



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG



Textanalyse in Zeiten der *Digital Humanities*

Zur Annotation und Auswertung
geisteswissenschaftlicher Textdaten mit CATMA

Evelyn Gius, Universität Hamburg

Inhalte

- Zur Genese und Konzipierung von CATMA
- Anwendungsbeispiel 1:
 - Zur narrativen Konstitution von Konflikterzählungen
- Anwendungsbeispiel 2:
 - Zur Automatisierung narratologischer Analysen im Projekt heureCLÉA
- Hands-On-Einführung in CATMA

Die Genese und die Konzipierung von

CATMA 

Textanalyse durch Annotationen

- (digitale) Annotationen von Texten für verschiedene Zwecke (Edition, Aufbereitung digitaler Ressourcen/Information Retrieval, philologische/linguistische Textanalyse)

The screenshot displays a digital text analysis workflow. On the left, a document image is shown with a navigation bar (Bild: 0014, << vorherige Seite) and a zoom level of 63%. The document text is in German and discusses mineralogy. In the center, a flow diagram shows the process: Document → Document Representation → Query → Information Need. A 'Match' arrow points from the 'Query' box back to the 'Document Representation' box. On the right, a search results panel is visible, showing a search for 'Gneiß' and a list of results with various annotations and links.

Document Representation (XML snippet):

```

9 <pb fa
10 <fw pl
11 Ein Th
12 Granit
13 <div r
14 <he
15 </he
16 <he
17 fche
18 <p>I
19 fta
20 weit
21 ne (
22 mu&
23 ko&
24 nenr
25 fch
26 eig
27 nite (S. 6.) ziemlich u&#x0364;berein. Diefte Bergart fu&#x0364;hrt gleichfalls
28 Metalle.<lb/>
29 Sie ko&#x0364;mnt mit wenigen Glimmer zu <placeName>Scharfenberg</placeName>,
30 <placeName>Meißen</placeName> und <placeName>Alten-<lb/>
31 berg</placeName>, ganz ohne Glimmer aber zu <placeName>Miltitz</placeName>, unweit
32 <placeName>Meißen</placeName>, und befon-<lb/>
33 ders im <placeName>Plaufifchen Grunde</placeName> und zu
34 <placeName>Priesnitz</placeName>, unweit <placeName>Dresden</placeName>,
35 desglei-<lb/>
36 chen auch in der <placeName>Oberlaufitz</placeName> vor.</p>
37 </div>
38 </div><lb/>
39 <div n="4">
40 <head>8. <lb/>
41 </head>
42 <head>2.) Gneiß.</head><lb/>
43 <p>Der Gneiß befteht ebenfalls aus Feldfpath,
44 <note xml:id="note-0014" next="#note-0015" place="foot" n="e">Alle Mineralogen u&#x0
45 den Erkl&#x0364;rungen, die fie von dem Gneiß gaben; auch nicht einer er-<lb/>
46 wa&#x0364;hnt derfelben. Dafu&#x0364;r fetzen aber einige Steinmark, andere Speck-<lb/>
47 ftein, und noch andere verha&#x0364;rteten Thon in das Verzeichniß der Theile<lb/>
48 feines Gemenges, fo wie auch die meiften Neuern noch eines der drey<lb/>
49 letztern beybehaltten, und alfo 4 Theile feines Gemenges angeben. Ich,<lb/>
50 als ich im Jahre 1775. bey hiefiger Bergakademie das erfte mal Vor-<lb/>
51

```

Search Results Panel:

Normalisierung: <choice><orig><neg>
Abkürzung: <choice><abbr><expan>
Hervorhebungen: #b #g #i #k #sub #sup #aq #fi #in #blue #red
Verknüpfungen: @next @next/@xml:id @prev @prev/@xml:id

Suche:

Anmerkung zu dieser Seite anlegen

Textanalyse durch Annotationen

washed him down. I played for it
I've dark green pebbles for eyes,
the back of my hands I could strangle with.
I stink and remember.
I'm in bed cawing Nooooo at the wall; the dress yellowing, the
wardrobe;
ed mirror, full-length, her, myself, who did this
to believe
comes back
wakes up from
ate behind a white veil; a red balloon bursting in my face.
tabbed at a wedding-cake.
nk it's only the heart that b-b-b-breaks.

p. 9 - Right eye for American. So we're seen
in 1st 70s. So why Tony arrived at Luke's
garden, Sidelin at the?
p. 9-11 - Luke: faithful & self-del.
21 - The father-child (Shelton) Am. p. 10 on time
is not adequate. Also a child's 2000 a child's mind
pulling with
21 - Luke: "Freaking skin."
21 - Luke: "Lay over here so I can? What not here
pretty. Be a what not guard
31 - All these references to Thales at
Necrotic
36 - Rome read past. - first-time plus time
p. 57 - Aunt Sumner - first-time of edge
like - Billie's foundation of
gains body function Smith
(200, ed.) (p. 1) - Science &
Medicine for the general
public at the hospital?
p. 101 - Billie - "I think
of mine - a real kind
of virility, always in
relation to mine -
Revolutionary consciousness
116-117 - Luke: More writing of writing
is L.P.'s only measure of beauty"
132 - Luke: Victim of the
Vegetal in Pappas, Kitzabo
1954 - What do they know him to do
of control of chapter? - "Opposite" - Billie's Return
257-261 - Ab - Same as Billie's old Chinese man

Don DeLillo's
RATNER'S STAR
"DeLillo's latest and best meditation on the excesses of consumer
thought... Neither outline nor message begins to sustain the
tiness of *Ratner's Star*... Nor do plot and meaning confer the
dazzling capacity to make farfetched matters of the most alluring
scientific theory... Nor the lyric poetry that Mr. DeLillo makes
easily as breathing." - Christopher Lehmann Haug, *The New York Times*
"Prestigious, accomplished, undoubtedly brilliant."
- Joe David Bellamy, *Chicago Daily News*
"DeLillo's amazing skill with the language lets him move from sane
renderings of intellectual jargon to the lyrical flow of interior monologue
He has a chillingly persuasive view of who we are in the twentieth Ce
tury and where we might be going."
- Bruce Allen, *Chicago Tribune Book Rev*
"Brilliant... *Ratner's Star* (1979) plays with science and science fiction
... The wit, elegance, and economy of Don DeLillo's art are equal to it
biter clarity of his perceptions."
- Diane Johnson, in *Terrorism and Novels*
"Eerie and intriguing... funny as well as instructive... DeLillo's aim is
to show how the codification of phenomena as practiced by scientists leads
aburdity and madness."
- Paul Gray, *Tar*
"Hilariously funny... Yet, as is evident in so much modern comic fiction
humor and terror are close to one another. And though DeLillo write
about a not too futuristic society, he is looking at those realities that dai
surround us."
- John D. Bryer, *Chicago Sun-Tim*

3rd person
Tension
Small community
family dispute
honour
property/ownership
inheritance
land
A = stubborn, self-confident

1. (a)
business
3rd person
Tension
Small community
family dispute
honour
property/ownership
inheritance
land
A = stubborn, self-confident

