



Annotation als Markup *avant la lettre*

Jan Christoph Meister

1 **Status quo: Das methodologische Delta**

In der Textthermeneutik als zentralem und für das traditionelle Selbstverständnis der Philologien wesentlichem Arbeitsbereich spielte der Einsatz digitaler Verfahren bislang kaum eine Rolle.¹ Zwar profitiert auch die literaturwissenschaftliche Auslegung von Textbedeutungen mittelbar durchaus vom *Digital Turn*: Software und webbasierte Kollaborationsumgebungen können den philologischen Arbeitsprozess von der Edition bis zur Lektüre des Einzeltextes unterstützen; Kollokationen und Datenbanken werden zu Recherchezwecken genutzt; Verfahren des *Data Mining*, der *Computational Stylometry* und des *Topic Modeling* bieten

¹Die nachfolgend skizzierte Bestandsaufnahme basiert u. a. auf bisherigen Erhebungen im Zeitraum 2014–2015, die im Rahmen von DARIAH-DE AP 5.4/Teilprojekt *Stakeholdergesellschaften* in Abstimmung mit dem Fachverband *DHd – Digital Humanities im deutschsprachigen Raum* (www.dig-hum.de [letzter Aufruf 8.1.2016]) vorgenommen wurden. Ziel dieses (2016 abgeschlossenen) Teilprojekts war „die Ermittlung der Bedarfe an digitalen Tools und Services, die sich aus der Perspektive von Fachwissenschaftlern konkret im Kontext ihrer je spezifischen Forschungsfragen wie Fachkulturen ergeben. Dieser nutzerzentrierte Ansatz soll sowohl zur Formulierung von Entwicklungsdesideraten und Use-Case-Szenarien beitragen wie das Interesse und die Akzeptanz der Wissenschaftler in den verschiedenen geisteswissenschaftlichen Fachcommunities für das Methodenparadigma des Big-Data-Approach fördern“ (Meister u. a., *DARIAH-DE Folgeantrag, Arbeitspaket 5.4*, internes Dokument 2013).

In den Beitrag sind Teile des DFG-LIS-Antrages *forTEXT* eingegangen. Die im Anhang beigefügte Sichtung und Evaluation von Annotationsplattformen wurde im Rahmen der Antragstellung von Marco Petris erarbeitet.

J. C. Meister (✉)
Universität Hamburg, Hamburg, Deutschland
E-Mail: jan-c-meister@uni-hamburg.de

neue Möglichkeiten für die Exploration großer Textkorpora mit Blick auf die Verbreitung und Genese inhaltlicher wie formaler Textmerkmale und eröffnen so nicht nur für die empirische Validierung von traditionell erarbeiteten Forschungshypothesen, sondern auch als Heuristiken neue Perspektiven, die für die Texthermeneutik durchaus relevant sein können. Aber wo es um die Textinterpretation als Bedeutungsauslegung im klassischen Sinne geht, finden computergestützte Ansätze bislang kaum Verwendung.²

Instrumentell bei vorbereitenden Verfahrensschritten durchaus in Anspruch genommen, ist das digitale Methodenparadigma dem philologischen Kerngeschäft der Texthermeneutik konzeptionell bislang suspekt geblieben. Diese Skepsis ist durchaus verständlich, stellt man die Spezifik des geisteswissenschaftlichen Wissensbegriffs selbst in Rechnung, der seit der Romantik durch zwei Leitideen bestimmt wird:

1. Historizität des Forschungsgegenstandes. – Das Erkenntnisobjekt ‚Text‘ wird in der hermeneutischen Praxis grundsätzlich als ein in stetig wandelbare soziale und kulturhistorische Kontexte eingebettetes Phänomen aufgefasst.
2. Diskursivität und Unabgeschlossenheit des hermeneutischen Erkenntnisprozesses. – Die Exploration und Auslegung von Textbedeutungen werden als rekursiv-iterative Prozesse praktiziert, die aufgrund der Dynamik ihrer Kontexte prinzipiell infinit sind; es gibt keinen hermeneutischen *Finite State* und keine absolut ‚wahre‘ Interpretation eines gegebenen Textes, sondern eine Bandbreite von mehr oder weniger plausiblen und anschlussfähigen Interpretationsvarianten.

