlichen Disziplinen und sind Literaturwissenschaft. Jetzt die zweite Bestimmung, die aus dem erwähnten Übersichtsbeitrag zu computationellen Methoden in den Computational Literary Studies stammt. Es gibt dort einen Abschnitt mit der Überschrift „CLS as applied NLP", also Natural Language Processing. Dort schreiben wir „The computational text analysis in Computational Literary Studies can be considered an application domain of natural language processing. Traditionally NLP in conjunction with adjacent fields like computational linguistics has been focused on analyzing all aspects of text," und weiter unten steht dann, „that is to say, CLS can offer a test bed for challenging tasks of natural language processing by providing the tools needed to facilitate literary analysis" (Hatzel et al. 2023: 2, 3). Das heißt, man kann Computational Literary Studies auch als Testumgebung für NLP sehen, also eine Testumgebung, um etwas innerhalb des NLP zu entwickeln.

Ich denke, beide Bestimmungen sind berechtigt, da es gute Gründe für beide gibt: Computational Literary Studies kann sowohl der Literaturwissenschaft als auch dem NLP zugeordnet werden. Trotzdem ist die Tatsache, dass beides behauptet wird, ein Indikator dafür, dass da eventuell doch ein Problem besteht.

Die methodologische Lücke

Wenn wir uns meinen Teilbereich der Computational Literary Studies, die Computational Narratology, genauer anschauen, dann wird das Problem leichter fassbar, weil wir einen kleineren Bereich in den Blick nehmen können. Ich habe letztes Jahr eine Keynote auf der europäischen Narratologie-Tagung gehalten, bei der das Tagungsthema Zeit und Raum war, und habe dafür eine kleine Recherche gemacht: Wie ist es eigentlich mit Zeit- und Raumanalysen in der Narratologie, und wie ist es mit narratologischen Zeit- und Raumanalysen in der NLP oder in der weiteren computationellen Community, die eben nicht literaturwissenschaftlich ausgerichtet ist? Dabei stellte ich fest, dass es sehr wohl so etwas wie Narratology in Computing gibt. Auch wenn man wie ich nur nach Zeit und Raum sucht, finden sich einige Publi-

kationen im Bereich der klassischen NLP, die Narratologie oder Narrative Theory explizit nennen oder sich auf narratologische Forschung berufen.

Das passt dann auch zu dem, was Inderjeet Mani, der Autor des Handbuchbeitrags zu „Computational Narratology" im „Living Handbook of Narratology" (Hühn et al. 2013), in seiner Definition gleich am Anfang schreibt: „Computational Narratology is a study of narrative from the point of view of computation and information processing" (Mani 2013). Hier geht es darum, dass aus der Informatik heraus Narratologie betrieben wird, wobei ich ergänzen würde: Dabei wird die Narratologie oft nur oberflächlich behandelt. Das würde ich deshalb eher unter Narratology in Computation fassen. Aber es gibt auch Beiträge, die stärker narratologisch ausgerichtet sind, also eine narratologische Frage computationell verfolgen. Diese Perspektive wird ebenfalls im „Computational Narratology"-Handbuchbeitrag erwähnt, auch wenn sie in der Definition selbst nicht auftaucht. Mani erklärt außerdem, dass es eine Bedeutung von Computational Narratology im Sinne eines „methodological instrument in the construction of narratological theories" (ebd.) gibt. Das heißt, hier gibt es dann die Rückbindung an die Theorie bzw. diese ist sogar zentral. Das Problem ist aber das, was auch Andrew Piper, Richard Jean So und David Bamman herausstellen, nachdem sie eine Übersicht über narratologisch informierte Beiträge zur Analyse von Erzähltexten aus der maschinellen Sprachverarbeitung erstellt haben: „this work (…) is by and large divorced from the large body of theoretical work on narrative within the humanities, social and cognitive sciences" (Piper et al. 2021: 298).

