


Ressourcenbeitrag: Tagset Interpretationstexte analysieren

Stefan Descher¹

1. Georg-August-Universität Göttingen

forTEXT

Thema:	Manuelle Annotation	DOI:	10.48694/fortext.3758
Jahrgang:	1	Ausgabe:	4
Erscheinungsdatum:	2024-08-07	Erstveröffentlichung:	2020-07-27 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. Kurzbeschreibung

Das Tagset „Interpretationstexte analysieren“ dient der Annotation literaturwissenschaftlicher Interpretationstexte. Es stellt Kategorien zur Analyse wichtiger Strukturmerkmale und anderer Eigenschaften von Interpretationen bereit (z. B. zur Markierung zentraler Thesen, von Formen des Umgangs mit Forschungsliteratur usw.). Darüber hinaus kann es auch für die Annotation von wissenschaftlichen Texten im Allgemeinen verwendet werden und lässt sich je nach Bedarf leicht ergänzen bzw. anpassen. Das Tagset erhalten Sie auf Zenodo (forTEXT 2020).

2. Anwendungsbeispiel

Das Tagset ist vor allem für Projekte geeignet, die sich mit der Erforschung der konkreten literaturwissenschaftlichen Praxis befassen. Ein typisches Einsatzfeld dürfte die manuelle Annotation (Jacke 2024) mittelgroßer Textkorpora (vgl. Korpus) darstellen, um Häufigkeitsanalysen oder diachrone und synchrone Vergleiche zwischen wissenschaftlichen Interpretationstexten bzw. Forschungstexten im Allgemeinen zu ermöglichen. Konkrete Fragen wären z. B., wie häufig und an welcher Stelle sich Verfasser*innen von Interpretationstexten auf andere Forschungsbeiträge beziehen, wie oft sie ein Forschungsdesiderat markieren oder wie häufig sie ihre Beurteilungskriterien für Interpretationen explizit machen. Ein konkretes Projekt könnte z. B. untersuchen, ob sich Argumentationspraktiken und strukturelle Merkmale von Interpretationen in zwei Untersuchungskorpora voneinander unterscheiden – z. B. ob die Anzahl der durchschnittlich zitierten Forschungsbeiträge in einem Korpus, das Interpretationen zu Texten männlicher Autoren enthält, signifikant von der durchschnittlichen Anzahl zitierter Forschungsbeiträge in einem Korpus abweicht, das Interpretationen zu Texten von Autorinnen enthält. Denkbar ist auch ein Einsatz in der universitären Lehre, etwa um einen reflektierten und problembewussten Umgang mit Forschungsliteratur einzuüben. So wird die Annotation der Hauptthese(n) eines Interpretationstexts bei verschiedenen Annotator*innen immer wieder einmal zu unterschiedlichen Resultaten führen. Solche Unterschiede in den Annotationsergebnissen können dazu einladen, über implizite Vorannahmen nachzudenken, die sowohl das Lesen als auch das Verfassen von Forschungstexten leiten – in diesem Fall implizite Vorannahmen darüber, was eine Hauptthese eigentlich ist und woran man sie erkennt.

3. Literaturwissenschaftlicher Kontext

In der Literaturwissenschaft werden Sie sich nicht nur mit literarischen Texten, sondern immer wieder auch mit Forschungsliteratur beschäftigen, insbesondere mit Textinterpretationen. Das Verfassen und Lesen von Interpretationen gehört geradezu zur literaturwissenschaftlichen Alltagspraxis. Diese Alltagspraxis wird seit einigen Jahren im Zuge der sogenannten ‚Praxeologie‘ verstärkt erforscht (Albrecht u. a. 2015). Praxeologische Forschungsprojekte zielen darauf ab, die konkreten Praktiken der Literaturwissenschaft zu beschreiben und zu untersuchen – das also, was Literaturwissenschaftler*innen tun, wenn sie Literaturwissenschaft betreiben, und wie sie es tun. Das Tagset Interpretationstexte analysieren ist dem Forschungsfeld der literaturwissenschaftlichen Praxeologie zuzuordnen. Es ging aus einem Projekt an der Universität Göttingen hervor, das sich der Erforschung der literaturwissenschaftlichen Interpretationspraxis widmet (<https://www.argulit.uni-goettingen.de>). Darin wird der Frage nachgegangen, auf welche verschiedenen Weisen Interpret*innen ihre Interpretationshypothesen plausibilisieren.