Beide Leitideen, die das Selbstverständnis der Literaturwissenschaften nachhaltig geprägt haben, sehen traditionell arbeitende Philologen infrage gestellt von einer Technologie, deren Anwendung auf genuin hermeneutische Fragestellungen scheinbar zwingend einen reduktionistischen Begriff von hermeneutischer Erkenntnis als formal operationalisier- und technologisch (re-)produzierbarem Resultat voraussetzt.

²Von der grundsätzlichen Problematik zeugen nicht zuletzt die mehrfach gescheiterten Versuche, in den Geisteswissenschaften multifunktionale virtuelle Forschungsumgebungen zu etablieren, die mehr als nur eine vorab genau spezifizierte Aufgabenstellung (Texteditorik; Konkordanzerstellung etc.) bewältigen können. Weder die *bottom-up* ansetzende Strategie von Project Bamboo (<http://www.projectbamboo.org/> [letzter Aufruf 8.1.2016]), die von einer breiten Bedarfsanalyse und Befragung der geisteswissenschaftlichen Nutzercommunities ausging (vgl. zu den Problemen des Projekts Quinn Dombrowski, „What Ever Happened to Project Bamboo?“, in: *Literary and Linguistic Computing* 29/3 [2014], 326–339), noch die spezifischer bei den Bedarfen einer definierten Nutzercommunity ansetzenden Projekte wie etwa *NINES* (*Nineteenth-Century Scholarship Online*, <http://www.nines.org> [letzter Aufruf 8.1.2016]) oder die ambitionierten *top-down*-Strategien nationaler und multinationaler Infrastrukturvorhaben wie *DARIAH* haben bislang breite Akzeptanz gefunden.

Diese grundsätzliche Problematik wird dabei nicht selten noch verschärft durch die Fehleinschätzung von Potenzial und Anspruch gegenwärtiger digitaler Verfahren und Modellierungen. Die Schimäre ‚künstliche Intelligenz‘ spielt dabei ebenso eine Rolle wie das wachsende Unbehagen angesichts von Big-Data-basierten Methoden, die mit opaken statistischen Verfahren probabilistische Erkenntnis und Prognosen menschlichen Verhaltens generieren.³ Im Vordergrund der gesellschaftlichen Debatte steht dabei in der modernen Informationsgesellschaft die Sorge um die Freiheit und die Privatsphäre des Individuums als eines autonomen Agenten im sozialen Gefüge. Im Hintergrund wird hingegen eine moralphilosophische Problemstellung aufgeworfen, die auf dem Umweg über den ästhetischen Gegenstand auch und gerade die Geisteswissenschaften betrifft: Je treffsicherer individuelles Handeln und Verhalten mathematisch modellierbar werden, desto tiefgreifender wird das seit der Aufklärung propagierte Ideal vom Menschen als einem rationalen, prinzipiell selbstbestimmt und damit selbstverantwortlich handelnden Agenten infrage gestellt. Gerade dieses Selbstbild des Menschen in seinem Verhältnis zu Geschichte und Umwelt aber ist seit der Neuzeit das Kernthema des geisteswissenschaftlichen Gegenstandes ‚Literatur‘.

Die traditionelle philologische Textthermeneutik sieht sich so angesichts der *DH* mit einem methodologischen Paradoxon konfrontiert: Ausgerechnet ihre privilegierte Forschungsdomäne – der ästhetisch-literarische Diskurs – soll nun mit Methoden untersucht werden, die das Geistesleben nicht mehr als existenziellen, kontingenten Phänomenkomplex, sondern als prinzipiell berechen- oder zumindest modellierbaren ‚Output‘ einer multifaktoriellen Funktion konzeptualisieren. Die Skepsis methodisch konservativer Philologen wird dabei weiter befördert durch die programmatische Selbstbeschränkung und Fokussierung der *DH* auf deskriptive oder textdatenbasiert-analytisch bearbeitbare Fragestellungen – oder pointierter gesagt: durch eine Art selbstverordnete hermeneutische Abstinenz der *DH*. Denn Fakt ist: Das Spektrum wie der Grad der methodischen Ausgereiftheit von digitalen Editionsprojekten, stilometrischen Untersuchungen oder diachronen wie systematischen Form-, Struktur- und Themenanalysen mit avancierten NLP-Verfahren haben sich seit der Jahrtausendwende zwar deutlich erweitert. Zu dieser rasanten Entwicklung des Forschungsfeldes *DH* beigetragen haben dabei methodische, technologische, institutionelle und in ganz besonderem Maße auch pädagogisch-curriculare Impulse, kurz: die Gesamtheit der Faktoren, die eine wissenschaftliche Disziplin ausmachen. Der Anschluss an genuin *hermeneutische* Fragestellungen ist hingegen vonseiten der *DH* bislang nur in Ausnahmefällen gesucht worden. Die Praxis digitaler Literaturwissenschaften hat durch die professionelleren, robusteren Verfahren der Datenanalyse und Datenmodellierung vielmehr ein vorwiegend empirisch-analytisches Gepräge gewonnen

