

Kollaboratives Annotieren literarischer Texte

,
janina.jacke@uni-hamburg.de
Universität Hamburg, Deutschland

,
evelyn.gius@uni-hamburg.de
Universität Hamburg, Deutschland

Kollaboratives Annotieren ist eine gute Möglichkeit, um mehr als bloß eine subjektive Perspektive auf den Untersuchungsgegenstand – beispielsweise einen Text – abzubilden: Sobald mehr als ein Individuum Markup an einem digitalen Objekt anbringt, kann deutlich werden, welche unterschiedlichen Aspekte des Objekts im Zentrum des individuellen Interesses stehen oder welche verschiedenen Sichtweisen in Bezug auf denselben Aspekt möglich sind. Soll die kollaborative Annotation jedoch nicht bloß die Pluralität von Perspektiven und Meinungen aufzeigen, sondern einem spezifischeren Erkenntnisinteresse dienen, so sollte die Annotationspraxis mithilfe von Guidelines strukturiert und reguliert werden. Für die Annotation linguistischer Phänomene (bspw. in Gebrauchstexten) werden solche bereits entwickelt (vgl. bspw. Pyysalo / Ginter 2014; Mamouri et al. 2008); dagegen existieren für die kollaborative Annotation semantischer Phänomene in literarischen Texten kaum *best practice*-Vorschläge. Eine direkte Übertragung linguistischer Annotationsanleitungen auf den literaturwissenschaftlichen Bereich scheint dabei aus mindestens drei Gründen nicht möglich:

- Literarische Texte sind in der Regel polyvalent, d. h. mehrdeutig: Während kollaboratives Annotieren im Bereich der Linguistik letztlich der Vermeidung individueller Annotationsfehler dient und eine autoritative Version zum Ziel hat, muss im Falle literaturwissenschaftlicher Annotation immer bedacht werden, dass auch unterschiedliche bzw. widersprüchliche Annotationen gleichermaßen legitime Lesarten ausdrücken können – ohne dabei jedoch in Beliebigkeit abzugleiten.
- Die Analysekategorien, mithilfe derer literarische Texte untersucht bzw. annotiert werden, sind häufig unterdefiniert. Grund hierfür scheint zum einen die Tatsache zu sein, dass Textanalyse und -interpretation in der traditionell-literaturwissenschaftlichen Praxis häufig weniger textnah ausgeführt werden als im digitalen Kontext, was dazu führt, dass vage Definitionen keine Anwendungsprobleme mit sich bringen. Zum anderen zeichnet sich die traditionelle Literaturwissenschaft durch individuelles Arbeiten und die Erschließung neuer Lesarten von Texten

aus. Dadurch bleibt häufig unbemerkt, dass die Analysekategorien nicht konsistent verwendet werden. Im Kontext der kollaborativen Annotation literarischer Texte muss also die Frage berücksichtigt werden, wie spezifisch und klar die Definition literaturwissenschaftlicher Annotationskategorien ausfallen kann und sollte.

- Literaturwissenschaftliche Analysekategorien stehen häufig in bisher untertheoretisierten Abhängigkeitsverhältnissen zueinander. Das kann dazu führen, dass Annotationen deshalb unterschiedlich ausfallen, weil die Annotatoren im Kontext impliziter Analyseschritte, die der Annotation logisch vorgeordnet sind, zu unterschiedlichen Ergebnissen gekommen sind. Auch für den Umgang mit solchen impliziten Abhängigkeitsverhältnissen muss im Kontext literarischer Annotation eine Regelung gefunden werden.

Im Folgenden möchten wir einen *best practice* Vorschlag für das kollaborative Annotieren literarischer Texte vorstellen, der die oben genannten Schwierigkeiten berücksichtigt. Da *best practices* – besonders im literaturwissenschaftlichen Bereich – allerdings nicht vollständig unabhängig vom zugrundeliegenden Erkenntnisinteresse sind, stellen wir unserem *How to* eine kurze Beschreibung des Projektkontextes voran, im Rahmen dessen die Guidelines entstanden sind. Mit kleineren, am jeweiligen Erkenntnisinteresse orientierten Modifikationen sollte diese Anleitung allerdings problemlos auf anders ausgerichtete Projekte übertragbar sein.