Ältere Studien zur Interpretationspraxis (Grewendorf 1975; von Savigny 1976) konzentrierten sich vorrangig auf die rein argumentative Dimension von literaturwissenschaftlichen Interpretationen. Im Zentrum standen

also Fragen wie „Was ist die These?“, „Was sind die Argumente?“. Im Unterschied zu diesen Studien geht das aktuelle Projekt davon aus, dass zur Plausibilisierung von Hypothesen auch viele ‚weichere‘ Faktoren beitragen, die über die rein argumentative Dimension hinausgehen. Relevant ist z. B. auch, wie Interpret*innen ihre Beiträge aufbauen, an welcher Stelle sie ihre zentralen Thesen positionieren, ob sie diese Thesen explizit als solche hervorheben oder nicht, ob sie ihre Ergebnisse zusammenfassen, ob und wie sie sich mit anderer Forschungsliteratur auseinandersetzen, ob sie ihre Maßstäbe für gelungene Interpretationen explizit machen usw. Das hier vorgestellte Tagset dient der Erfassung von Texteigenschaften dieser Art.

Die Anwendung einiger Tags basiert auf den Ergebnissen von argumentationstheoretischen, linguistischen u. a. Forschungen, auf die an den entsprechenden Stellen verwiesen wird. Da im Folgenden nur kurze Beschreibungen und grundlegende Hinweise zur Verwendung der Tags gegeben werden können, seien Annotator*innen zur Vertiefung auf diese Arbeiten verwiesen.

4. Tagset

Sie können das Tagset Interpretationstexte analysieren auf Zenodo als XML-Datei herunterladen (forTEXT 2020), um es dann in geeignete Tools (beispielsweise CATMA (Schumacher 2024)) zu importieren und es dort zu verwenden. Abbildung 1 zeigt die im Tagset enthaltenen Tags in ihrer hierarchischen Struktur sowie die Properties und Values.

Tagsets	Tags	Properties	Values
Interpretationstexte analysieren	1 Hauptthese		
	2 Argumentationssignale		
	2.2 Indikatoren normative Argumentation		
	2.1 Argumentationsindikatoren allgemein		
	3 Gliederung & Aufbau		
	3.1 Überblick Gliederung		
	3.3 Zentrale These(n) am Ende aufgegriffen	explizit_oder_umformuliert	explizit umformuliert
	3.2 Zusammenfassung der Ergebnisse		
	4 Geltungsmodifikation	Abschwächung_oder_Verstärkung_einer_These	Abschwächung Verstärkung
	5 Qualitätskriterien für Interpretationen		
	6 Umgang mit Forschung		
	6.3 Konkreter Forschungsbezug	Primärtextforschung_oder_andere_Forschung	Primärtextforschung andere Forschung
	6.2 Markierung eines Forschungsdesiderats	Primärtextforschung_oder_andere_Forschung	Primärtextforschung andere Forschung
	6.1 Forschungsüberblick	Primärtextforschung_oder_andere_Forschung	andere Forschung Primärtextforschung

Abb. 1: Tagset Interpretationstexte analysieren

5. Richtlinien zur Anwendung

Die folgenden Richtlinien enthalten nur spezifische Anwendungshinweise für die Tags, die sich auf der untersten Hierarchieebene befinden (also beispielsweise den Tag „Forschungsüberblick“). Hierarchisch höherliegende Kategorien (z. B. „Umgang mit Forschung“) dienen vor allem der Systematisierung.