³Viktor Mayer-Schönberger/Kenneth Cukier, *Big data. A revolution that will transform how we live, work, and think*, Boston 2013.

und sich damit immer weiter von den spekulativ-exemplarisch verfahrenen Methoden traditioneller Literaturwissenschaften entfernt.⁴

Anders als etwa in den Sprachwissenschaften trifft deshalb das stetig erweiterte Angebot an digitalen Ressourcen, Forschungsinfrastrukturen und immer robuster werdenden analytischen und modellierenden Methoden im texthermeneutischen Bereich auf eine nach wie vor geringe Nachfrage. Digitale Verfahren setzen grundsätzlich eine Konzeptualisierung des Forschungsgegenstandes voraus, die ausschließlich solche Gegenstands- und Prozess-Eigenschaften berücksichtigt, die in Form diskreter Daten erfassbar sind. Gegenstandsfunktionen lassen sich dank dieser Beschränkung damit formal als ein Effekt der Transformation und Verknüpfung von Daten modellieren. Was jedoch jenseits dieser epistemischen Schwelle liegt, ist für *DH*-Verfahren, solange sie primär informationstechnisch orientiert sind, nicht mehr operationalisierbar. Die methodische Konsequenz liegt auf der Hand: Ein Gegenstandsmodell, das beim diskreten Datum ansetzt und von dort aus strikt induktiv und ohne methodische Iteration oder Rückkoppelung *bottom-up* voranschreitet, kann phänomenologisch-hermeneutischen Problemstellungen nicht gerecht werden, die in der Regel ideengeleitet und *top-down*, darüber hinaus aber auch dynamisch und historisch-kontextuell entworfen sind.

Und doch lassen sich auch genuin hermeneutische Operationen des Erschließens, Verstehens und Interpretierens von Gegenständen durchaus mit *DH*-Verfahren unterstützen. Eine zentrale Rolle spielen dabei Verfahren der deklarativen Textauszeichnung und explikativen Textanreicherung, die in den *DH* unter den technischen Begriffen ‚Tagging‘ und ‚Markup‘ verhandelt werden. Das grundsätzliche hermeneutische Potenzial dieser Verfahren lässt sich allerdings erst erkennen, wenn wir die von der Computerlinguistik und der angewandten Informatik übernommene, vom elementaren String her und von vornherein mit Blick auf eine wohlgeordnete Syntax und Grammatik entworfene Konzeption des deklarativen wie des prozessualen Markups überschreiten. Es reicht dabei nicht, zähneknirschend die prinzipielle Interpretationshaltigkeit selbst einer grundlegenden schemakonformen Textauszeichnung mit einem Tag zu konzedieren. Ja, dieses Zugeständnis läuft geradezu auf eine Trivialisierung des Begriffes der ‚Interpretation‘ hinaus: In einem vorgegebenen Kategoriengerüst zwischen zwei

⁴Eine Stichprobe: Für den Zeitraum 2005–2015 werden in GEPRIS unter dem Suchbegriff ‚Philologie‘ 128 DFG-geförderte Projekte ausgewiesen; nur sieben dieser Projekte (= 5,5 %) weisen zugleich das Suchwort ‚digital‘ auf. (Eine Suchanfrage für ‚Literatur‘ ergibt 2825 Treffer; ‚Literatur & digital‘ ergibt 49 Treffer: eine Quote von 1,7 %.) Aus den online verfügbaren Kurzbeschreibungen geht hervor, dass nur 15 der beschriebenen Projekte einen *DH*-Bezug aufweisen. Von diesen 15 Projekten fallen wiederum zwölf in die Kategorie ‚Digitalisierung/Datenbank‘ und nur drei in die Kategorie ‚Analyse/Annotation‘. Die Suche nach ‚Digital Humanities‘ innerhalb der Literaturwissenschaft ergibt hingegen acht Treffer, wobei jedoch auch hier nur eines dieser Projekte u. a. hermeneutisch arbeitet. Einen ähnlichen Befund ergibt umgekehrt die Sichtung der Tagungsprogramme der regionalen und internationalen *DH*-Fachkonferenzen. Erfasst wurden dabei die Veranstaltungen, die im Zeitraum 2000–2017 von den maßgeblichen Verbänden initiiert wurden, die in der *Alliance of Digital Humanities Organisations (ADHO)* kooperieren.