Die Annotationsguidelines sind im Kontext des Projekts heureCLÉA entstanden (vgl. Bögel et al. im Erscheinen). Ziel des Projekts ist die Entwicklung einer digitalen Heuristik, d. h. eines Funktionsmoduls, das automatisch semantische Phänomene in literarischen Texten – hier: narratologische Phänomene der Zeitgestaltung – annotiert. Für die Generierung dieses Tools wurde zunächst ein Korpus literarischer Texte kollaborativ in Bezug auf die zu automatisierenden Phänomene annotiert. Basierend auf diesen Annotationen soll die Funktionalität dann mithilfe regelbasierter Verfahren und Machine Learning-Methoden entwickelt werden.

Nach einigen Anläufen hat sich folgende Annotationspraxis als *best practice* herausgestellt (vgl. Abbildung 1):

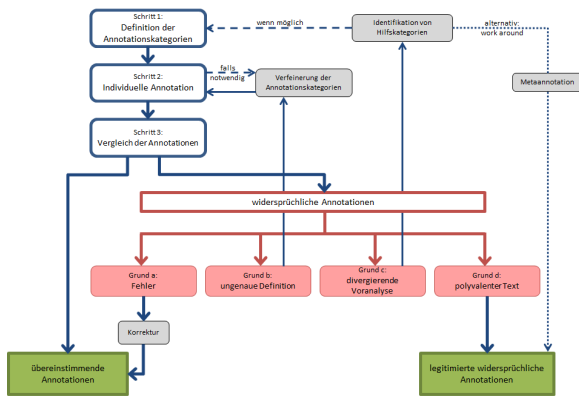


Abb. 1: Ablaufschema kollaborative literaturwissenschaftliche Annotation

- **Schritt 1: Möglichst genaue Definition der Annotationskategorien.** Es ist sinnvoll, die zugrundeliegenden Annotationskategorien bereits vor der Annotation so spezifisch wie möglich zu definieren. Geschieht dies nicht, so ist es letztlich schwieriger festzustellen, ob sich Annotationen tatsächlich auf Basis der Polyvalenz des literarischen Textes unterscheiden oder aufgrund vager Kategoriendefinitionen. Wenn die Definition einer traditionellen literaturwissenschaftlichen Kategorie auf zwei unterschiedliche Arten verstanden werden kann und beide Varianten interessante Textphänomene beschreiben, dann ist es problemlos möglich, beide Varianten zu operationalisieren – allerdings als sorgfältig getrennte Konzepte. Die Forderung der möglichst genauen Definition geht also nicht notwendigerweise mit einer Beschränkung des Erkenntnisinteresses einher. Zusätzlich zur Definition der Annotationskategorie sollten auch die Textoberflächenindikatoren festgehalten werden, die auf das jeweilige Phänomen hinweisen, sowie die Granularität der annotierten Zeichenkette (Fragment, Wort, Teilsatz, Satz etc.). Die Festlegungen werden in Form von Annotationsguidelines dokumentiert, die allen Annotatoren als Annotationsbasis dienen.
- **Schritt 2: Individuelle Annotation.** Jeder Annotator annotiert auf Basis der Guidelines die Korpustexte individuell. Dadurch soll gewährleistet werden, dass nicht gleich zu Beginn eine gegenseitige Beeinflussung der Annotatoren besteht, sondern unterschiedliche Lesarten auch tatsächlich in den Annotationen abgebildet werden.
- *falls notwendig: Verfeinerung der Annotationskategorien.* Eventuell stellt sich bereits in der individuellen Annotationsrunde heraus, dass einige Annotationskategorien noch nicht spezifisch genug definiert sind, um eine regelgeleitete Anwendung zu gewährleisten. Ist dies der Fall, so müssen die Kategorien spezifiziert werden, damit die individuelle Annotationsrunde sinnvoll beendet werden kann.

