

Methodenbeitrag: Projektkonzeption

Sandra Bläß

forTEXT

Thema:	Projektkonzeption	DOI:	10.48694/fortext.3796
Jahrgang:	1	Ausgabe:	12
Erscheinungsdatum:	30-11-2024	Erstveröffentlichung:	2019-12-16 auf fortext.net
Lizenz:			open & access

Allgemeiner Hinweis: Rot dargestellte Begriffe werden im Glossar am Ende des Beitrags erläutert. Alle externen Links sind auch am Ende des Beitrags aufgeführt.

1. Definition

Unter die Konzeption wissenschaftlicher Projekte, seien es Artikel, Abschlussarbeiten oder Vorträge, fallen das Ausmachen eines Forschungsgegenstandes (1. und 2. Fall in der folgenden Liste), die Wahl einer angemessenen Fragestellung (1. und 2. Fall) und Methodik (3. und 4. Fall) sowie die Koordination von Untersuchungsmaterial und zur Verfügung stehender Zeit (5. Fall). In der Projektkonzeption können Sie an einem oder mehreren dieser Punkte mit digitalen Tools und Hilfsmitteln ansetzen:

1. Digitale Organisation und Recherche von Textmaterial: Hierzu erhalten Sie genauere Informationen in den Artikeln Korpusbildung (Bläß 2024) und Bibliografieren (Flüh 2024b).
2. Exploration vorab: Sie untersuchen Textmaterial digital, um durch eine neue Perspektive auf potentielle Fragestellungen aufmerksam zu werden.
3. Digitale Unterstützung: Sie erweitern eine traditionell-literaturwissenschaftliche Arbeitsweise mit digitalen Methoden.
4. Distant-Reading-Einstieg (vgl. **Distant Reading**): Sie beginnen mit einer zu Ihrer Fragestellung passenden Distant-Reading-Technik. Mit den dabei gewonnenen Daten stellen Sie Ergebnishypothesen auf, die Sie danach direkt am Text prüfen können.
5. Koordination: Digitale Tools helfen Ihnen in organisatorischen Bereichen der Projektkonzeption, beispielsweise bei der Sortierung Ihrer Gedanken oder Ihrer Zeitplanung.

2. Anwendungsbeispiel

Sie haben bereits beschlossen, dass Sie Franz Kafkas *Der Bau* behandeln möchten, sind aber noch auf der Suche nach einem Thema (siehe Fall 2). Hier bietet es sich an, den Text zunächst digital zu explorieren. Sie finden ihn online (u. a. im Deutschen Textarchiv (DTA)) und können ihn ohne weiteren Aufwand etwa im Visualisierungstool Voyant (Flüh 2024a) einspeisen. Dabei sehen Sie in der **Wordcloud**, wie zentral beispielsweise die Begriffe „scheint“ und, etwas weniger häufig, „wahrscheinlich“ sind. Daraus könnten Sie die Idee ableiten, Unsicherheit und imaginierte Szenarien zu untersuchen. Falls Sie dem in einem **Close Reading** nachgehen möchten (siehe Fall 3), können Sie beispielsweise CATMA (Schumacher 2024) zur Unterstützung heranziehen und den Text manuell annotieren (Jacke 2024). In einem weiteren Schritt lassen sich die Ergebnisse hieraus in Queries abfragen (vgl. **Query**), um unter anderem deren Verteilung über den Text hinweg zu prüfen, **Kollokationen** zu betrachten oder zu visualisieren.

Ein möglicher Distant-Reading-Einstieg (siehe Fall 4) wäre, ein **Korpus** aus Kafka-Erzählungen zu erstellen und dieses mittels Topic Modeling (Horstmann 2024c) daraufhin zu untersuchen, ob Sie ein Unsicherheits- oder Möglichkeitstopic finden, oder in Verbindung mit welchen Begriffen und Topics zugehörige Schlüsselwörter fallen. Möglicherweise kommen derartige Phänomene in bestimmten Texten im Vergleich zu anderen häufig vor. Basierend auf den so gewonnenen Daten können Sie Hypothesen aufstellen und diese am Text überprüfen. Zur Koordination (Fall 5) können Sie z. B. Trello, Word oder Excel zu Hilfe nehmen, entweder zur Erstellung eines Zeitplans oder als Gliederungshilfe. Darin ordnen Sie Ihre Aufgaben spezifischen Zeitrahmen oder inhaltliche Punkte zu schreibenden Kapiteln zu.

3. Literaturwissenschaftliche Tradition

Über die letzten Jahrhunderte wurde eine große Fülle an Arbeitstechniken entwickelt, mit denen Wissen strukturiert, archiviert und kontextualisiert werden kann, um es für unterschiedliche Projekte aufzubereiten. Im Folgenden sind einige davon beschrieben.

Conrad Gessner versuchte sich ab 1545 in seiner *Bibliotheca Universalis* an einem Forschungsüberblick über alle seit dem Buchdruck erschienenen Texte, in dem er nicht nur Titel, sondern auch kompakte Inhaltsangaben sammelte. In seinem sogenannten Zettel- oder Karteibuch notierte er hierfür Kernthesen, zerschnitt diese Exzerpte und sortierte sie zu kontingenten Zusammenfassungen. Mit seinen so erarbeiteten Bibliografien gab er einen Überblick über mehr als 10.000 Bücher (Krajewski 2002, 16f.; Haarkötter 2013, 29f.).