oder mehr Optionen zu wählen, um ein Phänomen taxonomisch zu deklarieren, ist ein Akt des Sortierens und keiner der Interpretation. Interpretative Akte sind Akte der Kritik, die nicht nur den Gegenstand vermessen und bewerten, sondern immer zugleich reflexiv den Akt des Interpretierens selbst und dessen normative wie methodische Setzungen und Grenzbedingungen thematisieren.

Das Potenzial und die Anschlussfähigkeit von *DH*-Verfahren auch für die hermeneutische Literaturwissenschaft zu erkunden, setzt insofern voraus, die in den *DH* unter dem Terminus ‚Markup‘ etablierten Praktiken der Textauszeichnung mit einem anspruchsvolleren Maßstab zu konfrontieren und sie einmal konsequent vom komplexen, dynamischen Ganzen des Text-Leser-Kontext-Gefüges her zu denken. In dieser Perspektive erweist sich eine Markup-Praxis, die vom lokalen Tagging eines Strings maximal allenfalls bis zur schemakonformen Deklaration eines Gesamttextes reicht, als funktional eingeschränkte, dafür jedoch besser formalisierte technische Variante einer deutlich ambitionierteren ‚weichen‘ Textpraxis, die allen Schriftkulturen seit Jahrhunderten vertraut ist: die der *Annotation von Texten im emphatischen Sinne von hermeneutisch-kommunikativ wie rhetorisch motivierter Informationsanreicherung*. Markup als Annotation in diesem emphatischen Sinne zu konzeptualisieren, bedeutet dabei zugleich, es als ‚methodischen Mittler‘ zu begreifen, der die Anschlussfähigkeit digitaler Verfahren auch und gerade für traditionell-hermeneutisch orientierte Literaturwissenschaftler unter Beweis stellen kann.

2 Annotation als Kommunikation von Lesarten⁵

Annotation als Kulturtechnik ist – methodengeschichtlich gesehen – nicht nur historisch vorgängig, sondern auch variantenreicher als das, was wir gegenwärtig unter den Begriff ‚Markup‘ rechnen – aus der *DH*-Perspektive gesehen mithin eine Praxis *avant la lettre*, die schon lange vor dem *Digital Turn* je nach den spezifischen Anforderungen von Textsorten und Gebrauchskontexten (u. a. religiöse, juristische und ästhetische Exegese) spezifiziert und teilweise auch formalisiert worden ist. Als informationstechnologische Spezifikation stellt Markup insofern nur die jüngste Ausprägung eines übergreifenden Annotationskonzepts dar – eine technisch wie methodisch bedingte Spezifikation, die möglicherweise nicht die letzte bleiben wird.

Die wissenschaftliche Kontextualisierung von Markup als bislang letzte Ausprägung einer lange etablierten Kulturtechnik ist bislang weder ins Bewusstsein der traditionellen Literaturwissenschaften getreten noch umgekehrt von den *DH* als Herausforderung zur Entwicklung eines anspruchsvolleren digitalen Annotationskonzepts angenommen worden. Die im Wesentlichen technologiegetriebene

⁵Ich verwende den Begriff ‚Lesart‘ hier wie nachfolgend im allgemeineren Sinne von ‚Auslegung‘ oder ‚Auffassung‘, also nicht im Sinne des enger gefassten editionsphilologischen Terminus, der den Wortlaut einer Textstelle gemäß eines spezifischen Textzeugens bezeichnet.