5.1. Hauptthese

- Tag *Hauptthese*: Dieser Tag dient der Markierung der zentrale(n) These(n) eines Interpretationstextes. Was eine Hauptthese ist und an welchen Merkmalen man sie erkennt, ist nicht leicht zu beantworten. Hauptthesen werden in literaturwissenschaftlichen Interpretationstexten nicht immer explizit als solche markiert, z. B. durch Formulierungen wie „Meine zentrale These lautet...“ oder „In diesem Aufsatz soll dafür argumentiert werden, dass...“. Im Extremfall können sie sogar vollständig implizit bleiben und müssen

von den Annotator*innen aufgrund der Kenntnis des gesamten Textes erschlossen werden. Daher ist bei der kollaborativen Annotation durchaus zu erwarten, dass Annotator*innen nicht immer zu denselben Ergebnissen kommen werden. Zur Identifikation von Hauptthesen können sich Annotator*innen aber an folgenden Fragen orientieren: Worauf läuft die Argumentation des gesamten Aufsatzes hinaus? Welche Thesen sind für den bzw. die Verfasser*in am wichtigsten? Geben der Veröffentlichungskontext, der Titel oder Kapitelüberschriften weitere Hinweise auf die zentrale(n) These(n)?

- Länge der annotierten Passage: i. d. R. ein Satz.
- Indikatoren: explizite Markierungen wie „Meine zentrale These lautet...“; allerdings häufig nur indirekt erschließbar aufgrund der Kenntnis des gesamten Textes.
- Beispiel: „In Kleists ‚Michael Kohlhaas‘ irritiert gerade diese Distanz zwischen Autor und Leser, so die These dieses Beitrags, immer wieder inszenatorisch die Logik der Substitution, die sowohl jedem Akt der Lektüre als auch jedem Tauschakt im Sinne der episteme der Repräsentation zugrunde liegt.“ (Frey 2003, 296)

5.2. Argumentationssignale

Argumentationssignale: Bei Interpretationstexten handelt es sich wie bei allen wissenschaftlichen Texten in der Regel um argumentative Texte: Typischerweise werden darin Argumente für oder gegen bestimmte Thesen bzw. Konklusionen formuliert (Descher und Petraschka 2019, 37–48). Es gehört jedoch zu den charakteristischen Phänomenen der Argumentationspraxis, dass man Argumente nur selten an ihrer sprachlichen Oberflächenstruktur erkennen kann (van Eemeren und Grootendorst 2004, 95–122). In einigen Fällen allerdings gibt es auch explizite Argumentationssignale, d. h. sprachliche Hinweise darauf, dass an einer bestimmten Stelle argumentiert werden soll, dass es sich bei einer bestimmten Behauptung um ein Argument oder um eine Konklusion handelt usw. (Knott und Dale 1993; Stab und Gurevych 2014). Die Tags in der Kategorie „Argumentationssignale“ dienen der Annotation solcher Indikatoren.

- Tag *Argumentationsindikatoren allgemein:* Mit diesem Tag können sprachliche Indikatoren markiert werden, die darauf hinweisen, dass durch den bzw. die Verfasser*in ein argumentatives Ziel verfolgt wird. Dies sind insbesondere solche Textbestandteile, die explizit anzeigen, dass es sich um eine These (Konklusion) oder ein Argument handelt. (Hinweis: Für eine noch differenziertere Annotation kann der Tag ggf. mit einer Property versehen werden, die z. B. die Values „Indikatoren für Konklusionen“, „Indikatoren für Argumente“ und „sonstige Argumentationsindikatoren“ hat.)
 - Länge der annotierten Passage: Wörter oder kurze Wortgruppen.
 - Indikatoren: Worte wie „These“, „Argument“, „Beleg“, „Beweis“, „Folgerung“, „Gründe“ oder Formulierungen wie „daraus folgt, dass“ (zur Markierung einer Konklusion) oder „dafür spricht, dass“ (zur Markierung von Argumenten für eine Konklusion) etc.
 - Beispiel: „Die Ethik, so verstehe ich Deleuze, ist erst da möglich, wo die Moral überwunden ist. Und diese Überwindung findet sich in Kleists Erzählung, so meine These, in der Figur Lisbeths.“ (Schubarth 2011, 48)
- Tag *Indikatoren normative Argumentation:* Unter „normative Argumentation“ werden solche Argumentationen verstanden, die auf die Begründung einer Norm oder einer Handlungsanweisung abzielen. Die zu begründende These bzw. die Konklusion ist i. d. R. ein ‚Sollens-Satz‘ („Also sollte man x tun.“) oder ein Satz, der bestimmte Handlungen vorschreibt oder verbietet („Also darf man keinesfalls x annehmen.“). Der Tag dient der Markierung von sprachlichen Merkmalen, die auf eine solche normative Argumentation hinweisen.
 - Länge der annotierten Passage: i. d. R. einzelne Worte.
 - Indikatoren: Worte, die Handlungsanweisungen bzw. die Aufstellung einer Norm anzeigen, wie z. B. „soll“, „darf“ oder „muss“.
 - Beispiel: „Will man unserm Dichter [Wieland; S.D.] gerecht werden, so *darf* man dieses Werk [*Der Goldne Spiegel*; S.D.] nicht allzu ernst nehmen. Es ist ein Zeugnis seines Benehmens in einer bestimmten Situation, ein Abbild der schillernden Außenseite, welche der berühmte Dichterphilosoph für die tonangebende Gesellschaft zurecht machte, nicht ein Bild dessen, was er wirklich war.“ (Sengle 1949, 260)