Das heißt, es gibt eine Lücke zwischen Theorie und computationellen Analysen. Und wie kann man diese Lücke beschreiben? Das mache ich spezifisch für die Narratologie, aber ich glaube, dass es in der computationellen Literaturwissenschaft ein Stück weit ähnlich ist. Wenn man davon ausgeht, dass man in der Forschung von einer Theorie, die die geeigneten konzeptuellen Grundlagen bereitstellt, über ein Modell, das die Konzepte für die Bestimmung der Phänomene nutzbar macht, zur eigentlichen Textanalyse kommt, dann kann man Unterschiedliches beobachten. In der Computational Narratology ist es oft so, dass Forscher:innen in der Theorie etwas modellieren, lange darüber nachdenken, auch manuell annotieren, aber dann nur wenig als Analyse umsetzen, also den Weg nicht ganz bis zur eigentlichen Textanalyse gehen. Die NLP-Forscher:innen machen hingegen eher Folgendes: Sie gehen von einer narratologischen Theorie aus, wählen dann eine z. T. nicht weiter begründete NLP-Methode und machen dann die Analyse. In diesem Fall geht es also von der Theorie zur Analyse, aber über einen Umweg, oder wohl eher eine Abkürzung, die das narratologische Modell ignoriert und stattdessen direkt auf ein NLP-Verfahren zugreift.

Das heißt, das Modell ist auf der einen Seite für die Theorie vielleicht eine Art Grenze, weil es doch eine Herausforderung ist, die Theorie so in ein Modell zu überführen, dass man damit dann umfangreiche Analysen machen kann. Und auf der

anderen Seite klappt das mit der Analyse, aber die Modellierung narratologischer Phänomene wird umgangen.

Was wir hier eigentlich bräuchten, ist das: Wir bräuchten einen Weg von der Theorie, der nach dem, was wir Operationalisierung nennen, direkt übergeht in die Auswahl oder vielleicht auch die Etablierung eines geeigneten NLP-Modells, um die Analyse umzusetzen. In Einzelfällen gelingt das zwar jetzt schon, aber grundsätzlich ist das eher nicht so. Ich denke, dass es zwei Strategien gibt, die das erleichtern könnten. Die erste ist die Steuerung von Komplexität eines Forschungsvorhabens, die eine mögliche Reduktion von Komplexität und vor allem eine bewusste Setzung von Komplexität bedeutet. Die zweite Strategie ist, die eigene Forschung als Forschungsprozess zu denken, der sich an Messverfahren orientiert.

Lösungsvorschläge

Steuerung von Komplexität

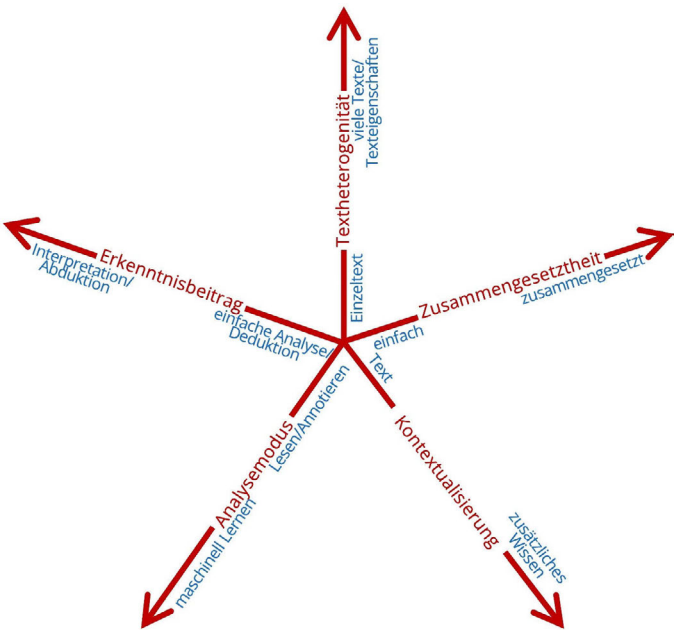
Ich möchte ein Modell vorstellen, das ich zu den Dimensionen von Komplexität der computationellen literaturwissenschaftlichen Textanalyse entwickelt habe, und von dem ich hoffe, dass es vielleicht auch darüber hinaus für computationelle Textanalysen anwendbar ist (Gius 2019). Es sind fünf Dimensionen: (1) die Zusammengesetztheit der analysierten Phänomene, (2) die Kontextualisierung der Phänomene, (3) die Heterogenität der betrachteten Texte, (4) der Analysemodus und (5) der Erkenntnisbeitrag der computationellen Analyse (vgl. Abb. 1).