- Schritt 2: *Individuelle Annotation*. Jeder Annotator annotiert auf Basis der Guidelines die Korpustexte individuell. Dadurch soll gewährleistet werden, dass nicht gleich zu Beginn eine gegenseitige Beeinflussung der Annotatoren besteht, sondern unterschiedliche Lesarten auch tatsächlich in den Annotationen abgebildet werden.
- *falls notwendig: Verfeinerung der Annotationskategorien*. Eventuell stellt sich bereits in der individuellen Annotationsrunde heraus, dass einige Annotationskategorien noch nicht spezifisch genug definiert sind, um eine regelgeleitete Anwendung zu gewährleisten. Ist dies der Fall, so müssen die Kategorien spezifiziert werden, damit die individuelle Annotationsrunde sinnvoll beendet werden kann.

- *falls notwendig: Verfeinerung der Annotationskategorien.* Eventuell stellt sich bereits in der individuellen Annotationsrunde heraus, dass einige Annotationskategorien noch nicht spezifisch genug definiert sind, um eine regelgeleitete Anwendung zu gewährleisten. Ist dies der Fall, so müssen die Kategorien spezifiziert werden, damit die individuelle Annotationsrunde sinnvoll beendet werden kann.

- **Schritt 3: Vergleich der Annotationen.** Nach Abschluss der individuellen Annotation vergleichen die Annotatoren ihre Arbeit. Im Fall diskrepanter Annotationen werden die jeweiligen Gründe der Annotationsentscheidung diskutiert. Generell können vier Arten von Gründen für widersprüchliche Annotationen auftreten, die unterschiedliche Maßnahmen erfordern:
 - *falsche Annotation:* Eine der Annotationen basiert auf einem eindeutig falschen Verständnis der fraglichen Textstelle oder der Kategoriedefinition. In diesem Fall muss die falsche Annotation korrigiert werden.
 - *ungenau Definition:* Es ist möglich, dass der Vergleich der Annotationen weitere Defizite der Kategoriedefinition offenbart, die im Kontext vorheriger Stadien nicht deutlich geworden sind. Ist dies der Fall, muss die Definition (ein weiteres Mal) überarbeitet werden. Es folgt eine weitere individuelle Annotationsphase, gefolgt von einem Vergleich der Annotationen.
 - *divergierende Voranalyse:* Ein Vergleich der diskrepanten Annotationen kann ergeben, dass die Anwendung bestimmter Annotationskategorien vorbereitende Analyseschritte notwendig macht. Diese Schritte sind von den Annotatoren jeweils implizit ausgeführt worden. Wenn diese Analysen zu unterschiedlichen Ergebnissen führen, dann können auch die darauf aufbauenden Annotationen variieren. Ist dies der Fall, so müssen die für die vorbereitenden Analyseschritte notwendigen Hilfskategorien identifiziert und definiert werden. Wenn der Projektrahmen es zulässt, dann sollte das gesamte Textkorpus auch unter Rückgriff auf die Hilfskategorien nach dem hier dargestellten Ablaufschema annotiert werden. Die Hilfsannotationen können dann ebenfalls auf ihre Richtigkeit hin überprüft werden, wodurch die Hauptannotationen in der Regel deutlich einheitlicher ausfallen. Ist es aus arbeitsökonomischen Gründen nicht möglich, eine Voranalyse des Korpus anhand von Hilfskategorien durchzuführen, so ist ein *work around* möglich: Anstatt den Versuch zu unternehmen, die vorbereitenden Analysen zu vereinheitlichen, können die Annotatoren ihre Hauptannotationen mit Metaannotationen versehen, die die Ergebnisse der Voranalysen festhalten. Auf diese Weise wird nicht die Frage nach der Korrektheit der Voranalysen gestellt, wohl aber der Grund für die divergierenden Hauptannotationen explizit gemacht.
 - *polyvalenter Text:* Wird als Grund für eine widersprüchliche Annotation textuelle Mehrdeutigkeit herausgestellt, so müssen die Annotationen nicht überarbeitet werden, sondern die widersprüchlichen Annotationen werden als legitimiert verstanden.