5.3. Gliederung und Aufbau

Gliederung und Aufbau: Diese Kategorie umfasst Tags, mit denen allgemeine Strukturmerkmale von Interpretationstexten annotiert werden können. Darunter fallen Passagen, in denen ein Überblick über den Aufbau des Texts gegeben wird, zentrale Gedankengänge zusammengefasst oder die Hauptthese(n) abschließend erneut aufgegriffen werden.

- Tag *Überblick Gliederung:* Hier wird annotiert, ob und wo Verfasser*innen einen Überblick über den Aufbau bzw. die Gliederung ihres Interpretationstextes geben.

- Länge der annotierten Passage: i. d. R. ganze Sätze oder Absätze.
- Indikatoren: Formulierungen wie „im Folgenden...“ oder „In Abschnitt 1 werde ich..., in Abschnitt 2...“ usw.
- Beispiel: „In diesem [...] Sinne handelt es sich bei der Erzählung [Ulrike Draesners ‚Rosakäfer‘; S.D.] um ein Beispiel von *Meta-Fantastik* [...]. *Unser Beitrag gliedert sich dabei wie folgt: Zunächst beschreiben wir Draesners ‚Rosakäfer‘ als Kontrafaktur von Kafkas ‚Die Verwandlung‘ (1); anschließend erläutern wir, inwiefern Draesners Text auch den Fantastik-Diskurs zum Gegenstand hat (2).*“ (Ertel und Köppe 2017, 191)
- Tag *Zusammenfassung der Ergebnisse*: Hier wird annotiert, ob und wo Verfasser*innen die Ergebnisse ihrer Interpretation zusammenfassen.
 - Länge der annotierten Passage: i. d. R. ganze Sätze oder Absätze.
 - Indikatoren: Formulierungen wie „Zusammenfassend lässt sich sagen, dass...“ oder „In den vorangehenden Abschnitten wurde gezeigt, dass...“.
 - Beispiel: „*Zusammenfassend lässt sich sagen, dass das lyrische Ich [in Dürrenmatts Schweizerpsalm III; S.D.] vor allem die schweizerische Politik und nicht die Schweiz im Allgemeinen oder die einzelnen Menschen kritisiert. Es ist die Politik, die als eine heuchlerische und korrumpierte entlarvt und demzufolge entmythologisiert wird.*“ (Conterno 2014, 303)
- Tag *Zentrale These(n) am Ende aufgegriffen*: Hier wird annotiert, ob die Verfasser*innen ihre zentrale(n) These(n) bzw. Hauptthese(n) am Ende des Aufsatzes noch einmal explizit oder in Umformulierung aufgreifen.
 - Länge der annotierten Passage: i. d. R. ein Satz.
 - Indikatoren: Formulierungen wie „Ich habe hier für die These argumentiert“ oder „In den vorangehenden Abschnitten wurde gezeigt, dass...“.
 - Property: Der Tag kann zwei Werte annehmen: ‚explizit‘ und ‚umformuliert‘, je nachdem, ob die zentrale(n) These(n) am Ende explizit (d. h. in der Regel wörtlich) oder in anderer Formulierung erneut erwähnt werden.
 - Beispiel: „*Die Untersuchung von Frühlings Erwachen hat bewiesen, daß die Dialoggestaltung des Dramas der sozialkritischen Intention Wedekinds durchaus untergeordnet ist. Seinem didaktischen Ziel gemäß ist das Stück stark publikumsgerichtet.*“ (Kuhn 1981, 79)