She was almost through with her figures when she heard a cart drive up to the
gate, and looking out of the window she saw her two older brothers. They had seemed
to avoid her ever since Carl Linnum's arrival, four weeks ago that day, and she hurried
to the door to welcome them. She saw at once that they had come with some very
definite purpose. They followed her stiffly into the sitting-room. Oscar sat down, but
Lou walked over to the window and remained standing, his hands behind him.
"You are by yourself?" he asked, looking toward the doorway into the parlor.
"Yes. Carl and Emil went up to the Catholic fair."
"For a few moments neither of the men spoke."
Then Lou came out sharply. "How soon does he intend to go away from here?"
"I don't know, Lou. Not for some time, I hope." Alexandra spoke in an even,
quiet tone that often exasperated her brothers. "They felt that she was trying to be
superior with them."
Oscar spoke up grimly: "We thought we ought to tell you that people have begun
to talk," he said meaningly. "Alexandra looked at him. "What about?"
Oscar met her eyes blankly. "About you, keeping him here so long - it looks bad
for him to be hanging on to a woman this way. People think you're getting taken in."
Alexandra shut her account-book firmly. "Boys," she said seriously, "don't let's
go on with this. We won't come out anywhere. I can't take advice on such a matter. I
know you mean well, but you must not feel responsible for me in things of this sort. If
we go on with this talk it will only make hard feeling."
Lou whipped about from the window. "You ought to think a little about your
family. You're making us all ridiculous."
"How am I?"
"People are beginning to say you want to marry the fellow."
"Well, and what is ridiculous about that?"
Lou and Oscar exchanged outraged looks.
"Alexandra! Can't you see he's just a tramp and he's after your money? He
wants to be taken care of, he does!"
"Well, suppose I want to take care of him? Whose business is it but my own?"
"Don't you know he'd get hold of your property?"
"He'd get hold of what I wished to give him, certainly."
Oscar sat up suddenly and Lou clutched at his bristly hair.
"Give him?" Lou shouted. "Our property, our homestead?"
"I don't know about the homestead," said Alexandra quietly. "I know you and
Oscar have always expected that it would be left to your children, and I'm not sure but
what you're right. But I'll do exactly as I please with the rest of my land, boys."
"The rest of your land!" cried Lou, growing more excited every minute. "Didn't
all the land come out of the homestead? It was bought with money borrowed on the
homestead, and Oscar and me worked ourselves to the bone paying interest on it."

Brothers
= awkward
determined
superior
attitude
Money
She's
independent -
They're
influenced
by community

Unser Ausgangspunkt:

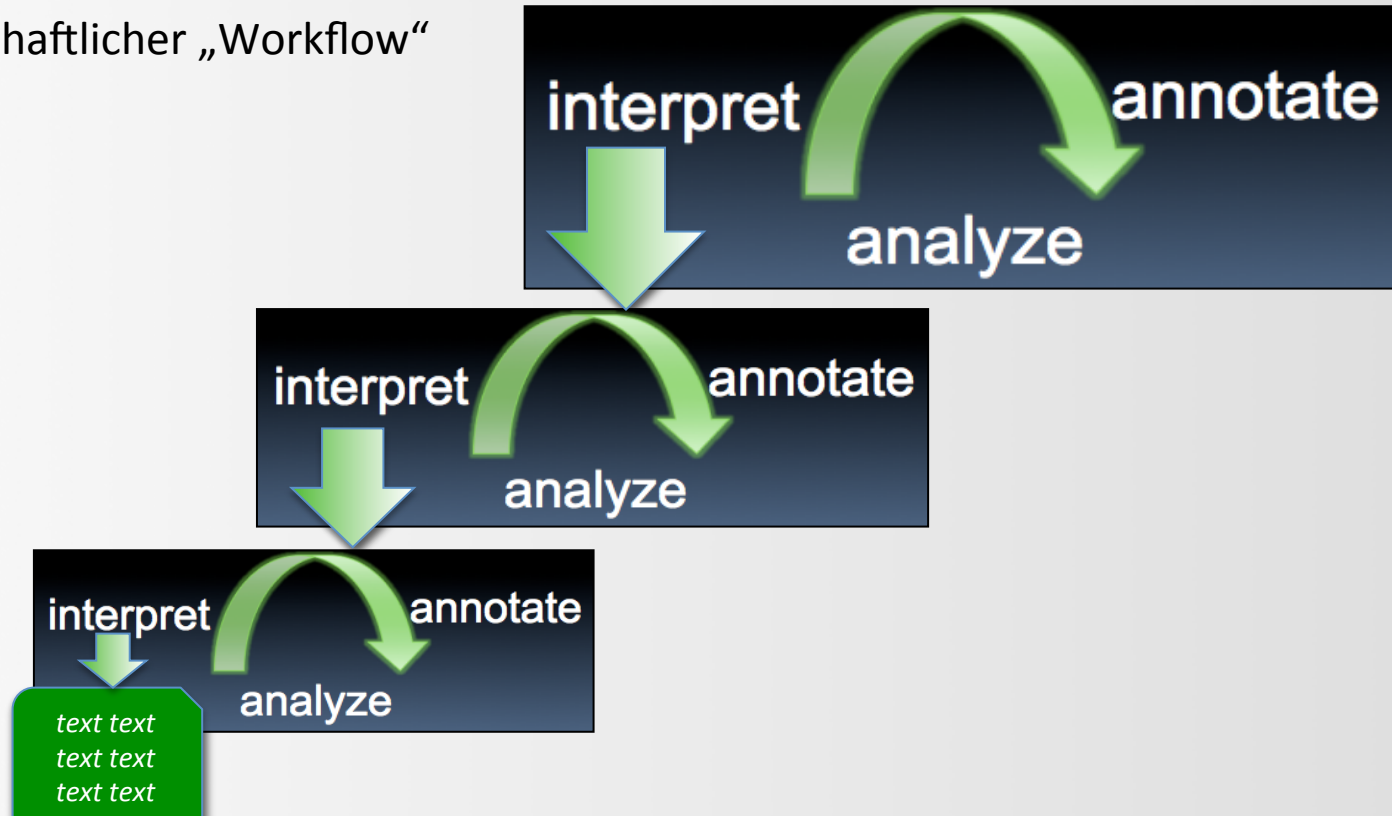
- philologische Textanalyse
- Funktion: Annotationen als ...
 - Anmerkungen
 - Hervorhebung besonders relevanter Textstellen
 - Markierung weiter zu analysierender Textstellen

→ Annotation als wesentliches Mittel der philologischen Textanalyse

Textanalyse in der Literaturwissenschaft

Ausgangsfrage:

- Was bedeutet dieser Text?
→ literaturwissenschaftlicher „Workflow“



CATMA-Vorgeschichte

Textanalyse mit TACT (Textual Analysis Computing Tools; entwickelt von 1986-1989 von John Bradley, Ian Lancashire, Lidio Presutti und Michael Stairs an der University of Toronto; Version 2.1: 1996)

[\[Home\]](#) -- [\[TOC\]](#) -- [\[Next\]](#)

● **Hardware and software requirements**

TACT requires a standard MS-DOS platform with 640K RAM (preferably up to 16Mb additional memory); DOS 3.x or above; a large hard disk (the uncompressed installed programs occupy about 3.2Mb); and a -386 or faster processor.

TACT will work fine under Windows 3.11 and Windows 95/98. It however does NOT work under Windows NT.

TACT

The screenshot displays the TACT software interface. At the top is a menu bar with the following items: File, Select, Displays, Group, Help (F1), and 224 K. The main window has a blue background and contains several reference tags represented by white lines and letters. Below the tags, the text 'UseBase V' is visible. Further down, the text 'Text: (DERUG' and '.GIX: ----' is shown. At the bottom of the main window, the text 'Copyright (c) 1994 by Centr' is partially visible. On the left side of the main window, the text 'F1: F2: F3: F4: Reference Tags' is displayed. Overlaid on the bottom right of the main window is a 'Query Dialog Box' with a black background and white text. The dialog box contains a 'Query:' field, a 'Span Context:' field, and a 'Unit:' field set to 'word'. Below these fields, the text 'Before: 5' and 'After: 5' is displayed. At the bottom of the dialog box, the text 'For a list of Reference Names for "Unit", type "?" in the Unit field.' is shown. Below this text, the text 'Ignore Errors? Yes: _ No: X' and 'Manual Select? Yes: X No: _' is displayed.