- *falsche Annotation*: Eine der Annotationen basiert auf einem eindeutig falschen Verständnis der fraglichen Textstelle oder der Kategoriedefinition. In diesem Fall muss die falsche Annotation korrigiert werden.
- *ungenau Definition*: Es ist möglich, dass der Vergleich der Annotationen weitere Defizite der Kategoriedefinition offenbart, die im Kontext vorheriger Stadien nicht deutlich geworden sind. Ist dies der Fall, muss die Definition (ein weiteres Mal) überarbeitet werden. Es folgt eine weitere individuelle Annotationsphase, gefolgt von einem Vergleich der Annotationen.
- *divergierende Voranalyse*: Ein Vergleich der diskrepanten Annotationen kann ergeben, dass die Anwendung bestimmter Annotationskategorien vorbereitende Analyseschritte notwendig macht. Diese Schritte sind von den Annotatoren jeweils implizit ausgeführt worden. Wenn diese Analysen zu unterschiedlichen Ergebnissen führen, dann können auch die darauf aufbauenden Annotationen variieren. Ist dies der Fall, so müssen die für die vorbereitenden Analyseschritte notwendigen Hilfskategorien identifiziert und definiert werden. Wenn der Projektrahmen es zulässt, dann sollte das gesamte Textkorpus auch unter Rückgriff auf die Hilfskategorien nach dem hier dargestellten Ablaufschema annotiert werden. Die Hilfsannotationen können dann ebenfalls auf ihre Richtigkeit hin überprüft werden, wodurch die Hauptannotationen in der Regel deutlich einheitlicher ausfallen. Ist es aus arbeitsökonomischen Gründen nicht möglich, eine Voranalyse des Korpus anhand von Hilfskategorien durchzuführen, so ist ein *work around* möglich: Anstatt den Versuch zu unternehmen, die vorbereitenden Analysen zu vereinheitlichen, können die Annotatoren ihre Hauptannotationen mit Metaannotationen versehen, die die Ergebnisse der Voranalysen festhalten. Auf diese Weise wird nicht die Frage nach der Korrektheit der Voranalysen gestellt, wohl aber der Grund für die divergierenden Hauptannotationen explizit gemacht.
- *polyvalenter Text*: Wird als Grund für eine widersprüchliche Annotation textuelle Mehrdeutigkeit herausgestellt, so müssen die Annotationen nicht überarbeitet werden, sondern die widersprüchlichen Annotationen werden als legitimiert verstanden.

- *ungenauae Definition:* Es ist möglich, dass der Vergleich der Annotationen weitere Defizite der Kategoriedefinition offenbart, die im Kontext vorheriger Stadien nicht deutlich geworden sind. Ist dies der Fall, muss die Definition (ein weiteres Mal) überarbeitet werden. Es folgt eine weitere individuelle Annotationsphase, gefolgt von einem Vergleich der Annotationen.
- *divergierende Voranalyse:* Ein Vergleich der diskrepanten Annotationen kann ergeben, dass die Anwendung bestimmter Annotationskategorien vorbereitende Analyseschritte notwendig macht. Diese Schritte sind von den Annotatoren jeweils implizit ausgeführt worden. Wenn diese Analysen zu unterschiedlichen Ergebnissen führen, dann können auch die darauf aufbauenden Annotationen variieren. Ist dies der Fall, so müssen die für die vorbereitenden Analyseschritte notwendigen Hilfskategorien identifiziert und definiert werden. Wenn der Projektrahmen es zulässt, dann sollte das gesamte Textkorpus auch unter Rückgriff auf die Hilfskategorien nach dem hier dargestellten Ablaufschema annotiert werden. Die Hilfsannotationen können dann ebenfalls auf ihre Richtigkeit hin überprüft werden, wodurch die Hauptannotationen in der Regel deutlich einheitlicher ausfallen. Ist es aus arbeitsökonomischen Gründen nicht möglich, eine Voranalyse des Korpus anhand von Hilfskategorien durchzuführen, so ist ein *work around* möglich: Anstatt den Versuch zu unternehmen, die vorbereitenden Analysen zu vereinheitlichen, können die Annotatoren ihre Hauptannotationen mit Metaannotationen versehen, die die Ergebnisse der Voranalysen festhalten. Auf diese Weise wird nicht die Frage nach der Korrektheit der Voranalysen gestellt, wohl aber der Grund für die divergierenden Hauptannotationen explizit gemacht.