Die ersten beiden Dimensionen betreffen die untersuchten Phänomene: Erstens, werden diese als einfach oder zusammengesetzt konzipiert? Zweitens, wie stark sind sie kontextualisiert? Dann geht es um die Textgrundlage und da ist die Frage: Drittens, wie heterogen ist die Textbasis, die ich analysieren möchte? Die Komplexität reicht vom einfachen Einzeltext bis hin zu vielen Texten mit verschiedenen Merkmalen. Die letzten beiden Dimensionen betreffen die Erkenntnisgenerierung: Viertens, wie setze ich die Analyse um, durch menschliches Lesen oder durch avancierte maschinelle Verfahren? Fünftens, wie umfangreich ist der computationell erzeugte Erkenntnisbeitrag bzw. ist das ein deduktives, induktives oder abduktives Vorgehen? Das ist eine recht kompakte Darstellung, aber ich hoffe, dass sie für unsere Zwecke reicht.

Der Witz an diesem Modell ist nicht, dass man damit die Qualität eines Vorhabens im emphatischen Sinne messen kann. Vielmehr ist es ein Modell, um ein Vorhaben zu planen. Alles, was man damit festlegt, sind Setzungen. Es geht um die Frage: Wie operationalisiere, wie modelliere ich etwas? Das heißt zum Beispiel Folgendes, wenn ich mir die beiden Dimensionen der Zusammengesetztheit und der Kontextualisierung der analysierten Phänomene anschaue. Nehmen wir an, ich

möchte Geschlechterrollen analysieren und dafür auf das Konzept von Geschlechterrollen von Simone de Beauvoir zurückgreifen. Wenn ich das mache, muss ich natürlich über das Konzept Bescheid wissen, mir überlegen, welche Textmerkmale von Figuren ich den Rollen zuweise, und muss auch wissen, wie ich spezifische Texte in der jeweiligen historischen Zeit interpretieren kann, um die Bezüge auf Geschlechterrollen zu erkennen. Das heißt, wenn ich mir so etwas wie Geschlechterrollen anschaue und das relativ komplex operationalisiere, dann bin ich eine Weile damit beschäftigt. Ich kann aber auch einfach sagen, Geschlechterrollen kann ich in Wortfeldern festlegen oder ich schaue mir nur die Häufigkeit von geschlechtsspezifischen Pronomen an. Sie sehen schon: Man kann die Komplexität steigern oder reduzieren, je nachdem, wie man ein Phänomen formalisiert. Je mehr Aspekte Sie berücksichtigen und je mehr zusätzliche, eventuell textexterne Informationen Sie integrieren, desto komplexer wird das Vorhaben. Übrigens denkt natürlich jeder und jede immer: Mein Phänomen ist besonders komplex. Aber da würde ich dringend empfehlen, zu überlegen, ob das stimmt, und wenn ja, ob das so sein muss. Wenn Sie nämlich in allen fünf Dimensionen eine hohe Komplexität haben, dann werden Sie Ihr Forschungsprojekt wahrscheinlich nicht beenden.

Abb. 1: Die fünf Komplexitätsdimensionen der computationellen literaturwissenschaftlichen Textanalyse nach Gius (2019).



Bei der Dimension der Heterogenität der betrachteten Texte geht es nicht nur um die Frage, wie viele Text man analysiert. Die Frage ist: Wie heterogen ist die Grundlage, die ich mir angucke? Also etwas genereller gefragt: Wie heterogen ist das Korpus in Bezug auf relevante Texteigenschaften? In der Literaturwissenschaft habe ich zum Beispiel verschiedene Genres, Epochen oder Strömungen, Autor:innen, Themen und noch viel mehr, einige Merkmale vielleicht sogar innerhalb eines Textes. Ein idealer romantischer Text enthält zum Beispiel alle drei Großgattungen in einem Text, während wiederum ein Text der neuen Sachlichkeit diesbezüglich weniger komplex ist, weil er sprachlich einfacher und nur Prosa ist. Das heißt, zehn Kurzgeschichten einer Autor:in zu analysieren ist wahrscheinlich einfacher als alle Dramen des 19. Jahrhunderts und zusätzlich noch 30 poetologische Texte zu analysieren. Da haben Sie deutlich mehr Probleme mit hoher Komplexität, die Sie dann an anderen Stellen einfangen müssen, um die Umsetzbarkeit des Projekts nicht zu gefährden.