An mittelalterliche Bibliothekskataloge („Archen“) angelehnt erschienen ab dem 18. Jahrhundert Anleitungen, mit welchen Maßen man einen „gelehrten Kasten“ bauen und in welche Größe das Papier dafür gebracht werden sollte. In ihnen sollte Wissen notiert, eingeordnet, wiedergefunden und in einen Kontext mit anderen Erkenntnissen gebracht werden (Haarkötter 2013, 30–32). Diese Kästen, heute als Zettelkästen bekannt, waren zentral für die Arbeiten unter anderem von Jean Paul, Vladimir Nabokov sowie Hans Blumenberg und finden noch heute Verwendung (vgl. Abb. 1). Der berühmteste Zettelkasten ist jedoch Niklas Luhmanns, der insgesamt etwa 90.000 Zettel umfasst. Er unterstützte Luhmann als ein „Zweitgedächtnis“ (Luhmann 1981, 225) dabei, verschiedenste wissenschaftliche Gebiete in über 50 Monografien und 500 Aufsätzen komplex bis ins Detail miteinander zu verknüpfen (Schmidt 2013, 86 und zur Veranschaulichung Abb. 3).



Abb. 1: Zettelkasten Jochen Missfeldts. Fotografie von DLA Marbach (Chris Korner). In: Heike Gfrereis und Ellen Strittmatter (Hrsg. 2013): Zettelkästen. Maschinen der Phantasie. Marbach: Deutsche Schillergesellschaft.

Auch Umberto Eco empfiehlt in seinem 1977 erschienenen Ratgeber zu wissenschaftlichen Abschlussarbeiten Karteikarten: nämlich je eine Kartei zum Bibliografieren, eine Lektürekartei (eine Karte hält pro Buch alles Essentielle für das Projekt fest) und eine Ideenkartei (Zitate nach Themen geordnet). Zudem wird im Ratgeber erkennbar, wie die Themeneingrenzung weit stärker als heute davon bestimmt war, welche Medien in lokalen oder per Reise zu erreichenden Bibliotheken vorliegen (Eco 1993, 65–67).

1940 bis 1946 arbeitete Roberto Busa seine Dissertation zu „praesentia“ bei Thomas von Aquin (was auf eine Untersuchung zur Kollokation des Wortes „in“ hinauslief) auf 10.000 Karteikarten aus. Er hielt ein Konkordanzverzeichnis für weitere Forschungsvorhaben zwar für sinnvoll, erkannte jedoch, dass sich ein solches von Hand nicht anlegen lassen würde. Aus diesem Grund überzeugte er das IT-Unternehmen IBM von der Erstellung entsprechender Lochkarten. Manfred Thaller bezeichnet diese Begebenheit als „Gründungsmythos“ der Digital Humanities“ (Thaller 2017, 3).

4. Diskussion

Je nachdem, wie Sie digitale Methoden innerhalb Ihres Projekts und dessen Konzeption einsetzen wollen (siehe Fälle 3 und 4), müssen das Projekt sowie Ihr Textmaterial verschiedene Voraussetzungen erfüllen und unterschiedliche Vorarbeiten durchlaufen. Das digitale Textkorpus muss passend zum Projekt ausgewählt werden und je nach Methode möglicherweise in einem bestimmten Format vorliegen und/oder ein **Preprocessing** durchlaufen. Darüber hinaus sollten Sie sich in die Methode selbst einarbeiten. Für die Anwendung suchen Sie ein angemessenes Tool und erlernen dessen Bedienung. Weiterhin gehen viele Methoden mit spezifischem oder möglicherweise ungewohntem Fachvokabular einher. Idealerweise lesen Sie daher Artikel zu Projekten, in denen die von Ihnen gewählte Methode verwendet wird, nicht nur daraufhin, worauf Sie inhaltlich achten können und in welchen Schritten oft gearbeitet wird, sondern auch, um Beispiele zu finden, wie sich Praktiken und Ergebnisse formulieren lassen. Für all diese Einarbeitungen sollten Sie in Ihrer Projektkonzeption je nach Kenntnisstand genügend Zeit einplanen: Nicht immer ist ein Projekt schneller abgeschlossen, wenn Sie es (teilweise)

digital durchführen. Gelegentlich können zudem Kosten aufkommen, beispielsweise, wenn Sie sich für ein kostenpflichtiges Tool oder die Auslagerung einzelner Arbeitsschritte entscheiden. Um den Methodeneinstieg möglichst effizient zu gestalten und gegebenenfalls für die einzelnen Aufgaben auch kostenfreie Tools zu finden, bietet Ihnen die forTEXT-Seite zahlreiche Anlaufmöglichkeiten.