File Select Displays Group Help (F1) 224 K

UseBase V

Text: (DERUG
.GIX: ----

Copyright (c) 1994 by Centr

F1: F2: F3: F4: Reference Tags

Query Dialog Box

Query:

Span Context: Unit: word
Before: 5 After: 5

For a list of Reference Names for "Unit", type "?" in the Unit field.

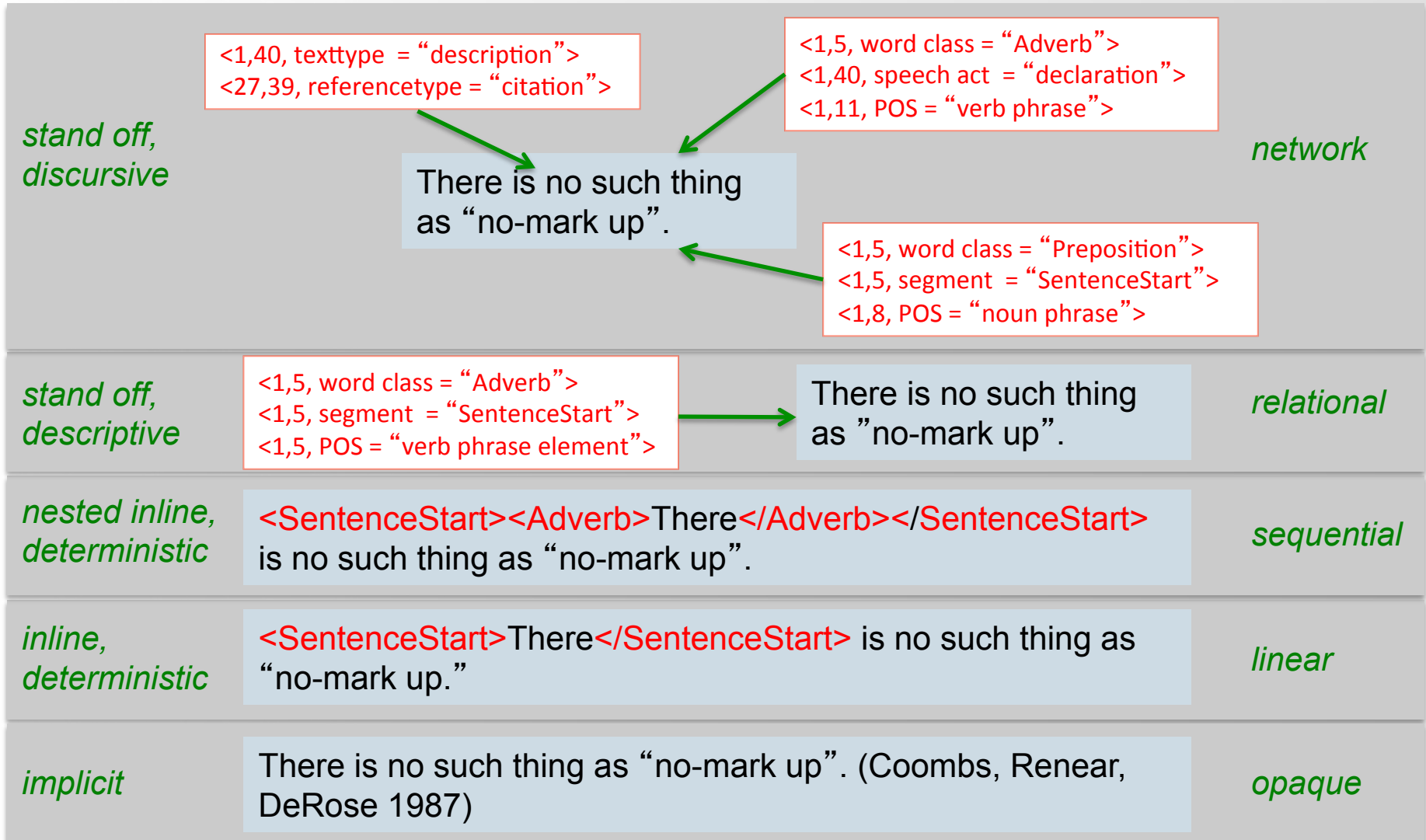
Ignore Errors? Yes: _ No: X
Manual Select? Yes: X No: _

Von TACT zu CATMA

Grundsätzlich: Funktionalitäten aus TACT + Standoff Markup

- Tagging mit frei wählbaren Konzepten (Tagnamen)
- Abfragen von:
 - einem oder mehreren Wörtern (getrennt durch Kommata)
 - Reguläre Ausdrücke
 - Abfragen aufeinander folgender Wörter (Phrasen)
 - Kollokationen
 - ähnliche Wörter
 - Frequenz
- + verfeinerte Abfragen (z.B. unter Einbezug von Tags),
sowie einstellbarer Suchkontext („Span Context“)

Markup-Typen und Datenmodelle



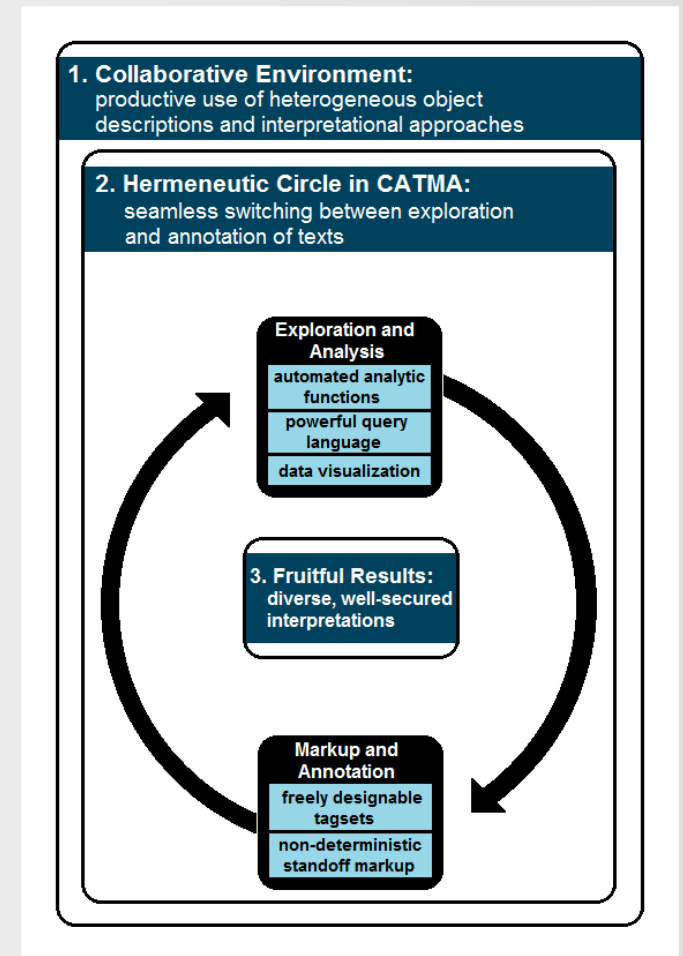
Hermeneutische „must haves“ des diskursiven Markups

Dynamisches und diskursives Markup:

- leicht erweiterbare Tags und Tagsets
- multiples (widersprüchliches) Markup
- Unterstützung des hermeneutischen Zirkels durch Integration von Markup und Analyse

Kollaboratives Arbeiten und nondeterministisches Tagging:

- mehrfaches Taggen derselbe Textstelle
- sich überschneidende Tags
- widersprüchliche Tags



Der CATMA-Zugang zu Markup

- Ein Tag referenziert einen Textbereich (*text range*) durch Angabe des Anfangs- und Endpunktes (*character offset*)
- Externes Standoff Markup ermöglicht...
 - überlappendes Tagging
 - durch mehrere Nutzer gleichzeitig
- TEI-basiert durch Nutzung von *feature structures*:
 - Codierung der Tagdefinitionen: *feature system declaration*
 - <fsdDecl> groups a set of tags
 - <fsDecl> defines a tag
 - <fDecl> defines a property of a tag
 - Codierung der Taginstanzen: *feature structures*
 - <fs> defines a Tag instance
 - <f> defines the instance's properties

Computer Aided Textual Markup and Analysis (CATMA)

The screenshot displays the CATMA software interface, divided into two main windows: the Repository Manager and the Tag Manager.