Bei der Dimension der Erkenntnisgenerierung geht es um die Frage: Wird die Textbasis durch Menschen oder durch Computer erschlossen? Vielleicht muss ich noch einmal dazu sagen, dass dieses Modell für die Planung computationeller Zugänge, nicht für generelle Komplexität konzipiert ist. Wenn Menschen die Forschungsschritte umsetzen, ist das potenziell komplexer, als wenn diese algorithmisch umgesetzt werden. Darum geht es hier aber nicht. Hier geht es um die Frage der computationellen Komplexität. Die Pole sind Lesen und automatisches Erschließen, wobei wir, weil es eben um computationelle Verfahren geht, bei beiden Daten erzeugen. Beim Lesen erzeugen wir die Daten typischerweise durch Annotation und nachdem das zwar ein komplexer Vorgang ist, die Komplexität aber nicht die computationelle Umsetzung betrifft, ist manuelle Annotation im Modell das am wenigsten komplexe Verfahren der Erkenntnisgenerierung.

Die letzte Dimension, der computationelle Erkenntnisbeitrag, ist ein bisschen die Königsdisziplin im Komplexitätswettbewerb. Die Frage ist, wie weit geht der Erkenntnisbeitrag der computationellen Methode? Also was soll der Rechner tun? Soll er interpretieren oder nicht? Die Komplexität reicht von einer einfachen Textanalyse bis zur Interpretation bzw. von einer Deduktion – bei der ich vorhersage, welche Kategorien zu welchen Befunden führen, und das wird dann so umgesetzt – über die Induktion – bei der erst aus der Analyse Kategorien entstehen – bis zur Abduktion – bei der es extrem schwierig ist, die Prozesse zu beschreiben, die die neuen Erkenntnisse generieren. Übrigens, eine kleine Randbemerkung dazu: Im Zuge der aktuellen KI-Entwicklungen erscheint es vielleicht so, dass man jetzt einfacher computationelle Abduktion durchführen kann. Ich würde sagen, dass man das vorgeschlagene Modell zur Komplexitätsbestimmung nutzen kann, um zu zeigen, dass das doch nicht so einfach ist, weil wir in dem Fall nicht wissen, was die darunterliegenden Konzepte sind, mit denen wir hantieren, wenn wir LLMs nutzen. Da gibt es unabhängig von der Komplexitätsfrage ein großes Evaluationsproblem. Für

die Frage der Komplexität der Dimension des computationellen Beitrags ist es auch ohne LLMs interessant zu überlegen: Was müsste ich tun, um den Rechner wirklich die Abduktion machen zu lassen? Ich glaube, dass das theoretisch möglich ist, aber ich habe nach wie vor keine Vorstellung davon, wie man das konkret angehen kann.

Was das Modell zur Komplexitätsbestimmung insgesamt angeht, sieht man idealerweise auch, wo am meisten Arbeit entstehen wird. Und wenn ich ein Vorhaben habe, wo alles auf maximale Komplexität gesetzt ist, dann habe ich entweder ein Riesenprojekt oder ein Problem oder vielleicht auch beides. Es entsteht möglicherweise auch für Projekte mit vielen Ressourcen ein Problem, weil sich die Komplexität der einzelnen Dimensionen potenziert und dadurch das Vorhaben zu wackelig werden kann.