Schruhl zufolge ist das deutliche Hervorheben der eigenen Methodik typisch für die Digital Humanities, insbesondere, um anderen zu ermöglichen, in ihrer Forschung darauf aufzubauen (Schruhl 2017, 249–252). Neben der Interpretation können Sie sich dementsprechend auch eine methodische Optimierung als ein (weiteres) Ziel setzen und darüber reflektieren, mit welchen alternativen Arbeitsschritten Sie Ihre Fragestellung hätten verfolgen können (siehe Fälle 3 und 4). So benennen auch Kirilloff u. a. (2018) in ihrer eigenen Studie ein Forschungsziel damit, „benefits and limitations of using computational tools“ zur Untersuchung des gewählten Themas herauszufinden und schließen mit Ideen ab, welche Methoden bereichernd für Anschlussprojekte wären. Wenn Sie digitale und klassische Methoden miteinander kombinieren möchten, können sich verschiedene Möglichkeiten und Voraussetzungen ergeben, mit welchen Schritten und Zielsetzungen Sie Ihre Fragestellung planen und verfolgen. Neben den Ergebnissen muss deswegen der gesamte Projektverlauf beschrieben werden, wofür eine ausführliche Dokumentation Ihrer Planung und Arbeitsweise im Projekt von Anfang an vorteilhaft ist (siehe Fall 5). Hierbei können unter anderem folgende Leitfragen helfen: Welches Korpus haben Sie gewählt und weshalb? Haben Sie es im Lauf des Projekts angepasst? Mit welcher Methodik haben Sie es untersucht, welche Vorarbeiten mussten Sie dafür durchführen und war Ihr Vorgehen letztlich geeignet für Ihre Zielsetzung? Je nach Methode bieten sich auch Abbildungen zur Veranschaulichung des Projektverlaufs an.

Für die ersten Recherchen und Ideenstrukturierungen zu Ihrem Thema befindet Eco, der absolute Idealfall wäre, alle für das Projekt benötigten Werke selbst zu besitzen und darin markieren, kommentieren und verweisen zu können. Dies sei jedoch unrealistisch, sodass ein Zurückgreifen auf Karteien und Fotokopien der Bücher vonnöten sei [Eco (1993), 82f.; 150–153; 161f.]. Mit einem digitalisierten Text haben Sie allerdings alle Vorteile aus Ecos Szenario: Sie können Notizen, Kommentare und Markierungen (vgl. [Annotation](#); [Metadaten](#)) anbringen und diese auch problemlos wieder löschen, insbesondere wenn Sie den Text für ein anderes Thema wiederverwenden wollen. Exzerpieren und Kontextualisieren von Textstellen aus verschiedenen digitalisierten Büchern ist durch die „copy & paste“-Funktionalität einfacher als in analog vorliegenden Texten. Suchen nach Begriffen und Textstellen lassen sich mit digitalen Abfragen schneller durchführen als händisch. Zudem können Sie, falls Sie noch keine gefestigten Ideen für Themen oder Fragestellungen haben, Texte explorativ mit verschiedenen Tools untersuchen und dadurch neue Perspektiven gewinnen. Dies ist insbesondere nützlich, um sich einen ersten Überblick über ein Korpus mit vergleichsweise großer Textmenge zu verschaffen (siehe Fall 2). Viele Texte liegen bereits in Onlinetextsammlungen und -datenbanken vor; je nach Textmenge kann es andernfalls auch lohnen, das Material selbst zu digitalisieren (siehe Textdigitalisierung ([Horstmann 2024a](#)); Transkribus ([Horstmann 2024b](#))), um von den genannten Vorteilen zu profitieren.

Bezüglich Fragestellungen erweitern sich Ihre Möglichkeiten, wenn Sie digitale Methoden mit in Ihre Projektkonzeption einbeziehen. Jochen Vogt erklärt zum Unterschied zwischen natur- und literaturwissenschaftlichen Projekten, letzteren gehe es „weniger um das Allgemeine als um das Besondere; [...] um den einzelnen Text. Ihn müssen wir, oder wollen wir verstehen, so vollständig und vielseitig wie möglich“. Dies unterstreicht er mit einer Aussage von Jürgen Habermas, derzufolge Naturwissenschaften zu eigen sei, methodischer zu arbeiten, nämlich empirisch-analytisch (Vogt 2002, 195f.). In den digitalen Literaturwissenschaften lässt sich diese Grenze zur Naturwissenschaft nicht mehr derart eindeutig ziehen, da empirische Ansätze aus den Natur- und Sozialwissenschaften für die Anwendung auf literarische Texte adaptiert (vgl. [Domäneadaptation](#)) werden. Auf den ersten Blick arbeiteten digitale Geisteswissenschaften zwar häufig mit demselben Text wie traditionelle Literaturwissenschaften, so Trilcke und Fischer (2018), genauer betrachtet seien die Forschungsgegenstände jedoch zwei verschiedene. In einem Fall würden Texte als Daten (vgl. [Text Mining](#)) tendenziell empirisch, im anderen Textinhalte interpretierend untersucht. Beides lässt sich jedoch auch zusammenzuführen, um für beide Gegenstände einen Erkenntnisgewinn zu verzeichnen ([Herrmann 2017](#)). So erhalten Sie neue Perspektiven auf Ihre Forschungsgegenstände, Fragestellungen und darauf, wie Sie sich mit Ihnen jeweils auseinandersetzen können. Praktische Beispiele hierfür finden Sie unter anderem in den Studien von Herrmann (2017) und Kirilloff u. a. (2018).