5.4. Geltungsmodifikation

- Tag *Geltungsmodifikation*: Durch sprachliche Signale können Verfasser*innen anzeigen, wie sicher sie sich mit einer bestimmten These sind bzw. wie stark der Geltungsanspruch ist, mit dem sie eine These vertreten (Toulmin 2003, 93f.). Sie können die Geltung ihrer Thesen also modifizieren, d. h. entweder abschwächen („Vielleicht ist es daher so, dass...“) oder verstärken („Zweifelloso ist es daher so, dass...“). Der Tag dient der Markierung von sprachlichen Signalen, die Geltungsmodifikationen dieser Art anzeigen.
 - Länge der annotierten Passage: i. d. R. ein Wort oder kürzere Wortgruppen.
 - Indikatoren: Worte und Formulierungen wie „vielleicht“, „möglicherweise“, „Es lässt sich vermuten, dass“ usw. (Abschwächung der Geltung) oder „zweifelloso“, „ganz sicher“, „Daher folgt zwingend“ usw. (Verstärkung der Geltung).
 - Property: Der Tag kann zwei Werte annehmen: ‚Abschwächung‘ und ‚Verstärkung‘.
 - Beispiel: „*Ohne Zweifel* allegorisiert diese anthropomorphistische Darstellung des willkürlichen Apparates und seiner Entscheidungsprozesse [in Kafkas *Das Schloss*] eine fundamentale totalitäre Struktur.“ (Sng 2007, 217)

5.5. Qualitätskriterien für Interpretationen

- Tag *Qualitätskriterien für Interpretationen*: Interpret*innen können explizit machen, was für sie eine gute Interpretation auszeichnet, d. h. welche Wertmaßstäbe bzw. Qualitätskriterien ihrer Interpretation zugrunde liegen. Wertmaßstäbe für Interpretationen können sehr unterschiedlich beschaffen sein und z. B. logisch-argumentative Aspekte (etwa ‚Plausibilität‘), aber auch ästhetische (‚Eleganz‘) oder rezeptionsseitige Aspekte (‚wirkt anregend‘) umfassen (Descher u. a. 2017, 46–49). Der Tag dient der Markierung von expliziten Nennungen solcher Qualitätskriterien.
 - Länge der annotierten Passage: i. d. R. ein Wort oder kürzere Wortgruppen.
 - Indikatoren: Wertprädikate oder Beschreibungen von Leistungen einer Interpretation, z. B. „plausibel“, „erhellend“, „fruchtbar“, „gut begründet“, „regt zum Nachdenken an“, „erklärt eine unverständliche Stelle“ usw.
 - Beispiel: „Das Verb ‚starren‘ [am Beginn von Alfred Döblins Erzählung *Die Ermordung einer Butterblume*; S.D.] bedeutet hinsichtlich des Blickes zweierlei: Entweder der Erzähler und wir mit ihm schauen dem Mann direkt in die Augen, was der Hinweis auf die Farbe der Augen nahelegt, oder wir starren mit ihm auf den Boden. Für die ersten zehn Wörter des Satzes scheint allein die erste

Möglichkeit *plausibel*, die Worte danach jedoch ziehen auch dem Leser buchstäblich den Boden unter den Füßen weg.“ (Kocher 2017, 99)

5.6. Umgang mit Forschung

Umgang mit Forschung: Diese Kategorie umfasst Tags, mit denen Bezüge auf Forschungsliteratur annotiert werden können.