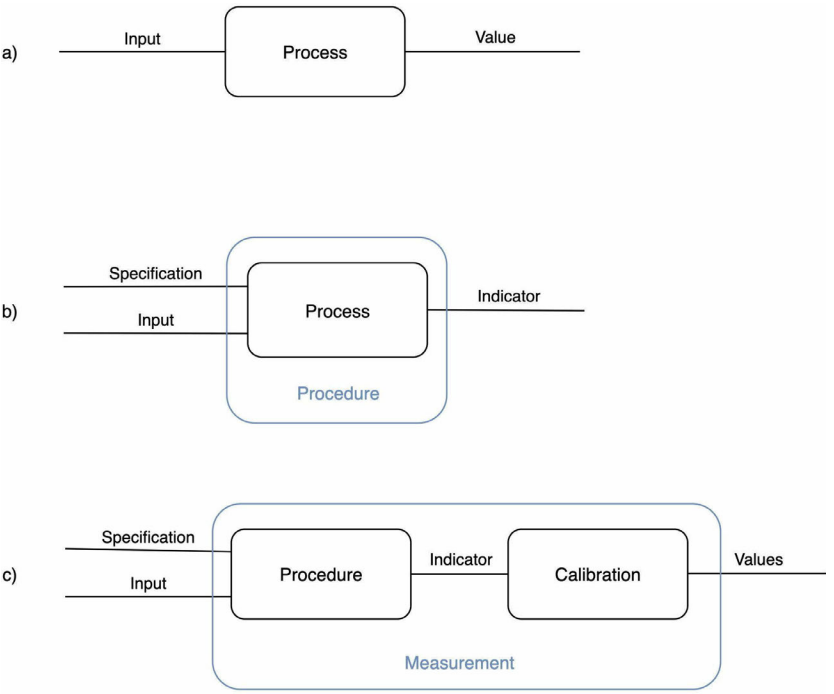
Der Forschungsprozess als Messprozess

Damit komme ich zu meinem letzten Teil. Es ist ein kurzer Vorschlag zum Forschungsprozess, der weiterhin aus der literaturwissenschaftlichen Textanalyse heraus gedacht ist. Aber auch hier hoffe ich, dass er genereller gültig ist und Sie etwas damit anfangen können. Er basiert auf dem Buch „Measurement Across the Sciences" von den drei Messtheoretikern Luca Mari, Mark Wilson und Andrew Maul (2023). Bevor Sie jetzt sagen: schon wieder Sciences, sage ich schnell: Sciences meint uns hier mit. Es geht in dem Buch nicht nur um die harten Wissenschaften. Außerdem glaube ich, dass die Literaturwissenschaft und auch alle Sozial- und Geisteswissenschaften bei bestimmten Forschungsfragen und vor allem Vorgehen durchaus mit den sogenannten harten Wissenschaften vergleichbar sind. Ich erinnere nochmal an die Messverfahren für Temperatur oder Gewicht. Wenn das harte Wissenschaft ist, dann betreiben wir auch harte Wissenschaft, weil wir auch Indikatoren messen, wenn auch vielleicht ein bisschen aufwendiger. Aber die Hoffnung ist, dass wir in einigen Jahrzehnten bei ein paar Dingen vielleicht sagen können: ja, früher wusste man nicht, wie man das macht, aber jetzt haben wir es herausgefunden und jetzt können wir das relativ einfach analysieren. Ein Beispiel für so eine Entwicklung sind übrigens Word Embeddings als computationelles Verfahren, das wiederum wesentlich zur Entwicklung der aktuellen LLMs beigetragen hat. Das Konzept dahinter kommt aus der Linguistik, genauer: der distributionellen Semantik, und ist je nachdem, wem es zugeschrieben wird, bald 100 Jahre alt. Seine Umsetzung ist hingegen erst Anfang dieses Jahrtausends gelungen und das Verfahren, was wir heute typischerweise nutzen, wurde erst 2013 implementiert.

