GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

GCDH GÖTTINGEN CENTRE FOR DIGITAL HUMANITIES

Reflexive Passagen und ihre Attribution

Hanna Varachkina*, Florian Barth+, Luisa Gödeke*, Anna Mareike Weimer*, Tillmann Dönicke+

*Seminar für Deutsche Philologie, +Göttingen Centre for Digital Humanities

Definitionen

Kommentare – Textstellen, in denen die erzählte Zeit unterbrochen wird und die Erzählung, Figuren, Handlung oder der Akt des Erzählens kommentiert wird. (2), (3), (5)

Nicht-fiktionale Rede – Passagen in fiktionalen Texten, die Behauptungen bzw. Hypothesen über die reale Welt nahelegen. (4), (5)

Generalisierungen – quantifizierte Aussagen über angenommene Instanzen einer Klasse oder Gruppe von Objekten, Individuen oder (Zeit-)Räumen verstanden, auf die nicht kontextuell referiert wird. (1), (5)

Nicht-fiktionale Rede, Kommentare und Generalisierungen werden mit reflexiven Passagen assoziert.

Attribution bestimmt, wem die in der Passage enthaltene Information zugeschrieben werden kann: einer Figur, der Erzählinstanz und/oder der Autorin. Unterschiedliche Interpretationen (Zuschreibungen) sind dabei möglich. Aus diesem Grund wird für Attribution kein Goldstandard erstellt.

Beispiele

(1) [ach wie ich oft da saß mit klopfendem Herzen, auf den Höhen von Tina, und den Falken und Kranichen nachsah, und den kühnen fröhlichen Schiffen]$_{\text{GI-DIV - ERZÄHLINSTANZ: alle AnnotatorInnen}}$, [wenn sie hinunterschwanden am Horizont]$_{\text{GI-BARE - ERZÄHLINSTANZ: alle AnnotatorInnen}}$! (Hölderlin)

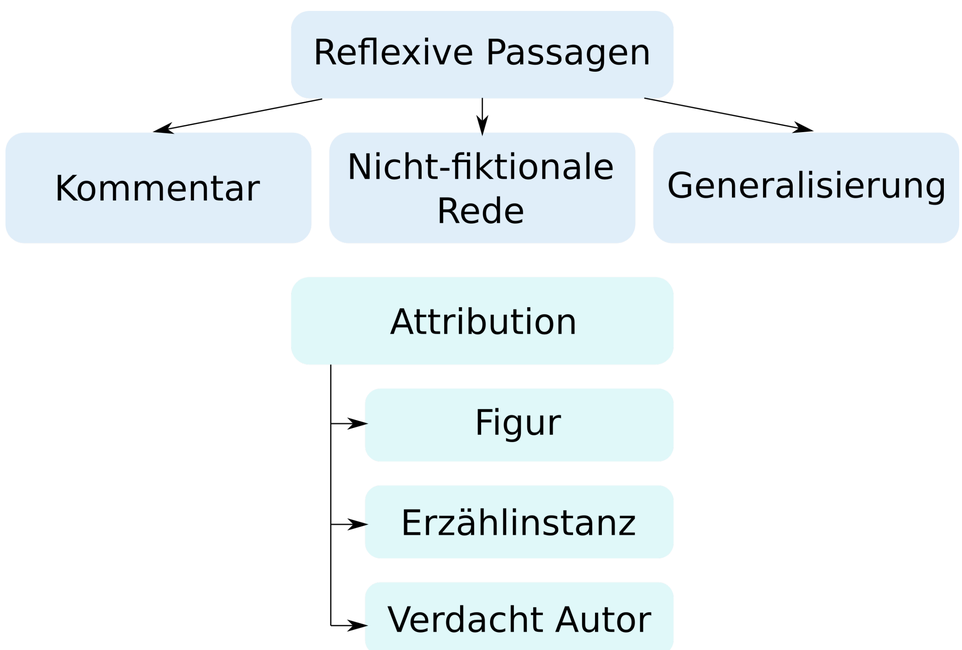
(2) [Ich hingegen war der Ansicht, daß man grad nur einer solchen See sich anvertrauen müsse]$_{\text{COMMENT-EINSTELLUNG - ERZÄHLINSTANZ: alle AnnotatorInnen}}$. (May)