Die Koordination Ihres Projekts, d. h. dessen Planung und Strukturierung (siehe Fall 5), gestaltet sich digital flexibler. Wenn Sie beispielsweise Ihren Zeitplan oder Ihre Gliederung in Papierform organisieren, sich dann aber Ihre Pläne ändern und Sie etwas umstellen möchten, müssen Sie entweder alles neu notieren oder laufen durch Streichungen, Pfeile und schriftliche Ergänzungen Gefahr, den Überblick zu verlieren. Digitale Programme hingegen ermöglichen Ihnen das Löschen, Verschieben, Umformulieren und Umfärben von Text. Zudem lassen sich je nach Programm Querverweise und Verlinkungen anbringen, mit denen Sie Ihre Literatur und zugehörige Gedanken ordnen und denen Sie mit nur einem Mausklick folgen können.

Literaturverwaltungsprogramme wie Zotero ([Flüh 2024c](#)) etwa ermöglichen, Literatureinträge mit Tags zu kategorisieren und so projektübergreifend wieder zu finden. In Projektmanagement-Tools wie Trello können die erstellten Aufgaben- oder Themenkarten mithilfe von selbst definierbaren Listen und Labels geordnet werden (vgl. Abb. 2). Das ist zwar kein Pendant zu einem Zettelkasten, der alles enthalten kann, was Sie in Ihrer wissenschaftlichen Laufbahn rezipieren (vgl. Abb. 3), jedoch können damit größere Projekte grob strukturiert werden, kleinere oder Einzelaspekte von Projekten auch detaillierter. So können Sie Ihre Literatur beispielsweise erst thematisch

Kapiteln zuordnen und Exzerpte sowie Zitate dann kapitelweise in einem digitalen Projektmanagement-Tool organisieren. Darüber hinaus helfen diese Tools Ihnen dabei, in kollaborativen Projekten mit Ihren Kolleg*innen zu kommunizieren, Daten auszutauschen und Aufgaben zu koordinieren.

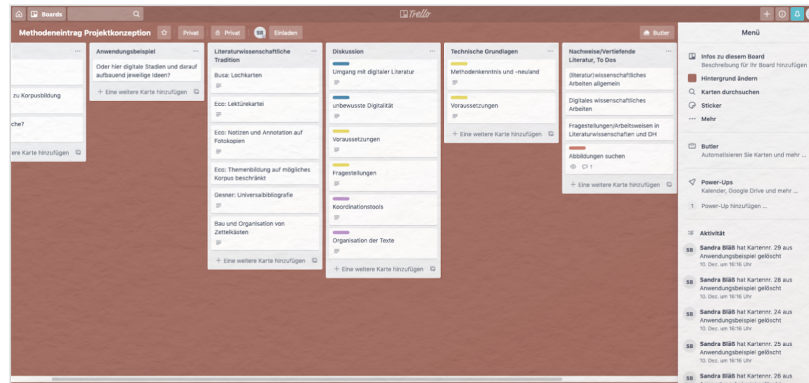


Abb. 2: Nutzung des Koordinationstools Trello zur Gliederung eines Projekts.

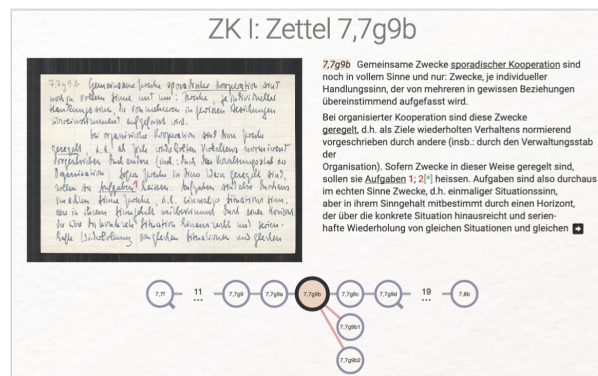


Abb. 3: Luhmann, Niklas: Zettel 7,7g9b aus Luhmanns Zettelkasten und, visualisiert durch das Niklas-Luhmann-Archiv, seine Verortung im Zettelsystem. Screenshot. CC BY-NC-SA 4.0.

Abhängig davon, welche weiteren Arbeitstechniken Sie nutzen, finden sich inzwischen zahlreiche Methoden ins Digitale übertragen, sodass es unter anderem auch Programme zur Schaubild- oder Mindmaperstellung gibt, zum Beispiel [draw.io](#) oder [Coggle](#). Diese lohnen sich vor allem für komplexere Visualisierungen und können zur Ideenfindung und mehr verwendet werden.