- *divergierende Voranalyse*: Ein Vergleich der diskrepanten Annotationen kann ergeben, dass die Anwendung bestimmter Annotationskategorien vorbereitende Analyseschritte notwendig macht. Diese Schritte sind von den Annotatoren jeweils implizit ausgeführt worden. Wenn diese Analysen zu unterschiedlichen Ergebnissen führen, dann können auch die darauf aufbauenden Annotationen variieren. Ist dies der Fall, so müssen die für die vorbereitenden Analyseschritte notwendigen Hilfskategorien identifiziert und definiert werden. Wenn der Projektrahmen es zulässt, dann sollte das gesamte Textkorpus auch unter Rückgriff auf die Hilfskategorien nach dem hier dargestellten Ablaufschema annotiert werden. Die Hilfsannotationen können dann ebenfalls auf ihre Richtigkeit hin überprüft werden, wodurch die Hauptannotationen in der Regel deutlich einheitlicher ausfallen. Ist es aus arbeitsökonomischen Gründen nicht möglich, eine Voranalyse des Korpus anhand von Hilfskategorien durchzuführen, so ist ein *work around* möglich: Anstatt den Versuch zu unternehmen, die vorbereitenden Analysen zu vereinheitlichen, können die Annotatoren ihre Hauptannotationen mit Metaannotationen versehen, die die Ergebnisse der Voranalysen festhalten. Auf diese Weise wird nicht die Frage nach der Korrektheit der Voranalysen gestellt, wohl aber der Grund für die divergierenden Hauptannotationen explizit gemacht.

- *polyvalenter Text*: Wird als Grund für eine widersprüchliche Annotation textuelle Mehrdeutigkeit herausgestellt, so müssen die Annotationen nicht überarbeitet werden, sondern die widersprüchlichen Annotationen werden als legitimiert verstanden.

Wie deutlich geworden ist, werden die eingangs genannten drei Probleme literarischer Annotation im Kontext dieses Ablaufschemas berücksichtigt: Die Möglichkeit, einen literarischen Text unterschiedlich zu deuten, wird zum einen durch die individuelle Annotationsphase (vgl. Schritt 2) gewährleistet, zum anderen dadurch, dass widersprüchliche Annotationen erlaubt sind, sofern sie durch die Polyvalenz des Textes bedingt sind (vgl. Grund d). Dass die Berücksichtigung der Polyvalenz nicht in eine Beliebigkeit von Annotationsentscheidungen abgeleitet, wird dadurch erreicht, dass andere Gründe (d. h. mindestens Gründe a und b) für widersprüchliche Annotationen ausgeschlossen werden. Die Spezifikation literaturwissenschaftlicher Annotationskategorien wird schrittweise optimiert, um aussagekräftige Annotationsergebnisse zu gewährleisten (vgl. Schritt 1 sowie ggf. weitere Optimierungsschritte ausgehend von Schritt 2 und Grund b). Da klare Definiertheit nicht notwendig mit einer Einschränkung der Perspektive auf Texte einhergeht (vgl. Schritt 1), ist sie mit der literaturwissenschaftlichen Praxis bzw. mit der Pluralität möglicher Erkenntnisinteressen kompatibel. Die Abhängigkeit literaturwissenschaftlicher Kategorien untereinander wird schließlich, je nach verfügbaren Ressourcen, entweder in einem reduzierten oder in einem ausführlichen Ansatz explizit gemacht (vgl. Grund c).