(3) [Ein Mährchen will ich dir erzählen, horche wohl]$_{\text{COMMENT-META - FIGUR: 4/6 - ERZÄHLINSTANZ: 2/6}}$. (Novalis)

(4) [Es hatte denselben nüchternen Charakter wie fast alles, [was unter dem Soldatenkönig entstand]$_{\text{NICHTFIKTIONAL}}$]$_{\text{NICHTFIKTIONAL mit KONTEXT ERZÄHLINSTANZ: 4/6 - ERZÄHLINSTANZ + VERDACHT AUTOR: 2/6}}$, [...] (Fontane)

(5) [...][[[und augenblickliche Hülfe ist doch immer das, was auf dem Lande am meisten vermißt wird]$_{\text{GI-ALL}}$]$_{\text{COMMENT-INTERPRETATION}}$]$_{\text{NICHTFIKTIONAL - FIGUR: 3/6 - FIGUR + VERDACHT AUTOR: 3/6}}$. (Goethe)

Model



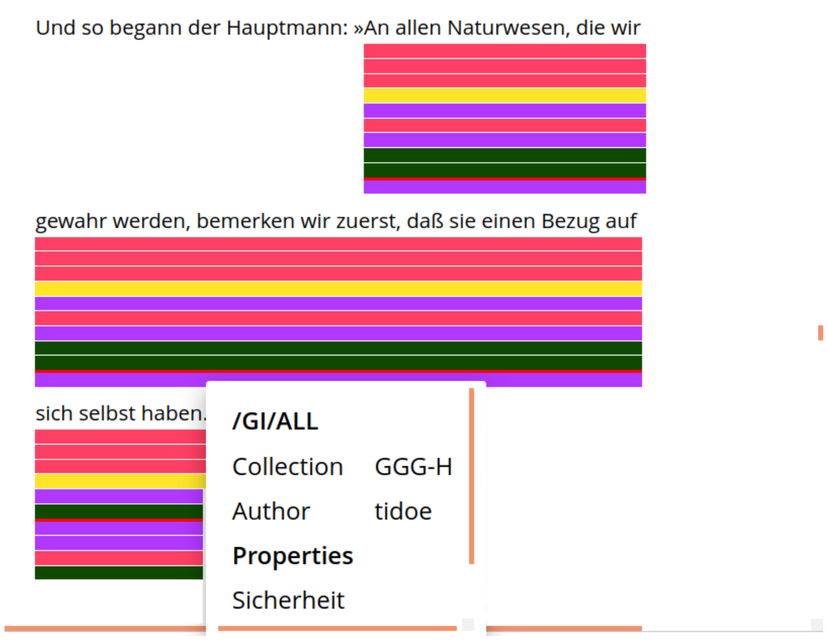
MONACO

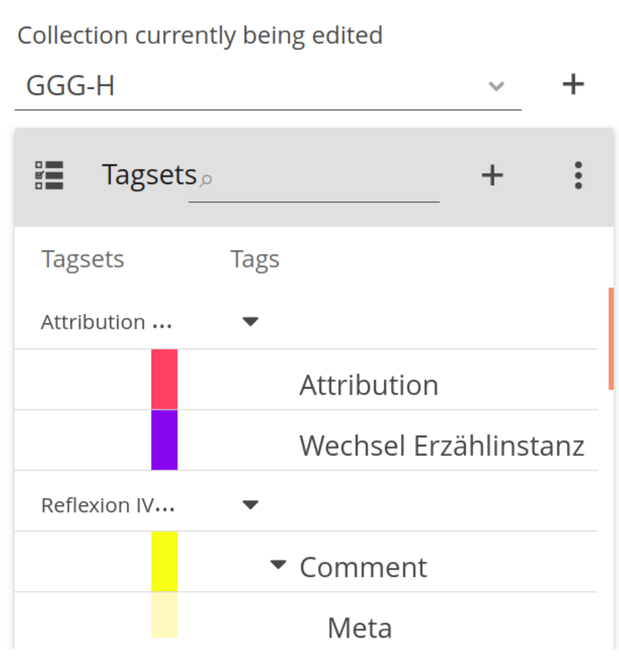
https://gitlab.gwdg.de/mona/korpus-public

Deutschsprachige fiktionale Erzähltexte von 1650–1950 (200 Anfangssätze)