5. Technische Grundlagen

Die Tools und Methoden, die Sie zur Projektkonzeption heranziehen können, werden in verschiedensten Bereichen eingesetzt, deren Bedürfnisse und Zielsetzungen sich teilweise stark unterscheiden. Dementsprechend sind viele Angebote technisch niedrigschwellig gestaltet, sodass Sie für deren Nutzung kaum Vorwissen benötigen. Wie viel technisches Vorwissen Sie für die Adaption digitaler Techniken in Ihre Fragestellung brauchen, hängt davon ab, welche Methode Sie interessiert. Abhängig davon, welches Tool Sie für Ihre Projektkonzeption einsetzen, ist möglicherweise ein Internetzugang vonnöten: Word und Excel lassen sich auch offline verwenden, Trello hingegen nur im **Browser**. Dies hat jedoch den Vorteil, dass Sie Ihre Pläne von überall aus bearbeiten können und bei Teamprojekten von verschiedenen Orten aus arbeiten und kommunizieren können. Darüber hinaus müssen Sie bedenken, welche Funktionen Sie sich vom genutzten Tool wünschen. Während einige kostenpflichtige Organisationstools wie [Asana](#) oder [Wrike](#) viele Features haben, die kollaborative Aufgabenkoordination erleichtern, ist das kostenfrei nutzbare Trello in seinen Grundfunktionen schlichter und allgemeiner ausgerichtet. Damit lässt es sich einfacher für verschiedene Anwendungsszenarien zusätzlich zur Kollaboration anpassen (Einzelarbeit, Gliederung etc.).

Je nachdem, wie Sie Planung und Abstimmung in einem Projektmanagementtool genau gestalten möchten, können Sie sich von etablierten Verfahren wie Scrum oder Kanban inspirieren lassen. Beide werden aktuell insbesondere in Informatik- und Wirtschaftsprjekten genutzt, lassen sich jedoch auch für andere Bereiche adaptieren. Sie werden einzeln sowie miteinander kombiniert verwendet und sowohl analog als auch digital umgesetzt (Parsons u. a. 2018). Scrum basiert darauf, sich bestimmte Ziele für zwei- bis vierwöchige „Sprints“

zu setzen. In diesen Zeiträumen wird konzentriert und meist ohne Planänderung am Ziel gearbeitet; die Teammitglieder melden in regelmäßigen kurzen „Scrum“-Treffen zurück, wie ihre Aufgabe fortschreitet. Am Ende des Sprints wird reflektiert, was erreicht wurde, wie der Plan optimiert werden kann und was die nächsten Sprintziele sein sollen. Kanban wiederum setzt auf die Visualisierung des Arbeitsprozesses – etwa über ein Trelloboard – und hat weniger eindeutige Zeitlimits. Die Arbeitsschritte können laufend neu organisiert werden, allerdings wird ein Maximum an Aufgaben pro Liste gesetzt, um die Konzentration auf die Fertigstellung einzelner Aufgaben pro Person gegenüber überforderndem Multitasking zu priorisieren (Rehkopf 2019).

Externe und weiterführende Links:

- Asana Work Management Plattform: <https://web.archive.org/web/20241106125534/https://asana.com/> (Letzter Zugriff: 06.11.24)
- Coggle Mindmap Drawing Plattform: <https://web.archive.org/web/20241106125623/https://coggle.it/> (Letzter Zugriff: 06.11.24)
- Creative Commons Lizenzrechte: <https://web.archive.org/web/20241106125658/https://creativecommons.org/licenses/by-nc-sa/4.0/deed.de> (Letzter Zugriff: 06.11.24)
- Draw.io Online Drawing Plattform: <https://web.archive.org/web/20241106125743/https://app.diagrams.net/> (Letzter Zugriff: 06.11.24)
- Niklas Luhmann-Archiv ZK I Zettel: https://web.archive.org/web/20241106125937/https://niklas-luhmann-archiv.de/bestand/zettelkasten/zettel/ZK_1_NB_7-7g9b_V (Letzter Zugriff: 06.11.24)
- Trello Homepage: <https://web.archive.org/web/20241106120534/https://trello.com/> (Letzter Zugriff: 06.11.24)
- Wrike Workflow Plattform: <https://web.archive.org/web/20241106120657/https://www.wrike.com/> (Letzter Zugriff: 06.11.24)