Aber zurück zum Forschungsprozess. Mein Vorschlag bzw. eigentlich der Vorschlag von Mari und Kollegen ist, dass wir einen Forschungsprozess wie folgt betrachten: Es gibt einen Input, es gibt einen Prozess und es gibt einen Output mit einem Wert. So weit, so einfach. Das ist die grundlegendste abstrakte Struktur eines Forschungsprozesses. Was bedeutet das aber, wenn man rechnet? Das bedeu-

tet, ein Computer empfängt Eingaben, führt Berechnungen (den Prozess) durch und gibt Ergebnisse aus. Ich denke, auch das ist irgendwie wenig überraschend. Das Problem ist aber, dass dieses Input-Prozess-Output-System von der Umwelt getrennt ist. Hier wird es ungemütlich für uns kontextbasierte Wissenschaften. Man kann das trotzdem umsetzen, es ist aber keine typische Perspektive für die meisten Sozial- und Geisteswissenschaften.

Abb. 2: Der Forschungsprozess als einfacher (a) und weiter ausdifferenzierter (b, c) Messprozess (vgl. Gerstorfer und Gius 2025: 347).



Die Frage ist ja: Wie bekomme ich am Ende des Prozesses das heraus, was ich wissen will? Dazu gehört: Was muss ich als Input bereitstellen? Und: Was muss im Prozess passieren? In unserer Forschung ist das, was als Prozess passiert, oft mit dem, was ich herausfinden will, oder dem, was ich vorne hineingebe, nicht in dem strengen Sinne des Modells des Forschungsprozesses verbunden. Das habe ich vorhin an der Computational Narratology und der Lücke zwischen Theorie und Anwendung zu zeigen versucht.

Ich denke, man kann das Modell zu den Komplexitätsdimensionen ein Stück weit nutzen, um den eigenen Forschungsprozess zu durchdenken und sich dem Ab-

lauf als Input-Prozess-Output zu nähern. Als Input brauchen wir neben den Daten auf jeden Fall Konzepte und Phänomene sowie Informationen zur Kontextualisierung. Da trage ich jetzt Eulen nach Athen, aber ich sage es trotzdem: Metadaten zu Daten und alle weiteren Informationen, die im Prozess wichtig sind, gehören als Input – also von vornherein – dazu. Der Prozess muss dann darauf abgestimmt sein.

Das ist ein Denkmodell, das den Forschungsprozess als Messverfahren ernst nimmt. Wenn wir das machen, gehört dazu auch, und das habe ich bisher nur angedeutet, dass wir uns der Herausforderung der Entwicklung von neuen Messverfahren und Instrumenten stellen. Wobei die Verfahren und Instrumente nicht immer neu sein müssen, manchmal gibt es sie auch bereits. Das gilt es dann aber herauszufinden. Ich kann nicht einfach irgendeine Methode nehmen, weil ich sie beherrsche. Ich muss wissen, welche Methode für welche Texte geeignet ist. In qualitativer Forschung sind es noch ganz andere, noch komplexere Daten als Texte. Um bei den bisherigen Beispielen zu bleiben: Ich kann zwar mit einem Thermometer die Temperatur von ganz verschiedenen physikalischen Entitäten messen und so etwas über deren Temperatur herausfinden. Aber ich kann nicht mit einem Thermometer das Gewicht messen. Das klingt jetzt vielleicht ein bisschen absurd, aber genau das passiert gar nicht so selten, insbesondere wenn jemand gerade erst anfängt, computationell zu arbeiten.

Zu der Frage der Identifikation oder der Entwicklung eines geeigneten Instruments gehört auch, dass ich eine genauere Vorstellung davon haben sollte, wie die Phänomene, die mich interessieren, in den Daten zu finden sind. Dann muss ich sicherstellen, dass der Prozess auch das umsetzt, was in Anbetracht meiner Daten und meiner Forschungsfrage sinnvoll ist. Das muss alles festgelegt werden, damit das herauskommt, was ich will. Nicht im Sinne von „Ich wusste vorher schon, was rauskommt", sondern in dem Sinne, dass das Ergebnis adäquat ist in Bezug auf die gestellte Forschungsfrage. Ein gutes Kriterium dafür ist oft, wie ich als Mensch, wenn ich könnte, die Dinge analysieren würde.