Repository Manager: This window is titled "Repository Manager" and has tabs for "Repositories Overview" and "Local Repository". It features a "Document Manager" section with a "Corpora" list (including "All documents", "The national corpus", and "My super special corpus") and a "Documents" list (including file paths and "A Rose for Emily"). An "Information" panel on the right shows details for a document, such as "Author: William Fa", "Description: empty", "Publisher: empty", and "Title: A Rose for". Below the document lists are buttons for "Create Corpus", "Open Document", "Add Document", and "More actions...".

Tag Manager: This window is titled "Tag Manager" and shows a "My Tag Library various nice tagsets" view. It contains a table of tagsets and their colors:

Tagsets	Tag Color
Time	
Tenses	Purple
Order	Green
Dates	Cyan
Frequency	Red
Duration	Yellow
summary	Yellow
ellipsis	Yellow
scene	Orange
pause	Orange
stretch	Yellow
MoveTest	
TestTagset	
Standard Tagset	

To the right of the table are three sections of controls: "Tagset" (with "Insert Tagset", "Remove Tagset", and "Edit Tagset" buttons), "Tag" (with "Insert Tag", "Remove Tag", and "Edit Tag" buttons), and "Property" (with "Insert Property", "Remove Property", and "Edit Property" buttons).

Computer Aided Textual Markup and Analysis (CATMA)

- CATMA 1 (Juni 2009): Analyzer in Microsoft C++, Tagger in Java
- CATMA 2 (Januar 2010): Plattformunabhängige Java-Version
- CATMA 3 (Juli 2010): Version mit natürlichsprachlichem Query Builder, Hilfefunktion an diversen Verbesserungen
 - CATMA to go (CATMA 3.2, Mai 2011)
- CATMA 4 (Februar 2013): Webversion mit kollaborativer Funktion und Möglichkeit zur Korpusanalyse
- Aktuelle Version: CATMA 4.2 (Oktober 2014)
- weitere Informationen unter www.catma.de

Anwendungsbeispiel 1

Zur narrativen Konstitution von Konflikterzählungen

ERZÄHLEN ÜBER KONFLIKTE

Erzählen über Konflikte.

Eine computergestützte narratologische Untersuchung von narrativen Interviews zu Arbeitskonflikten

- **Das *Wie* als Indikator des *Was***

Ausgangsbeobachtungen:

- zwischen Erlebtem und dem Erzählen über das Erlebte bestehen systematische Zusammenhänge (vgl. Methoden der Psychotherapie)
- Dieser Zusammenhang ist auch für die Konfliktbearbeitung relevant:
Wahrnehmung [von Konflikten] hat bei der Konfliktdiagnose nicht nur auf das Was zu achten, d. h. auf die inhaltlichen Streitpunkte ("Issues"), sondern noch viel mehr auf das Wie: Auf welche Weise spricht eine Partei über ihren Gegner? [...] Wie spricht sie über ihre Streitpunkte?

(Glasl, 2011, S. 104)

Hypothese:

- Die Konflikthaftigkeit einer Situation kann an der speziellen Beschaffenheit der Erzählung über den entsprechenden (akuten) Konflikt festgemacht werden.

Fragestellung:

- Wie hängt die Art und Weise, wie ein Konflikt erzählt wird, mit der Beschaffenheit und insbesondere der Eskaliertheit des erzählten Konfliktes zusammen?

Reihenfolge der Analyse

	Inter-pretation	Text-umfang	Kontext	geordnet nach Abhängigkeiten
1 time	1	1	2	1 time
2 space	1	1	2	2 space
3 modes of representation	1	3	1	3 modes of representation
4 relation narrator–world	1	4	5	4 narrative levels
5 relation narrator–event time	1	5	5	5 self-reflexive narration
6 self-reflexive narration	2	2	2	6 relation narrator–world
7 narrator representation	2	2	3	7 relation narrator–event time
8 timerelation discours–histoire	2	4	4	8 character
9 narrative levels	2	4	5	9 narrator representation
10 narratee representation	3	2	3	10 narratee representation
11 perspective	4	3	3	11 plot
12 character	4	3	5	12 timerelation discours–histoire
13 information allocation	4	4	2	13 perspective
14 plot	4	5	5	14 information a
15 reliability	5	5	5	15 reliability

- Die Reihenfolge der Analyse
- unterstützt den hermeneutischen Prozess durch genau das Kontextwissen, das nötig ist, und
 - ermöglicht gleichzeitig eine textorientierte, interpretationsarme Analyse

Analyse

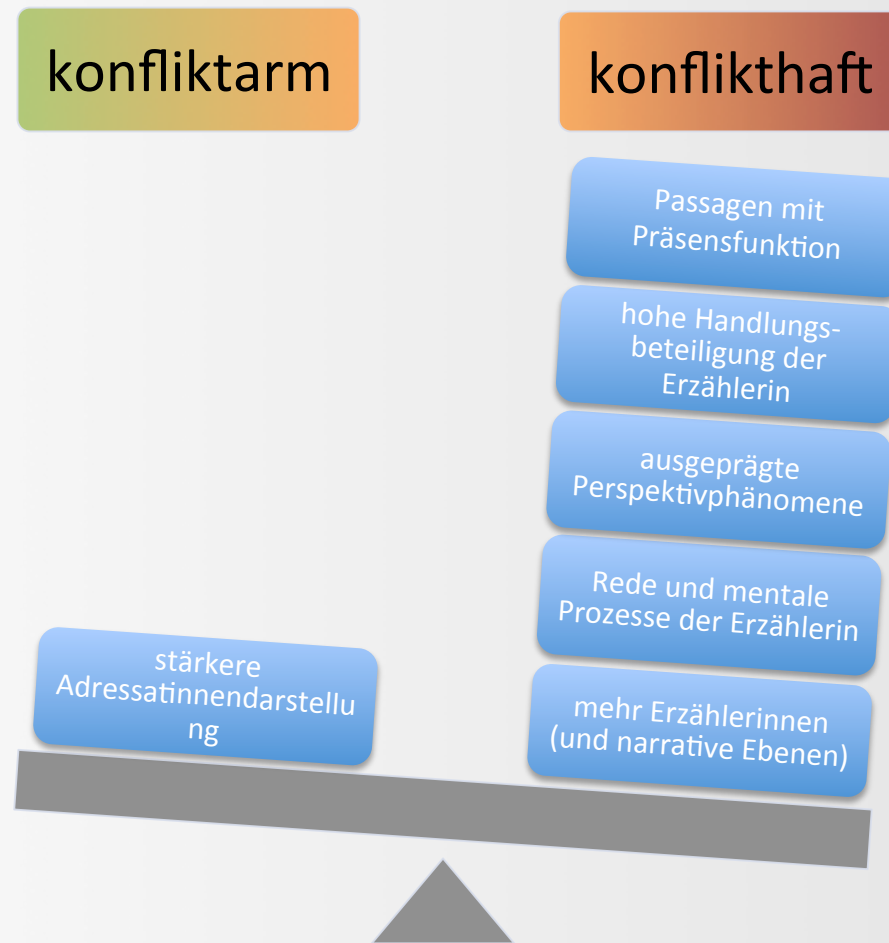
- Transkripte von narrativen Interviews mit 17 Personen, die einen akuten Arbeitskonflikt haben bzw. hatten
- Korpus mit insgesamt 39 Erzählungen:
 - 18 Erzählungen über akute Konflikte
 - 15 Erzählungen über vergleichbare konfliktlose Situationen
 - 6 Erzählungen über den nicht mehr akuten/gelösten Konflikt
- computergestützte Annotation anhand der narratologischen Konzepte in den 14 Phänomenbereichen (ca. 500 Konzepte bzw. Tags)
- kombinierter Zugang:
 - induktives Vorgehen (bottom up): Analyse aller Phänomene in einem exemplarischen Erzähltripel (akut, konfliktlos, nicht mehr akut)
 - deduktives Vorgehen (top down): Analyse von Phänomenen in Verbindung mit Zeit sowie Phänomenen der Rede- und Gedankenwiedergabe in allen Erzählungen