Bibliographie

- Bläß, Sandra. 2024. Methodenbeitrag: Korpusbildung. Hg. von Evelyn Gius. *forTEXT* 1, Nr. 2. Korpusbildung (12. Juni). doi: 10.48694/fortext.3708, <https://fortext.net/routinen/methoden/korpusbildung>.
- Eco, Umberto. 1993. *Wie man eine wissenschaftliche Abschlußarbeit schreibt. Doktor-, Diplom- und Magisterarbeit in den Geistes- und Sozialwissenschaften*. Heidelberg: Müller.
- Flüh, Marie. 2024a. Toolbeitrag: Voyant. Hg. von Evelyn Gius. *forTEXT* 1, Nr. 5. Textvisualisierung (7. August). doi: 10.48694/fortext.3775, <https://fortext.net/tools/tools/voyant>.
- . 2024b. Methodenbeitrag: Digitales Bibliografieren. Hg. von Evelyn Gius. *forTEXT* 1, Nr. 11. Bibliografie (29. November). doi: 10.48694/fortext.3786, <https://fortext.net/routinen/methoden/digitales-bibliografieren>.
- . 2024c. Toolbeitrag: Zotero. Hg. von Evelyn Gius. *forTEXT* 1, Nr. 11. Bibliografie (29. November). doi: 10.48694/fortext.3787, <https://fortext.net/tools/tools/zotero>.
- Gfrereis, Heike und Ellen Strittmatter, Hrsg. 2013. *Zettelkästen. Maschinen der Phantasie*. Marbach am Neckar: Deutsche Schillergesellschaft.
- Haarkötter, Hektor. 2013. Fäden und Verzettelungen. Eine kurze Geschichte des Zettelkastens. In: *Zettelkästen. Maschinen der Phantasie*, hg. von Heike Gfrereis und Ellen Strittmatter, 28–40. Marbach: Deutsche Schillergesellschaft.
- Herrmann, Berenike J. 2017. In a text bed with Kafka. Introducing a mixed-method approach to digital stylistics. *Digital Humanities Quarterly* 11, Nr. 4. <http://www.digitalhumanities.org/dhq/vol/11/4/000341/000341.html> (zugegriffen: 12. November 2018).
- Horstmann, Jan. 2024a. Methodenbeitrag: Möglichkeiten der Textdigitalisierung. Hg. von Evelyn Gius. *forTEXT* 1, Nr. 3. Textdigitalisierung und Edition (12. Juni). doi: 10.48694/fortext.3741, <https://fortext.net/routinen/methoden/moeglichkeiten-der-textdigitalisierung>.
- . 2024b. Toolbeitrag: Transkribus. Hg. von Evelyn Gius. *forTEXT* 1, Nr. 3. Textdigitalisierung und Edition (12. Juni). doi: 10.48694/fortext.3746, <https://fortext.net/tools/tools/transkribus>.
- . 2024c. Methodenbeitrag: Topic Modeling. Hg. von Evelyn Gius. *forTEXT* 1, Nr. 8. Topic Modeling (7. Oktober). doi: 10.48694/fortext.3717, <https://fortext.net/routinen/methoden/topic-modeling>.
- Jacke, Janina. 2024. Methodenbeitrag: Manuelle Annotation. Hg. von Evelyn Gius. *forTEXT* 1, Nr. 4. Manuelle Annotation (7. August). doi: 10.48694/fortext.3748, <https://fortext.net/routinen/methoden/manuelle-annotation>.
- Kirilloff, Gabi, Peter J Capuano, Julius Fredrick und Matthew L Jockers. 2018. From a distance „You might mistake her for a man“: A closer reading of gender and character action in Jane Eyre, The Law and the Lady, and A Brilliant Woman. *Digital Scholarship in the Humanities* 33, Nr. 4 (1. Dezember): 821–844. doi: 10.1093/lc/fqy011, <https://academic.oup.com/dsh/article/33/4/821/5004302> (zugegriffen: 6. November 2018).
- Krajewski, Markus. 2002. *Zettelwirtschaft. Die Geburt der Kartei aus dem Geiste der Bibliothek*. Berlin: Kadmos.

- . 2013. Elektronische Literaturverwaltungen: kleiner Katalog von Merkmalen und Möglichkeiten. In: *Die Technik wissenschaftlichen Arbeitens: eine praktische Anleitung*, hg. von Norbert Franck, 91–109. Paderborn: Verlag Ferdinand Schöningh.
- Lauer, Gerhard. 2013. Die digitale Vermessung der Kultur. Geisteswissenschaften als Digital Humanities. In: *Big Data. Das neue Versprechen der Allwissenheit*, hg. von Heinrich Geiselberger und Tobias Moorstedt, 99–116. Berlin: Suhrkamp.
- Luhmann, Niklas. 1981. Kommunikation mit Zettelkästen. Ein Erfahrungsbericht. In: *Öffentliche Meinung und sozialer Wandel*, hg. von Horst Baier, Hans Mathias Kepplinger, und Kurt Reumann, 222–228. Opladen: Westdeutscher Verlag.
- Parsons, David, Rochelle Thorn, Milla Inkila und Kathryn Maccallum. 2018. Using Trello to Support Agile and Lean Learning with Scrum and Kanban in Teacher Professional Development. In: *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 720–724. doi: 10.1109/TALE.2018.8615399.
- Rehkopf, Max. 2019. Kanban vs. scrum: which agile are you? Uncover the key considerations when choosing between scrum or kanban, and what to do if you can't decide. *ATLASSIAN Agile Coach*. <https://www.atlassian.com/agile/kanban/kanban-vs-scrum> (zugegriffen: 2. Dezember 2019).
- Schmidt, Johannes F. K. 2013. Der Zettelkasten als Kommunikationspartner Niklas Luhmanns. In: *Zettelkästen. Maschinen der Phantasie*, hg. von Heike Gfrereis und Ellen Strittmatter, 84–93. Marbach: Deutsche Schillergesellschaft.
- Schruhl, Friederike. 2017. Literaturwissenschaftliche Wissensproduktion unter dem Einfluss der Digitalisierung. *Zeitschrift für Germanistik* 27, Nr. 2: 239–260.
- . 2018. Objektumgangsnormen in der Literaturwissenschaft. *Zeitschrift für digitale Geisteswissenschaften* Sonderband 3. https://zfdg.de/sb003_012.
- Schumacher, Mareike. 2024. Toolbeitrag: CATMA. Hg. von Evelyn Gius. *forTEXT* 1, Nr. 4. Manuelle Annotation (7. August). doi: 10.48694/fortext.3761, <https://fortext.net/tools/tools/catma>.
- Thaller, Manfred. 2017. Geschichte der Digital Humanities. In: *Digital Humanities. Eine Einführung*, hg. von Fotis Jannidis, Hubertus Kohle, und Malte Rehbein, 3–12. Stuttgart: Metzler.
- Trilcke, Peer und Frank Fischer. 2018. Literaturwissenschaft als Hackathon. Zur Praxeologie der Digital Literary Studies und ihren epistemischen Dingen. In: *Wie Digitalität die Geisteswissenschaften verändert: Neue Forschungsgegenstände und Methoden*, hg. von Martin Huber und Sybille Krämer, Sonderband der Zeitschrift für digitale Geisteswissenschaften: doi: 10.17175/sb003_003.
- Vogt, Jochen. 2002. *Einladung zur Literaturwissenschaft*. München: Wilhelm Fink.