Andreae *Die chymische Hochzeit* 1616
Dahn *Kampf um Rom* (ab Kapitel 2) 1876
Fontane *Der Stechlin* 1898
Gellert *Das Leben der schwedischen Gräfin von G* 1747
Goethe *Die Wahlverwandtschaften* 1809
Grimmelshausen *Der abenteuerliche Simplicissimus* 1668
Hoffmann *Der Sandmann* 1816
Hölderlin *Hyperion oder der Eremit in Griechenland* 1816
Kafka *Der Bau* 1931
Kleist *Michael Kohlhaas* 1810
LaRoche *Geschichte des Fräuleins von Sternheim* 1771
Mann *Der Zauberberg* 1924
May *Winnetou II* 1893
Musil *Der Mann ohne Eigenschaften* 1930
Novalis *Die Lehrlinge zu Sais* 1802
Schnabel *Die Insel Felsenburg* 1731
Seghers *Das siebte Kreuz* 1942
Wieland *Geschichte des Agathon* 1766
Zesen *Adriatische Rosemund* 1645

Annotationen in CATMA



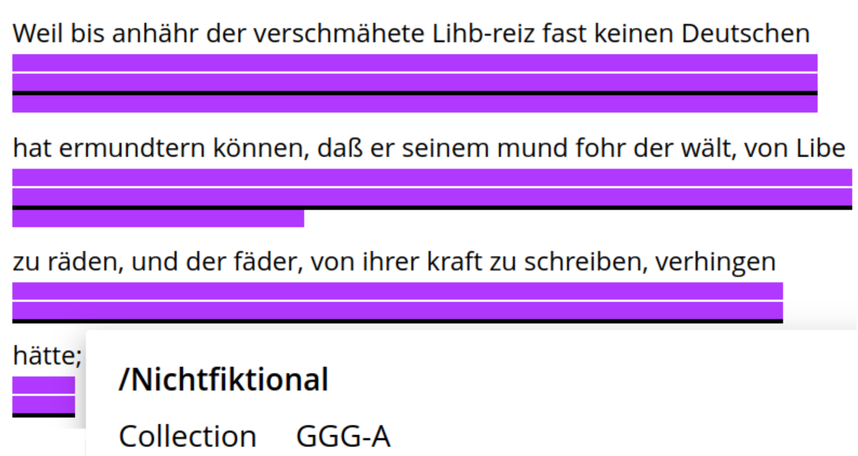


CATMA Merger

Automatische Vorerstellung des Goldstandards

Python Implementierung auf Grundlage der exportierten xml Dateien

GGG-A – **Vereinigungsmenge** der Annotationen



Inter-Annotator Agreement

Phänomen	Fleiss' κ M.	Fleiss' κ B.	Krippendorff's α M.	Krippendorff's α B.	Mathet's γ M.	Mathet's γ B.
Kommentar	0.43	0.46	0.44	0.46	0.56	0.58
Nicht-fiktionale Rede	0.63	0.68	0.64	0.68	0.72	0.78
Generalisierung	0.59	0.61	0.6	0.62	0.65	0.69
Attribution	0.46	0.89	0.43	0.43	nan	nan

Tab. 1: Multi-Label (M.) und binäres (B.) Agreement für die annotierten Phänomene

Klassifizierer

Phänomen	Train/Dev	Test	Tags	Methode	Mikro	Makro	Binär	Veröffentlichung
Kommentar	15/2	2	3	DT	–	0.38	0.54	Weimer et al. (2022, eingereicht)
Kommentar	15/2	2	3	LR	–	0.37	0.59	Weimer et al. (2022, eingereicht)
Nicht-fiktionale Rede	8/1	1	2	DT	–	0.30	0.42	–
Generalisierung	8/1	1	6	RB	0.23	0.28	–	Gödeke et al. (2022, angenommen)
Generalisierung	8/1	1	6	RF	0.20	0.39	–	Gödeke et al. (2022, angenommen)
Attribution	8/1	1	3	NN	0.85	0.80	–	Dönicke et al. (2022)

Tab. 2: Performanzen (F1) bisher implementierter Klassifizierer und zugehörige Veröffentlichungen.

✉ hanna.varachkina@stud.uni-goettingen.de florian.barth@uni-goettingen.de luisa.goedeke@uni-goettingen.de annamareike.hofmann@uni-goettingen.de tillmann.doenicke@uni-goettingen.de