- Tag *Forschungsüberblick:* Mit diesem Tag können Passagen markiert werden, in denen die Verfasser*innen eines Interpretationstextes einen Überblick über die bisherige Forschung zu ihrem Untersuchungsgegenstand geben.
 - Länge der annotierten Passage: Mindestens ein Satz, i. d. R. ein oder mehrere Absätze.
 - Indikatoren: Passagen, in denen relevante Forscher*innen und deren Forschungsbeiträge erwähnt und zentrale Thesen und Argumente referiert werden.
 - Property: Der Tag kann zwei Werte annehmen („Primärtextforschung“ oder „andere Forschung“), je nachdem, ob ein Überblick über die Forschung zum jeweils interpretierten Primärtext gegeben wird oder ein Überblick über andere Forschungsfelder.
 - Beispiel: „*Kaum eine Gestalt der neueren Dramatik hat derart widersprüchliche Bewertungen provoziert wie Meister Anton [...]. Ob die ‚größte Gestalt‘ des Stücks [Friedrich Hebbels Maria Magdalena; S.D.] auch ‚die größte Hebbels überhaupt‘ ist, wie Elise Dosenheimer mutmaßt [...], mag offen bleiben. Aber dies lehrt die Rezeptionsgeschichte [...] geradezu handgreiflich, wie nämlich die Vorentscheidung der Urteilenden als hermeneutische Bedingung je und je ins Spiel kommt. Gilt für den einen die Devise: ‚Hut ab vor dem Ehrenmann‘ (So Kurt May [...]), so schrumpft Meister Anton für den anderen zu einer moralischen Beschränktheit, die im ‚Zurechthobeln von Menschen zu Reputationszwecken‘ ihren kritisch-bedenklichen Nenner findet. (So L[u]dger Lütkehaus [...]).*“ (Reinhardt 1989, 223f.)
- Tag *Markierung eines Forschungsdesiderats:* Der Tag dient der Annotation von Textstellen, an denen die Verfasser*innen des Interpretationstextes auf ein Forschungsdesiderat hinweisen.
 - Länge der annotierten Passage: i. d. R. eine kurze Wortgruppe oder ein Satz.
 - Indikatoren: Formulierungen wie „Bisher unbeachtet blieb...“, „eine Untersuchung von Aspekt x steht noch aus“ usw.
 - Property: Der Tag kann zwei Werte annehmen („Primärtextforschung“ oder „andere Forschung“), je nachdem, ob ein Desiderat bezüglich der Forschung zum jeweils interpretierten Primärtext oder bezüglich anderer Forschungsfelder markiert wird.
 - Beispiel: „*Die Kafka-Forschung hat es bislang unterlassen, Kafkas Rückgriff auf Goethes ‚Ruhe‘ und symbolische Seh- und Erkenntnisweise bei der Gestaltung des Berichts [Kafkas Erzählung Ein Bericht für eine Akademie; S.D.] nachzuspüren.*“ (Guarda 2015, 226)
- Tag *Konkreter Forschungsbezug:* Dieser Tag kann verwendet werden, um explizite Nennungen von Forschungsliteratur zu annotieren, die in der Regel im Fließtext oder in den Fußnoten eines Interpretationstextes vorkommen. Annotator*innen sollten sich vorab darauf einigen, welche Fälle ausgeschlossen werden und ob z. B. Erwähnungen von Forschungsliteratur in Literaturverzeichnissen ebenfalls markiert werden sollen.
 - Länge der annotierten Passage: Hier gibt es unterschiedliche Möglichkeiten. Annotiert werden kann die vollständige Literaturangabe. Da jedoch erwähnte Forschungsbeiträge häufig nicht mit der vollständigen Angabe erwähnt werden, ist es empfehlenswert, die Namen der Verfasser*innen oder die Kurztitel des erwähnten Forschungsbeitrags zu annotieren. (Dabei ist es wichtig, nur einen Tag pro Forschungsbeitrag zu verwenden, um die Ergebnisse einer potenziellen quantitativen Auswertung nicht zu verfälschen.)
 - Indikatoren: Typische Literaturangaben, wie sie i. d. R. in Fußnoten, zuweilen auch im Fließtext vorkommen.
 - Property: Der Tag kann zwei Werte annehmen („Primärtextforschung“ oder „andere Forschung“), je nachdem, ob ein Forschungsbeitrag zum jeweils interpretierten Primärtext erwähnt wird oder zu anderen Forschungsfeldern.
 - Beispiel: „*Eine den poetologischen Charakter des Gedichts hervorkehrende Interpretation bietet Sigrid Weigel, wenn sie [Ingeborg Bachmanns Gedicht; S.D.] *Dunkles zu sagen* als Behandlung der Frage nach dem Ursprung der Kunst aus dem Klagelied liest und feststellt, daß ‚Wissen um die Unumkehrbarkeit des Sterbens (...) Voraus-Setzung des Vergleichs zwischen Dichtung und Orpheusscher Kunst‘ ist. Vgl. WEIGEL (1999), S. 135-142. Hier: S. 140.*“ (von Jagow 2003, 44)