Glossar

- Annotation** Annotation beschreibt die manuelle oder automatische Hinzufügung von Zusatzinformationen zu einem Text. Die manuelle Annotation wird händisch durchgeführt, während die (teil-)automatisierte Annotation durch **Machine-Learning-Verfahren** durchgeführt wird. Ein klassisches Beispiel ist das automatisierte **PoS-Tagging** (Part-of-Speech-Tagging), welches oftmals als Grundlage (**Preprocessing**) für weitere Analysen wie Named Entity Recognition (NER) nötig ist. Annotationen können zudem deskriptiv oder analytisch sein.
- Browser** Mit Browser ist in der Regel ein Webbrowser gemeint, also ein Computerprogramm, mit dem das Anschauen, Navigieren auf, und Interagieren mit Webseiten möglich wird. Am häufigsten genutzt werden dafür Chrome, Firefox, Safari oder der Internet Explorer.
- Close Reading** Close Reading bezeichnet die sorgfältige Lektüre und Interpretation eines einzelnen oder weniger Texte. Close Reading ist in der digitalen Literaturwissenschaft außerdem mit der manuellen **Annotation** textueller Phänomene verbunden (vgl. auch **Distant Reading** als Gegenbegriff).
- Data Mining** Data Mining gehört zum Fachbereich **Information Retrieval** und bezieht sich auf die systematische Anwendung computergestützter Methoden, die darauf abzielt, in vorhandenen Datenbeständen Muster, Trends oder Zusammenhänge zu erkennen. Textbasierte Formen des Data Minings sind u. a. **Text Mining**, **Web Mining** und **Opinion Mining**.
- Distant Reading** Distant Reading ist ein Ansatz aus den digitalen Literaturwissenschaften, bei dem computationale Verfahren auf häufig große Mengen an Textdaten angewandt werden, ohne dass die Texte selber gelesen werden. Meist stehen hier quantitative Analysen im Vordergrund, es lassen sich jedoch auch qualitative **Metadaten** quantitativ vergleichen. Als Gegenbegriff zu **Close Reading** wurde der Begriff insbesondere von Franco Moretti (2000) geprägt.
- Domäneadaption** Domäneadaption beschreibt die Anpassung einer in einem Fachgebiet entwickelten digitalen Methode an ein anderes Fachgebiet.
- HTML** HTML steht für *Hypertext Markup Language* und ist eine textbasierte Auszeichnungssprache zur Strukturierung elektronischer Dokumente. HTML-Dokumente werden von **Webbrowsern** dargestellt und geben die Struktur und Online-Darstellung eines Textes vor. HTML-Dateien können außerdem zusätzliche **Metainformationen** enthalten, die auf einer Webseite selbst nicht ersichtlich sind.