Fußnoten

1. Die zentralen zeitlichen Phänomene in heureCLÉA beziehen sich auf die temporale Relation zwischen erzählter Geschichte und ihrer Repräsentation. Diese Relation kann in dreierlei Hinsicht untersucht werden: *Ordnung* bzw. Reihenfolge (Wann findet ein Ereignis statt? – Wann wird es erzählt?), *Frequenz* bzw. Häufigkeit (Wie oft findet es statt? – Wie oft wird es erzählt?) und *Dauer* bzw. Geschwindigkeit (Wie lange dauert es? – Wie lange dauert es, davon zu erzählen?) (vgl. Genette 1994).

2. Ein Beispiel für eine solche Hilfskategorie in heureCLÉA ist die der *Erzählebenen* (Finden innerhalb einer Erzählung weitere eingebettete Erzählungen statt?, vgl. Ryan 1991). Die Identifikation von Erzählebenen beeinflusst die Analyse zeitlicher Phänomene wie *Dauer*: Wo immer eine eingebettete Erzählung auftaucht, kann die Erzähldauer entweder auf der Ebene der Haupterzählung oder auf der der eingebetteten Erzählung analysiert werden. In heureCLÉA wurden deswegen narrative Ebenen vor der finalen Annotation von Dauer annotiert. Ein *work around* hätte folgendermaßen aussehen können: Wann immer die Diskussion widersprüchlicher Annotationen ergibt, dass einer der Anntatoren eine eingebettete Erzählung identifiziert hat und der andere nicht bzw. dass beide Annotatoren eine eingebettete Erzählungen identifiziert haben, sich ihre Dauerannotationen sich jedoch auf unterschiedliche Ebenen beziehen, wird dies in Form einer Metaannotation festgehalten. – Das Ergebnis der

oben genannten Arbeitsschritte in heureCLÉA – die heureCLÉA Annotationsguidelines Version 1.0 – sind unter www.heureCLÉA.de/guidelines einsehbar.

Bibliographie

Bögel, Thomas / Gertz, Michael / Gius, Evelyn / Jacke, Janina / Meister, Jan Christoph / Petris, Marco / Strötgen, Jannik (im Erscheinen): "Collaborative Text Annotation Meets Machine Learning. heureCLÉA, a Digital Heuristic of Narrative", in: *DHCommons* 1.

Genette, Gérard (1994): *Die Erzählung*. München: Fink.

Gius, Evelyn / Jacke, Janina (2015): *Zur Annotation narratologischer Kategorien der Zeit*. Guidelines zur Nutzung des CATMA-Tagsets (Version 1.0) www.heureclea.de/guidelines [letzter Zugriff 11. Oktober 2015].

Maamouri, Mohamed / Bies, Ann / Kulick, Seth (2008): "Enhancing the Arabic Treebank. A Collaborative Effort toward New Annotation Guidelines", in: *7th International Conference on Language Resources and Evaluation*. Marrakech (LREC 2008) https://catalog.ldc.upenn.edu/docs/LDC2010T13/Enhancing_Arabic_Treebank.pdf [letzter Zugriff 11. Oktober 2015].

Pyysalo, Sampo / Ginter, Filip (2014): "Collaborative development of annotation guidelines with application to Universal Dependencies", in: *Fifth Swedish Language Technology Conference*. Uppsala (SLTC 2014) http://www2.lingfil.uu.se/SLTC2014/abstracts/sltc2014_submission_32.pdf [letzter Zugriff 11. Oktober 2015].

Ryan, Marie-Laure (1991): *Possible Worlds, Artificial Intelligence, and Narrative Theory*. Bloomington, Ind.: Indiana University Press.