Für die eigentliche Messprozedur gibt es ebenfalls eine Reihe von Fragen. Muss ich diese erst etablieren? Also: Muss ich erst Quecksilber irgendwo finden bzw. einen Indikator identifizieren, an dem ich das Messverfahren orientieren kann? Wie kann ich das normieren, wie bekomme ich eine Skala, die sinnvoll ist? Man sieht zum Beispiel bei der Temperatur, dass es verschiedene Nullpunkte in den verschiedenen Systemen gibt. Wie viel brauche ich? Was muss ich als Orientierung nehmen? Wie leicht kann ich es unterteilen? Ist es etwas, das wirklich linear steigt? Und so weiter und so fort. Das sind Detailfragen, die man sich stellen sollte, wenn man konkret wird.

Man sieht schon an den wenigen Gedanken, die ich bisher ausgeführt habe: Es ist meistens sehr aufwendig, neue Messverfahren zu entwickeln. Trotzdem ist mein Plädoyer, dass man sich einen Ruck gibt und sagt: Ich führe einen wissenschaftlichen Prozess in diesem strengen Input-Process-Output-Sinne durch. Denn ein

Messverfahren bietet einen Workflow aus explizit gemachten Schritten und macht, wenn das Vorgehen auch beschrieben wird, den jeweiligen Zugang nachvollziehbar und diskutierbar – womit es seine Wissenschaftlichkeit stärkt. Man kann den Mess-Workflow vielleicht nicht immer vollständig umsetzen, aber man sollte in diesem Modus denken und offenlegen, wo es noch Lücken gibt.

Ein kurzes Resümee

Damit bin ich fast am Ende meines Beitrags angelangt und möchte abschließend kurz resümieren, was ich gesagt habe bzw. mit dem Gesagten sagen wollte. Ich habe versucht, anhand der Quadratur des Kreises und der Frage der transzendenten Zahlen dazu anzuregen, über unsere eher hermeneutisch ausgerichteten Tätigkeiten und unsere Gegenstände nachzudenken. Dann habe ich über Messen gesprochen und versucht zu zeigen, dass das Messen etwas sehr Genaues ist, was einen komplexen Prozess erfordert. Zum anderen ist es aber eben nicht so stark an Wahrheit orientiert oder an Realität, wie man vielleicht denken möchte. Deshalb ist Genauigkeit nicht mit Wahrheit zu verwechseln. Ich habe gefragt, ob man Literatur messen darf, und versucht zu zeigen; ja, wobei das eigentlich die falsche Frage ist, denn die Antwort hängt, Stichwort Banane, eigentlich am disziplinären Selbstverständnis bzw. dem Blick auf den Gegenstandsbereich. Dann habe ich mich etwas intensiver mit den Verfahren und dem Selbstverständnis der Computational Literary Studies beschäftigt, wobei ich darauf verwiesen habe, dass es eine Lücke im Forschungsprozess gibt, die auch als fehlendes Interface zwischen qualitativen und quantitativen Ansätzen gesehen werden kann, wenn man Theoriearbeit als eher qualitativ einstuft. Und dann habe ich zwei Modelle vorgestellt, die vielleicht helfen können. Zum einen ging es darum, Komplexität genauer anzuschauen und sich auch klarzumachen, dass sie im Wesentlichen auf unseren Setzungen beruht. Das zweite Modell konnte ich nur kurz anreißen. Da ging es darum, den wissenschaftlichen Zugang als einen an Messverfahren orientierten Input-Process-Output-Forschungsprozess zu denken, der zwar aufwendiger ist, aber auch eine bessere Verknüpfung unserer Ausgangsfragen und -konzepte mit den Ergebnissen ermöglicht bzw. eigentlich sogar erzwingen sollte. Und damit bin ich am Ende angekommen und sage: Hier sollten wir weitermachen. Vielen Dank.

Literatur

Bunia, Remigius. 2011. „Das Handwerk in der Theoriebildung. Zu Hermeneutik und Philologie". In *Journal of Literary Theory* 5 (2). https://doi.org/10.1515/JLT.2011.014.

Chang, Hasok. 2004. *Inventing temperature: measurement and scientific progress*. Oxford studies in philosophy of science. Oxford University Press.

Computational Literary Studies Infrastructure. https://clsinfra.io/.