- Information Retrieval** Die Teildisziplin der Informatik, das Information Retrieval, beschäftigt sich mit der computergestützten Suche und Erschließung komplexer Informationen in meist unstrukturierten Datensammlungen.
- Kollokation** Als Kollokation bezeichnet man das häufige, gemeinsame Auftreten von Wörtern oder Wortpaaren in einem vordefinierten Textabschnitt.
- Korpus** Ein Textkorpus ist eine Sammlung von Texten. Korpora (Plural für „das Korpus“) sind typischerweise nach Textsorte, Epoche, Sprache oder Autor*in zusammengestellt.
- Lemmatisieren** Die Lemmatisierung von Textdaten gehört zu den wichtigen **Preprocessing**-Schritten in der Textverarbeitung. Dabei werden alle Wörter (**Token**) eines Textes auf ihre Grundform zurückgeführt. So werden beispielsweise Flexionsformen wie „schneller“ und „schnelle“ dem Lemma „schnell“ zugeordnet.
- Machine Learning** Machine Learning, bzw. maschinelles Lernen im Deutschen, ist ein Teilbereich der künstlichen Intelligenz. Auf Grundlage möglichst vieler (Text-)Daten erkennt und erlernt ein Computer die häufig sehr komplexen Muster und Gesetzmäßigkeiten bestimmter Phänomene. Daraufhin können die aus den Daten gewonnen Erkenntnisse verallgemeinert werden und für neue Problemlösungen oder für die Analyse von bisher unbekannten Daten verwendet werden.
- Markup Language** Markup Language bezeichnet eine maschinenlesbare Auszeichnungssprache, wie z. B. **HTML**, zur Formatierung und Gliederung von Texten und anderen Daten. So werden beispielsweise auch **Annotationen** durch ihre Digitalisierung oder ihre digitale Erstellung zu Markup, indem sie den Inhalt eines Dokumentes strukturieren.
- Metadaten** Metadaten oder Metainformationen sind strukturierte Daten, die andere Daten beschreiben. Dabei kann zwischen administrativen (z. B. Zugriffsrechte, Lizenzierung), deskriptiven (z. B. Textsorte), strukturellen (z. B. Absätze oder Kapitel eines Textes) und technischen (z. B. digitale Auflösung, Material) Metadaten unterschieden werden. Auch **Annotationen** bzw. **Markup** sind Metadaten, da sie Daten/Informationen sind, die den eigentlichen Textdaten hinzugefügt werden und Informationen über die Merkmale der beschriebenen Daten liefern.
- Named Entities** Eine Named Entity (NE) ist eine Entität, oft ein Eigenname, die meist in Form einer Nominalphrase zu identifizieren ist. Named Entities können beispielsweise Personen wie „Nils Holgerson“, Organisationen wie „WHO“ oder Orte wie „New York“ sein. Named Entities können durch das Verfahren der Named Entity Recognition (NER) automatisiert ermittelt werden.
- Opinion Mining** Unter Opinion Mining, oder Sentiment Analysis, versteht man die Analyse von Stimmungen oder Haltungen gegenüber einem Thema, durch die Analyse natürlicher Sprache. Das Opinion Mining gehört zu den Verfahren des **Text Minings**.
- POS** PoS steht für *Part of Speech*, oder „Wortart“ auf Deutsch. Das PoS- **Tagging** beschreibt die (automatische) Erfassung und Kennzeichnung von Wortarten in einem Text und ist ein wichtiger **Preprocessing**-Schritt, beispielsweise für die Analyse von **Named Entities**.
- Preprocessing** Für viele digitale Methoden müssen die zu analysierenden Texte vorab „bereinigt“ oder „vorbereitet“ werden. Für statistische Zwecke werden Texte bspw. häufig in gleich große Segmente unterteilt (*chunking*), Großbuchstaben werden in Kleinbuchstaben verwandelt oder Wörter werden **lemmatisiert**.
- Query** *Query* bedeutet „Abfrage“ oder „Frage“ und bezeichnet eine computergestützte Abfrage zur Analyse eines Textes. Um Datenbestände zu durchsuchen, werden Abfragesprachen eingesetzt, die *Queries* (Anfragen) an den Datenbestand senden. So bilden alle möglichen Queries zusammen die *Query Language* eines Tools.
- Text Mining** Das Text Mining ist eine textbasierte Form des **Data Minings**. Prozesse & Methoden, computergestützt und automatisch Informationen bzw. Wissen aus unstrukturierten Textdaten zu extrahieren, werden als Text Mining zusammengefasst.
- Type/Token** Das Begriffspaar „Type/Token“ wird grundsätzlich zur Unterscheidung von einzelnen Vorkommnissen (Token) und Typen (Types) von Wörtern oder Äußerungen in Texten genutzt. Ein Token ist also ein konkretes Exemplar eines bestimmten Typs, während ein Typ eine im Prinzip unbegrenzte Menge von Exemplaren (Token) umfasst.
Es gibt allerdings etwas divergierende Definitionen zur Type-Token-Unterscheidung. Eine präzise Definition ist daher immer erstrebenswert. Der Satz „Ein Bär ist ein Bär.“ beinhaltet beispielsweise fünf Worttoken („Ein“, „Bär“, „ist“, „ein“, „Bär“) und drei Types, nämlich: „ein“, „Bär“, „ist“. Allerdings könnten auch vier Types, „Ein“, „ein“, „Bär“ und „ist“, als solche identifiziert werden, wenn Großbuchstaben beachtet werden.
- Web Mining** Unter Web Mining versteht man die Anwendung von Techniken des **Data Mining** zur Extraktion von Informationen aus dem World Wide Web. Das Web Mining ist ein Teilbereich des Data Minings und zählt zu einem der wichtigsten Anwendungsgebiete für das **Text Mining**.
- Wordcloud** Eine *Wordcloud*, oder auch Schlagwortwolke, ist eine Form der Informationsvisualisierung, beispielsweise von Worthäufigkeiten in einem Text oder einer Textsammlung. Dabei werden unterschiedlich

gewichtete Wörter, wie die häufigsten Wörter, i.d.R. größer oder auf andere Weise hervorgehoben dargestellt. Die horizontale/vertikale Ausrichtung und die Farbe der dargestellten Wörter hat meistens allerdings keinen semantischen Mehrwert.