Externe und weiterführende Links

- Tagset „Interpretationstexte analysieren“ auf Zenodo: <https://web.archive.org/save/https://zenodo.org/records/10519632> (Letzter Zugriff: 03.07.2024)

Bibliographie

- Albrecht, Andrea, Lutz Danneberg, Olaf Krämer und Carlos Spoerhase, Hrsg. 2015. *Theorien, Methoden und Praktiken des Interpretierens*. Berlin, Boston: de Gruyter.
- Conterno, Chiara. 2014. *Die andere Tradition Psalm-Gedichte im 20. Jahrhundert*. Göttingen: V & R unipress.
- Descher, Stefan, Jan Borkowski, Felicitas Ferder und Philipp David Heine. 2017. Probleme der Interpretation von Literatur. Ein Überblick. In: *Literatur interpretieren: Interdisziplinäre Beiträge zur Theorie und Praxis*, hg. von Jan Borkowski, Stefan Descher, Felicitas Ferder, und Philipp David Heine, 11–70. Münster: mentis.
- Descher, Stefan und Thomas Petraschka. 2019. *Argumentieren in der Literaturwissenschaft. Eine Einführung*. Ditzingen: Reclam Verlag.
- Ertel, Anna und Tilman Köppe. 2017. Meta-Fantastik – Ulrike Draesners Erzählung Rosakäfer. In: *Funktionen der Fantastik. Neue Formen des Weltbezugs von Literatur und Film nach 1945*, hg. von Sonja Klimek, Tobias Lambrecht, und Tom Kindt, 191–208. Heidelberg: Universitätsverlag Winter.
- forTEXT. 2020. Tagset Interpretationstexte analysieren. Zenodo, 27. Juli. doi: 10.5281/zenodo.10519632, <https://doi.org/10.5281/zenodo.10519632>.
- Frey, Christiane. 2003. Spiegelfechtereien mit dem Leser. Trügerische Ökonomien der Schrift in Kleists Michael Kohlhaas. In: *Beiträge zur Kleist-Forschung*, hg. von Lothar Jordan, 17:296–317. Stiftung Kleist-Museum.
- Grewendorf, Günter. 1975. *Argumentation und Interpretation. Wissenschaftstheoretische Untersuchungen am Beispiel germanistischer Lyrikinterpretationen*. Kronberg (Ts): Scriptor Verlag.
- Guarda, Sylvain. 2015. Kafkas Akademiebericht: Die auflösende ‚Ruhe‘ als lebendige Varieténnummer. *Seminar: A Journal of Germanic Studies* 51, Nr. 3: 225–241.
- Jacke, Janina. 2024. Methodenbeitrag: Manuelle Annotation. *forTEXT Heft 1*, Nr. 4. Manuelle Annotation (7. August). doi: 10.48694/fortext.3748, <https://fortext.net/routinen/methoden/manuelle-annotation>.
- Knott, Alistair und Robert Dale. 1993. Using Linguistic Phenomena to Motivate a Set of Rhetorical Relations. *Discourse Processes* 18, Nr. 1: 1–40. doi: 10.1080/01638539409544883.
- Kocher, Ursula. 2017. Krankheit aus der Distanz. Alfred Döblins frühe Erzählungen als narrative Notate krankhafter Existenzen. *DIEGESIS. Interdisziplinäres E-Journal für Erzählforschung / Interdisciplinary E-Journal for Narrative Research* 6, Nr. 2: 91–106. <http://elpub.bib.uni-wuppertal.de/servlets/DerivateServlet/Derivate-7288/dej17060206.pdf> (zugegriffen: 1. Juli 2020).
- Kuhn, Anna Katharina. 1981. *Der Dialog bei Frank Wedekind. Untersuchungen zum Szenengespräch der Dramen bis 1900*. Heidelberg: Winter.
- Mochales Palau, Raquel und Marie-Francine Moens. 2009. Argumentation Mining: The Detection, Classification and Structure of Arguments in Text. In: *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, 98–109. doi: 10.1145/1568234.1568246.
- Reinhardt, Hartmut. 1989. *Apologie der Tragödie. Studien zur Dramatik Friedrich Hebbels*. Tübingen: Max Niemeyer Verlag.
- Schubarth, Caroline. 2011. Der Ausschluss aus der Gemeinschaft. Gewalt und Ethik in Kleists Michael Kohlhaas. In: *Kleist. Relektüren*, hg. von Branka Schaller-Fornoff, 45–62. Dresden: Thelem.
- Schumacher, Mareike. 2024. Toolbeitrag: CATMA. *forTEXT Heft 1*, Nr. 4. Manuelle Annotation (7. August). doi: 10.48694/fortext.3761, <https://fortext.net/tools/tools/catma>.
- Sengle, Friedrich. 1949. *Wieland. Mit 23 montierten Bildern und Beilagen*. Stuttgart: Metzler.
- Sng, Zachary. 2007. Das Fehlläuten der Nachtglocke. Zu Kafkas Erzählung Ein Landarzt. In: *Kafkas Institutionen*, hg. von Arne Höcker und Oliver Simons, 213–233. Bielefeld: transcript Verlag.
- Stab, Christian und Iryna Gurevych. 2014. Identifying Argumentative Discourse Structures in Persuasive Essays. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 46–56. doi: 0.3115/v1/D14-1006.
- Toulmin, Stephen. 2003. *The Uses of Argument*. Updated Edition. Cambridge: Cambridge University Press.
- van Eemeren, Frans H. und Rob Grootendorst. 2004. *A Systematic Theory of Argumentation. The Pragma-Dialectical Approach*. Cambridge (u.a.): Cambridge University Press.
- von Jagow, Bettina. 2003. *Ästhetik des Mythischen. Poetologien des Erinnerns im Werk von Ingeborg Bachmann*. Köln (u.a.): Böhlau Verlag.
- von Savigny, Eike. 1976. *Argumentation in der Literaturwissenschaft. Wissenschaftstheoretische Untersuchungen zu Lyrikinterpretationen*. München: Beck.