Annotation und Auswertung

- 23.904 annotierte Textstellen:
 - 15.546 narratologische Tags
 - 8.358 Hilfstags (Sprecherin, Erzähltyp etc.)
- plus: 4.147 zusätzliche Merkmale („properties“) für Tags
- Auswertung der Verteilung narratologischer Phänomene
 - in Bezug auf jeden narratologischen Phänomenbereich, und
 - in Abhängigkeit von den drei Erzählungstypen (akuter Konflikt, konfliktlose Situation, nicht mehr akuter Konflikt)

Active Tagsets	Active Markup Collections	Markup collections	Tag color
		time	blue
		order	pink
		frequency	purple
		duration	grey
		time	blue
		space	blue
		self-reflexive_narration	blue
		relation_narrator-world	blue
Tag Instance	Tag	Tag Path	Tag
main_narration	main_narration	/narratologic	CAT
repetitive	repetitive	/narratologic	CAT
plot_sequence	plot_sequence	/narratologic	CAT
mental_process_representation	mental_process_representation	/narratologic	CAT
narrator_self-description	narrator_self-description	/narratologic	CAT
primary_narration	primary_narration	/narratologic	CAT
value_system-based_perspective	value_system-based_perspective	/narratologic	CAT
narrated_character_thought	narrated_character_thought	/narratologic	CAT
marked_narrator_identity	marked_narrator_identity	/narratologic	CAT
propertyvalue_narrating_I	propertyvalue_narrating_I	/narratologic	CAT

Lineare Zusammenhänge Phänomen-Konflikthaftigkeit



Linear-opponierte Zusammenhänge Phänomen und Konflikthaftigkeit

wenig konflikthaft

- schneller erzählt
 - iterativ erzählt
- komplette Anachronien
 - externe Analepsen
 - erzählte Rede
 - Redewiedergabe
- Gefühle des Gegenüber

vs.

hoch konflikthaft

- langsamer erzählt
- repetitiv erzählt
- partielle Anachronien
- interne Analepsen
- transponierte Rede
- Wiedergabe mentaler Prozesse
- Gedanken des Gegenüber

Typische Phänomenkonstellationen

- Komparative Indikatoren:
 - relative Ausprägung der Phänomene

konfliktlos	nicht mehr akut/gelöst	akut
<ul style="list-style-type: none">• viel prospektives Erzählen• deutlichere Erzählerinnendarstellung	<ul style="list-style-type: none">• gemeinsame Rede und mentale Prozesse• Perspektive aller Beteiligten	<ul style="list-style-type: none">• weniger Variationen im Erzähltempo

Absolute Phänomenkonstellationen

konfliktlos	nicht mehr akut/gelöst	akut
simultan & retrospektiv	retrospektiv	simultan
chronologisch	anachronisch (invers-linear)	anachronisch
<ul style="list-style-type: none"> • wenig äußere Handlungen des Gegenüber, keine inneren Handlungen • mehr Rede als mentale Prozesse • Rede des Gegenüber häufig zitiert • Rede der Hauptbeteiligten ähnlich häufig • Haupterzählerin erzählt weitere narrative Ebenen • wenig interpersonelle Charakterisierungen • vollständige Skripte, kaum Abweichungen • Textanteil beschleunigter Passagen am höchsten • überwiegend kompakte Perspektive der Erzählerin • keine zeitliche Perspektive • keine Perspektive anderer 	<ul style="list-style-type: none"> • gemeinsames Wissen über kognitive Inhalte des Gegenüber • negative Bewertungen durch das Gegenüber • Darstellung problematischen Verhaltens 	<ul style="list-style-type: none"> • häufige multisingulative Passagen

Ansatzpunkte für Konfliktanalyse

- Phänomene als Ausgangspunkt für inhaltliche Betrachtungen:
 - Tempuswechsel
 - Negative interpersonelle Charakterisierungen
 - Zusammenfallen der figuralen und narratorialen Funktion der Erzählerin
 - Repetitive Passagen
 - Häufigkeit
 - iterativ-repetitive Passagen
 - explizite Adressatinnendarstellung in höheren Ebenen
 - zeitliche Perspektive
 - ideologische Perspektive

Ergebnisse für die narratologische Theoriediskussion

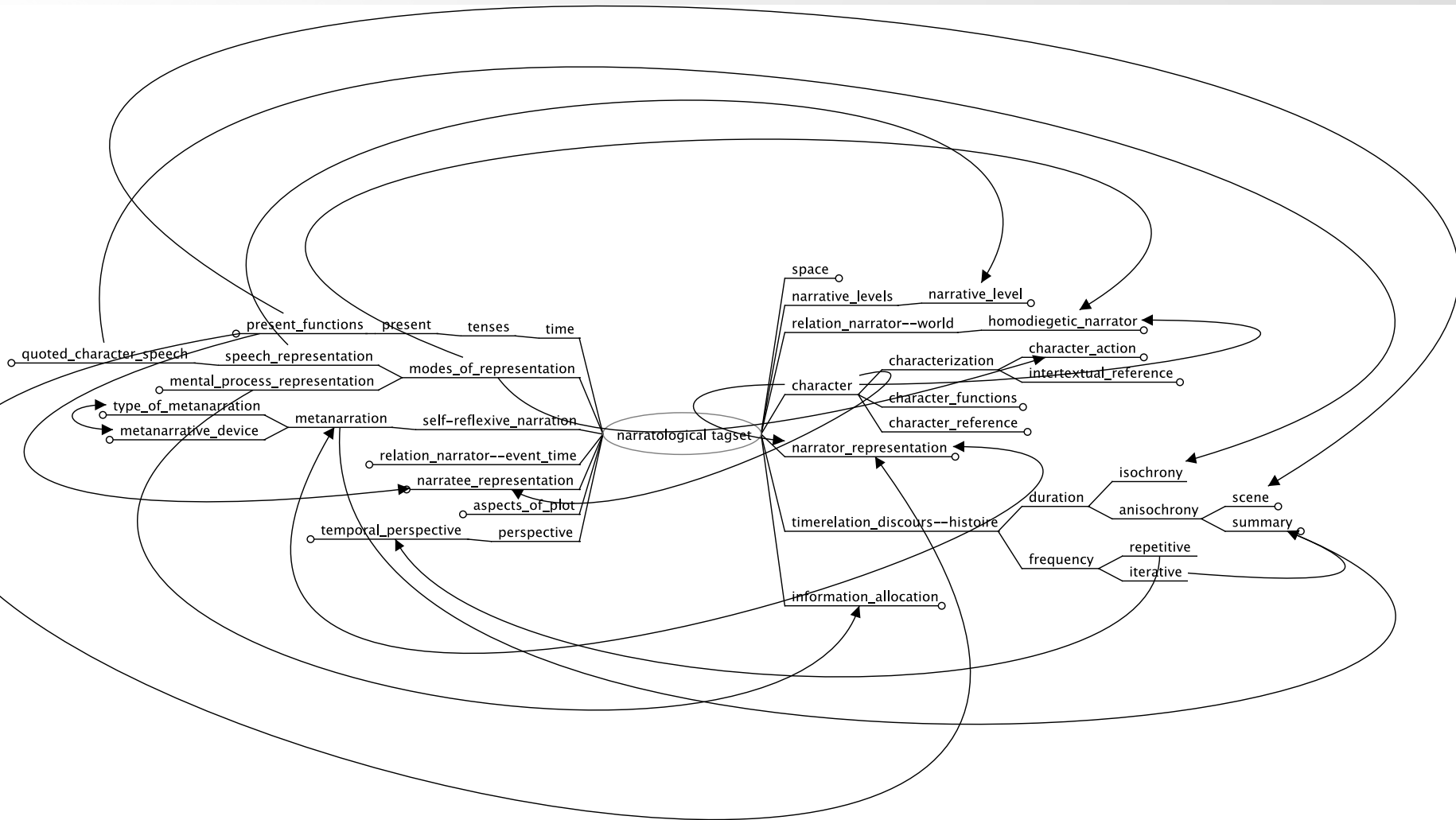
- Beitrag zur Narratologie:
 - Umfangreiche Anwendung narratologischer Konzepte
 - Überprüfung der Anwendbarkeit der einzelnen narratologischen Konzepte