Dobson, James E. 2019. *Critical Digital Humanities: The Search for a Methodology*. *Topics in the digital humanities*. University of Illinois Press.

Gerstorfer, Dominik, und Evelyn Gius. 2025. „Operationalizing Operationalizing". In *Digital Humanities im deutschsprachigen Raum (DHd2025)*, 345–48. https://doi.org/10.5281/zenodo.14943080.

Gius, Evelyn. 2019. „Computationelle Textanalysen als fünfdimensionales Problem: Ein Modell zur Beschreibung von Komplexität". In *LitLab Pamphlet*, Nr. 8: 1–20.

Hatzel, Hans Ole, Haimo Stiemer, Chris Biemann, und Evelyn Gius. 2023. „Machine Learning in Computational Literary Studies". In *It – Information Technology* 65 (4–5). https://doi.org/10.1515/itit-2023-0041.

Helling, Patrick, Kerstin Jung, und Steffen Pielström. 2022. „Pragmatisches Forschungsdatenmanagement – qualitative und quantitative Analyse der Bedarfslandschaft in den Computational Literary Studies". In *Digital Humanities im deutschsprachigen Raum (DHd 2022)*. https://doi.org/10.5281/zenodo.6328021.

Hühn, Peter, John Meister, Wolf Schmid, und Jörg Schönert. 2013. *„The living handbook of narratology."* Hamburg University Press. http://www.lhn.uni-hamburg.de/.

Jannidis, Fotis: Computational Literary Studies. In: https://dfg-spp-cls.github.io/.

Journal of Computational Literary Studies (JCLS). https://jcls.io/.

Kinder-Kurlanda, Katharina. 2026. „Data Politics: Ethnografische Forschung mit/zu digitalen Methoden." In *Digital Humanities und Empirische Kulturwissenschaft. Fluide Forschungsfelder und digitale Daten in methodischer Reflexion*, herausgegeben von Lina Franken und Sabina Mollenhauer. Transcript.

Köppe, Tilmann, und Simone Winko. 2013. „Theorien und Methoden der Literaturwissenschaft". In *Handbuch Literaturwissenschaft: Methoden und Theorien*, herausgegeben von Thomas Anz, Bd. 2. Metzler. https://doi.org/10.1007/978-3-476-01271-5.

Mani, Inderjeet. 2013. „Computational Narratology". In *the living handbook of narratology*, herausgegeben von Peter Hühn, Jan Christoph Meister, John Schmid, und Jörg Schönert. Hamburg University. https://www-archiv.fdm.uni-hamburg.de/lhn/node/43.html.

Mari, Luca, Mark Wilson, und Andrew Maul. 2023. *Measurement Across the Sciences: Developing a Shared Concept System for Measurement*. Springer Series in Measure-

ment Science and Technology. Springer Nature. https://doi.org/10.1007/978-3-031-22448-5.

Piper, Andrew, Richard Jean So, und David Bamman. 2021. „Narrative Theory for Computational Narrative Understanding". *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 298–311. https://doi.org/10.18653/v1/2021.emnlp-main.26.

Schöch, Christof. 2017. „Quantitative Analyse". In *Digital Humanities. Eine Einführung*, herausgegeben von Fotis Jannidis, Hubertus Kohle, und Malte Rehbein. Metzler.

Schöch, Christof, Julia Dudar, und Evgeniia Fileva. 2023. *CLS INFRA D3.2: Series of Five Short Survey Papers on Methodological Issues (= Survey of Methods in Computational Literary Studies)*. Zenodo. https://doi.org/10.5281/zenodo.7892112.

Swift, Jonathan. o. J.: *„Gullivers Reisen zu mehreren Völkern der Welt."* Übers. v. Franz Kottenkamp. Verlag Philipp Reclam jun.

Weimar, Klaus. 2000. „Hermeneutik". In *Reallexikon der deutschen Literaturwissenschaft: Neubearbeitung des Reallexikons der deutschen Literaturgeschichte*, 3., neubearb. Aufl., herausgegeben von Harald Fricke, Klaus Grubmüller, Jan-Dirk Müller, und Klaus Weimar, II. De Gruyter.