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.
- 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.
- 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 unbekanntem Daten verwendet werden.
- Markup (Textauszeichnung)** Die Textauszeichnung (eng. *Markup*) fällt in den Bereich der Daten- bzw. Textverarbeitung, genauer in das Gebiet der Textformatierung, welche durch **Auszeichnungssprachen** wie **XML** implementiert wird. Dabei geht es um die Beschreibung, wie einzelne Elemente eines Textes beispielsweise auf Webseiten grafisch dargestellt werden sollen.
- 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.
- 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**.
- Tagset** Ein Tagset definiert die Taxonomie, anhand derer **Annotationen** in einem Projekt erstellt werden. Ein Tagset beinhaltet immer mehrere Tags und ggf. auch Subtags. Ähnlich der **Type/Token**-Differenz in der Linguistik sind Tags deskriptive Kategorien, wohingegen Annotationen die einzelnen Vorkommnisse dieser Kategorien im Text sind.
- TEI** Die *Text Encoding Initiative* (TEI) ist ein Konsortium, das gemeinsam einen Standard für die Darstellung von Texten in digitaler Form entwickelt. Die TEI bietet beispielsweise Standards zur Kodierung von gedruckten Werken und zur Auszeichnung von sprachlichen Informationen in maschinenlesbaren Texten (siehe auch **XML** und **Markup**).
- 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.

XML XML steht für *Extensible Markup Language* und ist eine Form von **Markup Language**, die sowohl computer- als auch menschenlesbar und hochgradig anpassbar ist. Dabei werden Textdateien hierarchisch strukturiert dargestellt und Zusatzinformationen i. d. R. in einer anderen Farbe als der eigentliche (schwarz gedruckte) Text dargestellt. Eine standardisierte Form von XML ist das **TEI-XML**.