- Ergebnisse:
 - Kategorien für die Anwendung präzisiert
 - Zusammenhänge zwischen Kategorien innerhalb verschiedener Phänomenbereiche verdeutlicht
 - Ansatzpunkte für die Weiterarbeit: Überprüfung der narratologischen Konzepte auf Redundanzen

Für eine ausführliche Beschreibung vgl.

Gius, Evelyn (im Erscheinen) „Erzählen über Konflikte. Ein Beitrag zur digitalen Narratologie“. Narratologia, Berlin [u.a.]: De Gruyter.

Ergebnisse für die narratologische Theoriediskussion



Anwendungsbeispiel 2

Zur Automatisierung narratologischer Analysen



Computergestützte Erzähltextanalyse in heureCLÉA

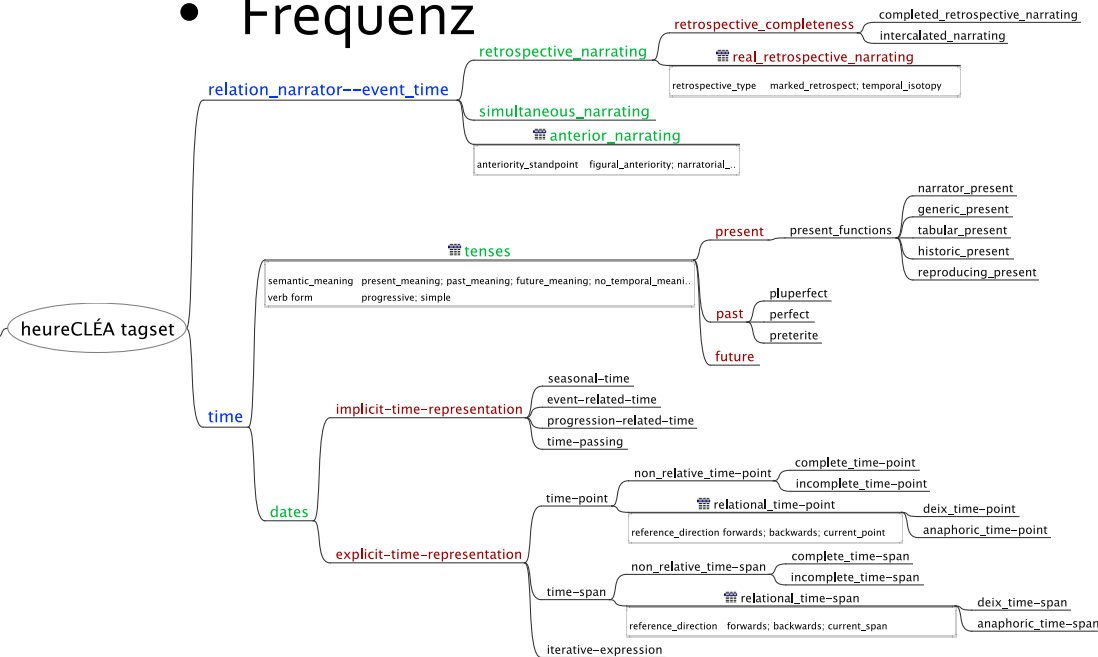
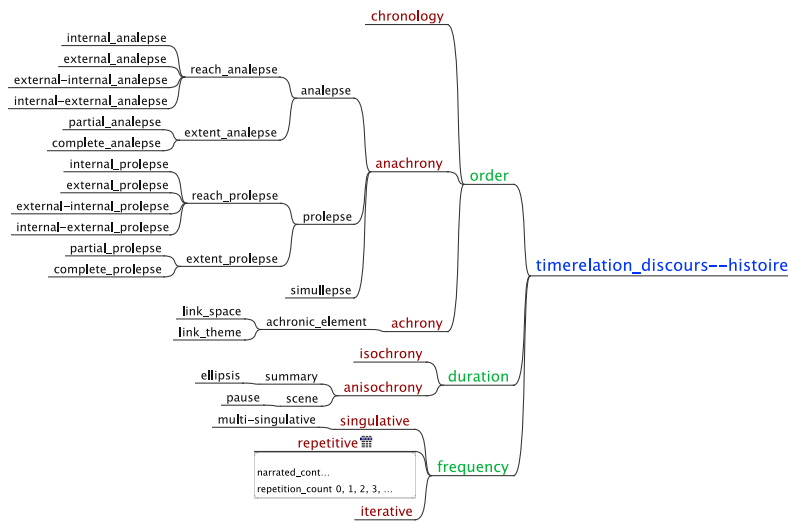
heureCLÉA (cf.: CLÉA= Collaborative Literature Exploration and Annotation):

- Entwicklung einer “digitalen Heuristik”: ein funktionales Modul, das Literaturwissenschaftler/innen bei der Annotation und Interpretation von Texten unterstützt
- Exemplarischer Zugang: semantische Analyse von Phänomenen der Zeit in deutschsprachigen Erzählungen um 1900
- Entwicklungsphasen:
 1. Automatisierung von niedrigkomplexen Annotationsaufgaben (HeidelTime, TreeTagger, Morphisto etc.)
 2. Exploration und Analyse komplexerer manuell und automatisch generierter Markup Varianten
 3. computergestützte Modellierung und Generierung von Markup Varianten

Narratologische Phänomene der Zeit

Ausgehend von Genette (1972) bzw. Meister & Lahn (2013):

- Zeitausdrücke und Tempus
- Dauer
- Ordnung
- Frequenz



- + zugrunde liegende Konzepte
- (Erzählebenen, Narrativität...)

Annotation in heureCLÉA

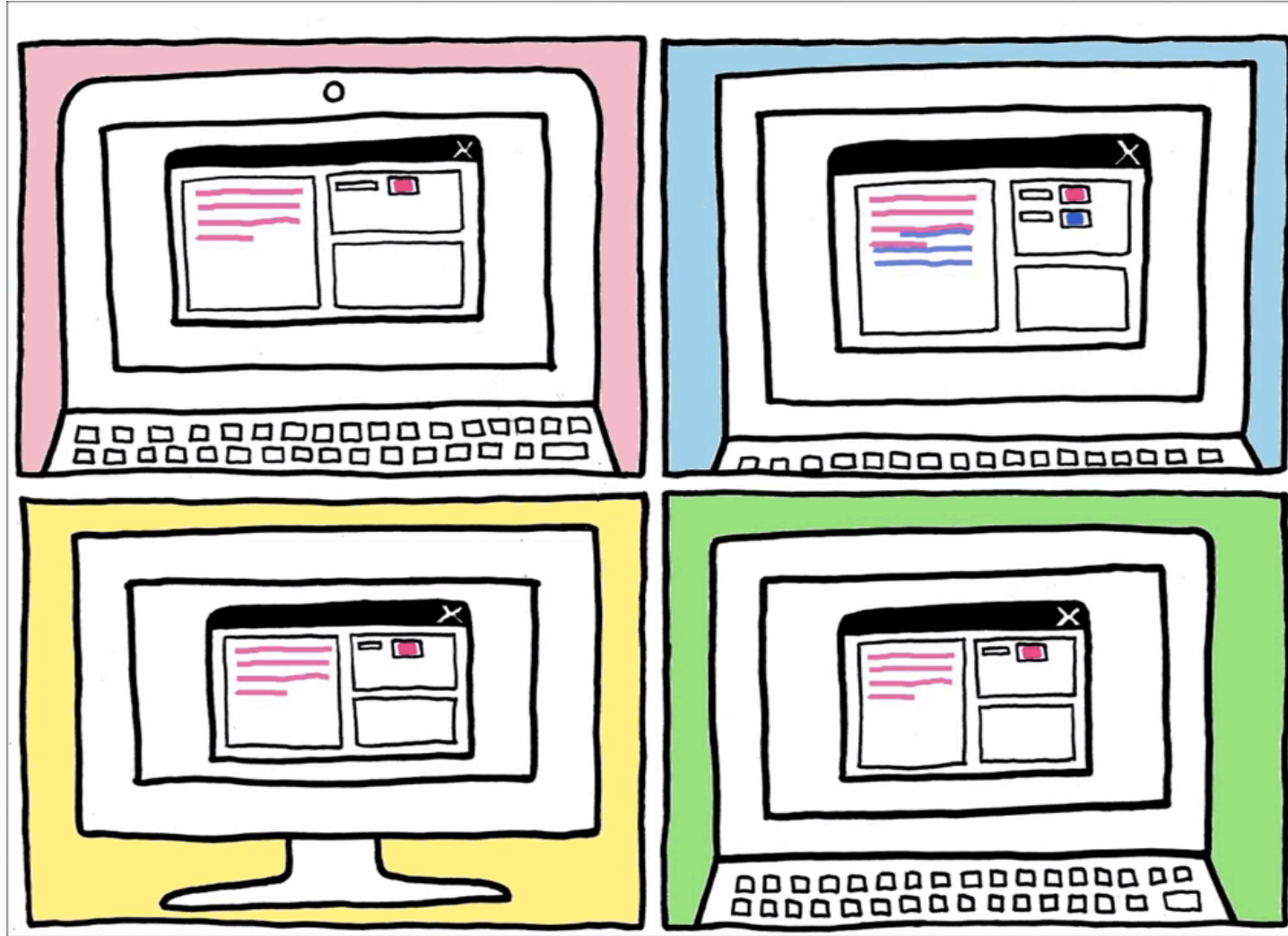
Auszug aus den Guidelines:
Zusammenfassung der Annotation von Ordnungsphänomenen

Tagstring <ul style="list-style-type: none"> • Textabschnitte – Mindestgröße: Teilsatz 	Unmarkierter Fall <ul style="list-style-type: none"> • Chronologisches Erzählen
Indikatoren auf der Textoberfläche <ul style="list-style-type: none"> • Zeitausdrücke, die Vorzeitigkeit, Gleichzeitigkeit oder Nachzeitigkeit ausdrücken • Tempuswechsel 	
Tagging-Routine <ol style="list-style-type: none"> 1. Annotation aller nicht chronologisch dargestellten Textpassagen als Prolepse, Analepse, Simullepse oder Achronie. 2. Bei Anachronien: Spezifizierung von Umfang und Reichweite. 3. Bei Achronien: Spezifizierung der Verknüpfungsart. 	
Beispiele <ul style="list-style-type: none"> • chronologisches Erzählen: „von ohngefähr erhob sie das Auge und traf mit dem blauesten Strahle in seinen Blick. Er ward wie von einem Blitz durchdrungen. Sie strauchelte, und so schnell er auch hinzusprang, konnte er doch nicht verhindern, daß sie nicht kurze Zeit in der reizendsten Stellung knieend vor seinen Füßen lag“ (Der Pokal) • Analepse: „Jetzt sah man, was geschehen war: <u>der Hansjörg hatte sich am mittleren Gelenk den Zeigefinger der rechten Hand abgeschossen</u>“ (Die Kriegspfeife) • Prolepse: „Zwanzig Jahre lang habe ich den Tod auf den Tag herbeigezogen, <u>der in einer Stunde beginnen wird [...]</u>“ (Der Tod) • Simullepse: „'Ich bin nicht allein', sagte ich [...]. Dabei <u>preßte sich mein Arm, der die Decke über ihren Kopf gelegt hatte, krampfhaft auf jene Stelle, wo ich den Mund vermutete [...]</u>“ (Die Schutzimpfung) • Achronie: „Vorliebe empfindet der Mensch für allerlei Gegenstände. Liebe, die echte, unvergängliche, die lernt er – wenn überhaupt – nur einma kennen.“ (Krambambuli) 	

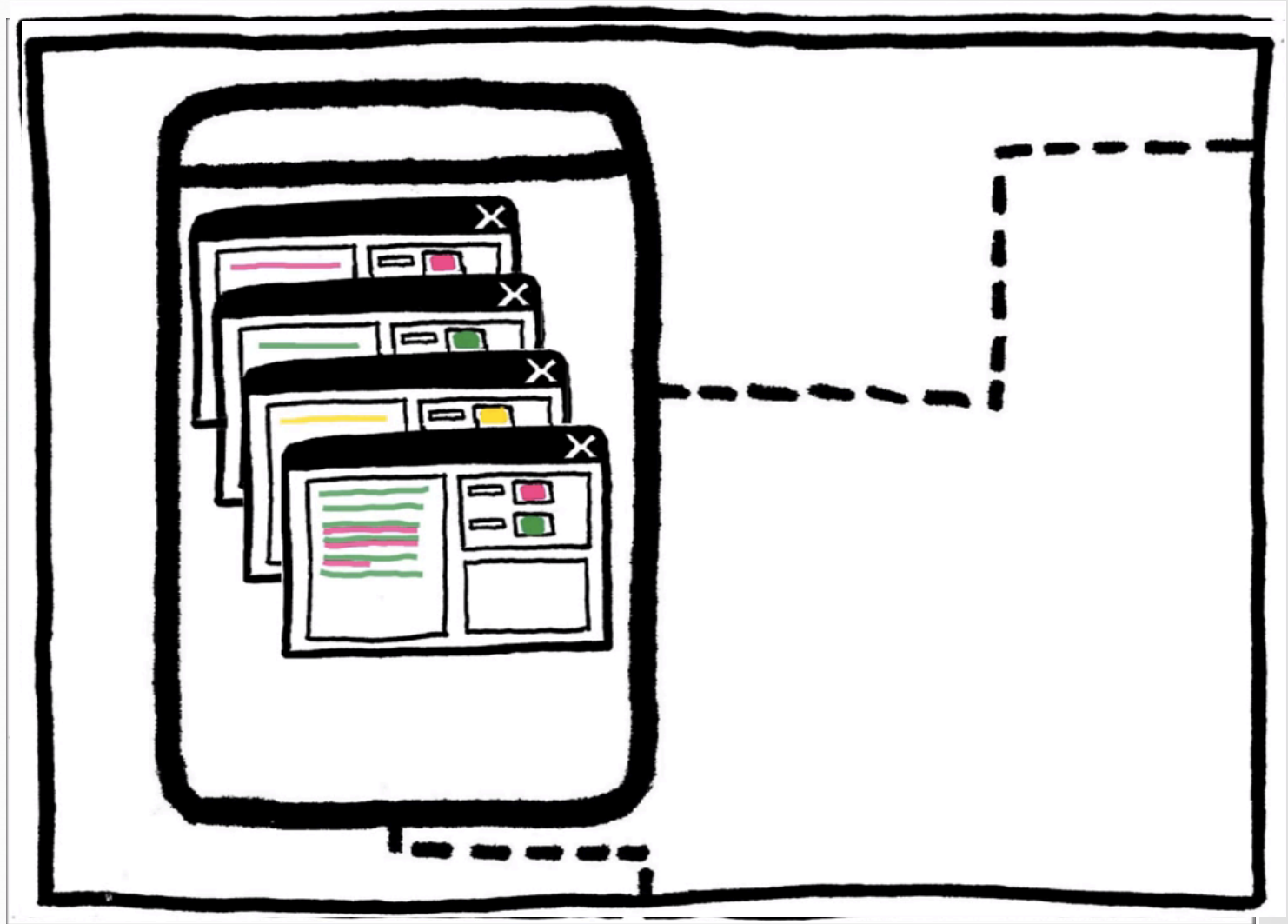
Annotationsworkflow



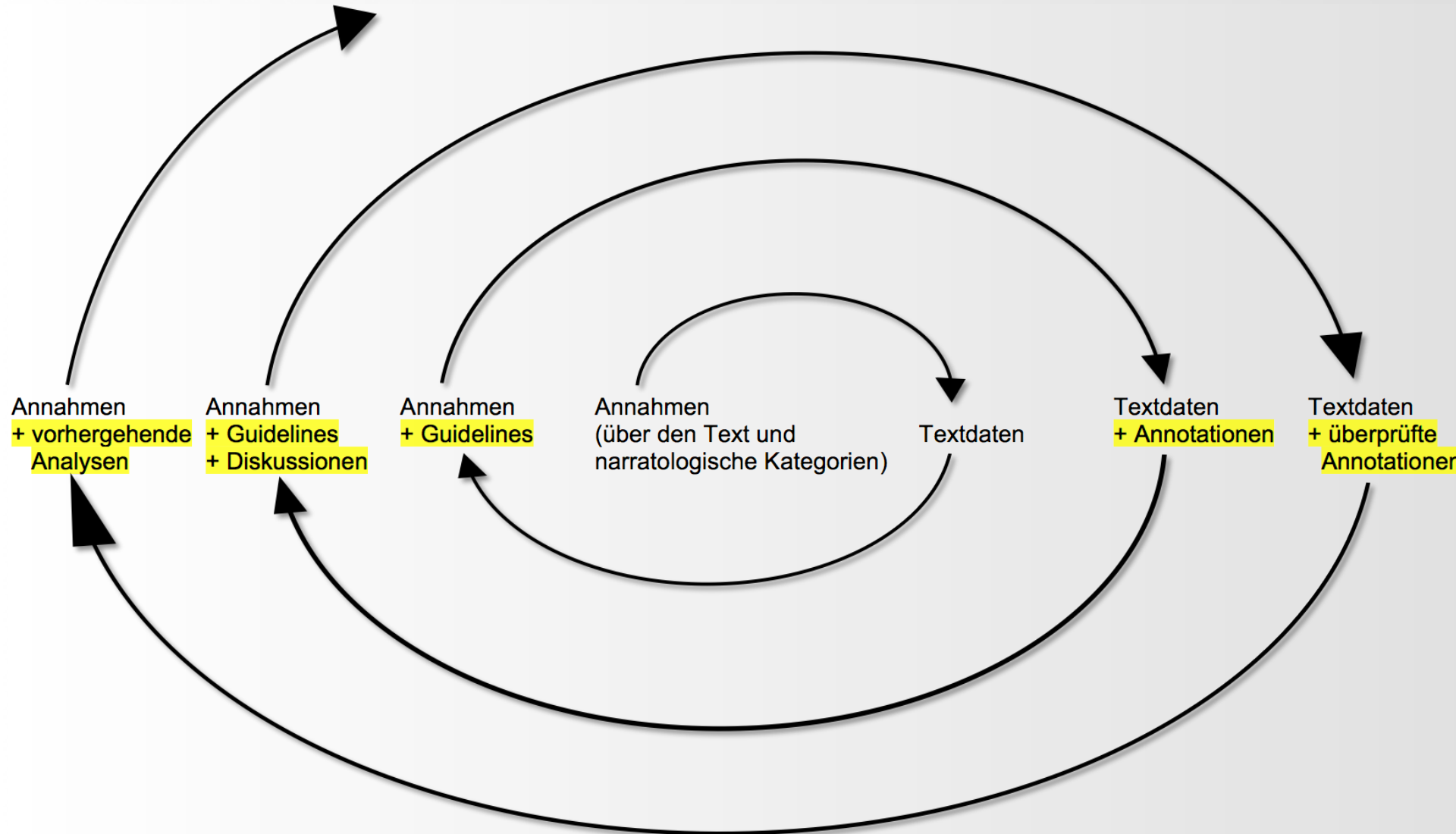
Annotationsworkflow



Annotationsworkflow



Der kollaborative hermeneutische Zirkel



Interpretation vs. Automatisierung?

Literaturwissenschaftliche Textanalyse:

- hermeneutisches Textverstehen
- Pluralität von Interpretationen, Intersubjektivität

INDETERMINISTISCH

Natural Language Processing (NLP):

- einheitliche Datenbasis, Reproduzierbarkeit
- gold standard, interannotator agreement

DETERMINISTISCH

Narratologische Erkenntnisse

- Computerbasierter Zugang – und seine kollaborative Anwendung – verdeutlicht:
 - unklare Einzelkonzepte:
 - Inkonsistenzen durch unterspezifizierte Konzepte
 - betrifft einzelne Konzepte
 - diese müssen spezifiziert werden
 - grundlegende theoretische Unklarheiten:
 - Inkonsistenzen durch (implizite!) Abhängigkeiten von grundlegenden narratologischen Konzepten
 - betrifft abhängige Konzepte
 - Theorieentwicklung notwendig

Grundlegende theoretische Unklarheiten: Dauer und Ereignis

Ich habe die ganze Nacht hinausgeblickt, und mich dünkte, so müsse der Tod sein
oder das Nach dem Tode: dort drüben und draußen ein unendliches, dumpf
brausendes Dunkel. Wird dort ein Gedanke, eine Ahnung von mir fortleben und
-weben und ewig auf das unbegreifliche Brausen horchen?

(Thomas Mann *Der Tod*)

▼ ◆ anisochrony	■
▼ ◆ scene	■
◆ pause	■
▼ ◆ summary	■
▶ ◆ ellipsis	■
◆ isochrony	■

- Das Konzept des **Ereignisses** beeinflusst das Konzept der Dauer
- Das kann zu widersprüchlichen Annotationen führen

heureCLÉA

COLLABORATIVE LITERATURE EXPLORATION & ANNOTATION



heureCLÉA @Hamburg

Jan Christoph Meister

Evelyn Gius

Janina Jacke

Marco Petris

heureCLÉA @Heidelberg

Michael Gertz

Thomas Bögel

Jannik Strötgen

Hilfskräfte:

Reza Kai Ghamsari, Rabea Kleymann,

Nicola Meier, Venera Musina,

Judith Niehaus, Nina Reder,

Hilke Scholz, Till Weinrich

Hilfskraft: Christoph Mayer

www.heureclea.de

gefördert im Rahmen des
eHumanities-
Förderprogramms des BMBF



- Fragen und Kommentare?

Eine kurze Hands-on-Einführung:

ERSTE SCHRITTE IN CATMA

Vorbereitung

- Login in CATMA: www.digitalhumanities.it/catma
- Download der XML-TEI-Versionen von Helene Langes Artikel „Frauenwahlrecht“ sowie Clara Zetkins Beitrag „Das Frauenstimmrecht“ aus dem Deutschen Textarchiv:
 - http://www.deutschestextarchiv.de/book/show/lange_frauenwahlrecht_1896
 - http://www.deutschestextarchiv.de/book/show/zetkin_frauenstimmrecht_1907
- Sharing-Interessierte: Angabe der Login-Emailadresse (optional)