Konzeptualisierung von Markup als bloße Anreicherung von Objekten mit deklarativen Metadaten geht vielmehr weiterhin von einem erheblich engeren, instrumentell gefassten Begriff der ‚Anreicherung‘ aus. Als Markup in diesem engeren Sinne gelten systematisch geregelte Beschreibungen des Gesamttextes wie lokale Textauszeichnungen auf Wort- und Zeichenebene, die unter Rückgriff auf Standards, Taxonomien und Objektontologien und in der Regel für definierte Anschlussoperationen erstellt werden. Für die Objektbeschreibung auf elementarer Ebene sind prototypisch die Verfahren der linguistischen Basisannotation wie z. B. *POS Tagging*; für die systematische wie historische Einordnung des Gesamtdokumentes sind es etwa die Hinzufügung von Metadaten in einem TEI-Header. Generische Schemata wie GML, SGML, HTML bis hin zu TEI und den diversen heute gebräuchlichen XML-basierten domänenspezifischen Markup-Sprachen setzen dabei konzeptionell ein zwar abstraktes, aber klar definiertes Objektmodell voraus – sei es das grammatische Modell einer natürlichen Sprache, das jedes *Part-of-Speech Tagging* organisiert, sei es das einer *Ordered-Hierarchy-of-Content-Objects-Sicht (OHCO)* auf Textdokumente, das bei der TEI-konformen Textauszeichnung im Hintergrund steht. In der konkreten Annotationspraxis wird dieses abstrakte Objektmodell dann mithilfe einer geregelten Syntax am konkreten Objekt *bottom-up* instanziiert. Das Objektmodell selbst steht dabei prinzipiell nicht mehr zur Debatte – es sei denn, man stößt bei der Anwendung auf grundsätzliche Probleme, wie die 1999 mit einer Kontroverse zwischen Renear, McGann und Buzzetti einsetzende Debatte um die Unvereinbarkeit des *OHCO*-Modells mit dem Faktum der sog. *Overlapping Hierarchies* gezeigt hat.⁶

Coombs u. a.⁷ zählen zu den ersten, die die auf GML aufbauenden konkreten technischen Umsetzungen der taxonomie- und modellgetriebenen Markup-Praxis systematisch und funktional differenziert haben. Sie unterscheiden zwischen den Formen *Punctuational*, *Presentational*, *Procedural*, *Descriptive*, *Referential* und *Meta-Markup*. Als für den menschlichen Leser besonders relevante Variante heben die Autoren dabei das sog. *Descriptive Markup* hervor.⁸ Ähnlich weist 25 Jahre später

⁶Siehe hierzu Susan Hockey/Jerome J.McGann/Allen Renear, „What is text? A debate on the philosophical and epistemological nature of text in the light of humanities computing research“, in: *1999 ACH-ALLC Conference Program*, <http://www2.iath.virginia.edu/ach-allc.99/proceedings/hockey-renear2.html> (letzter Aufruf 7.5.2018). Mit diesem Beitrag rückte die Thematik auf der ACH-ALLC-Konferenz 1999 erstmals ins Zentrum der Aufmerksamkeit. Dino Buzzetti trug schon auf der gleichen Konferenz mit dem Vortrag „Text Representation and Textual Models“ explizit zur Debatte bei, in: *1999 ACH-ALLC Conference Program*, <http://www2.iath.virginia.edu/ach-allc.99/proceedings/buzzetti.html> (letzter Aufruf 7.5.2018; im Beitrag „Digital Representation and the Text Model“, in: *New Literary History* 33 (2002), 61–88, legte Buzzetti seine Argumentation dann umfassender dar.

⁷James H. Coombs/Allen H. Renear/Steven J. DeRose, „Markup systems and the future of scholarly text processing“, in: *Communications of the ACM* 30 (1987), 933–947.

⁸Siehe jedoch Allen Renear, „The descriptive/procedural distinction is flawed“, in: *Markup Languages* 2/4 (2000), 411–420, der später feststellt: „the descriptive/procedural distinction is flawed“.

auch Nyhan darauf hin, dass die nicht bereits auf spezifischere Anschlussoperationen der maschinellen Transformation oder Präsentation eines Dokumentes zugeschnittene Variante einer beschreibenden Annotation flexibel und auf verschiedenste Textsorten, Genres und Medien übertragbar ist: „Descriptive markup can be applied to any kind or genre of text; indeed, any information that can be consistently represented using a symbol of some kind and then digitized can be marked up“.⁹

Nochmals zu betonen ist allerdings, dass das ‚beschreibende‘ Markup, von dem hier die Rede ist, *de facto* nicht mehr als ein kategorial sortierendes, allenfalls taxonomisch evaluierendes Verfahren ist. Das Beschreibungsschema, die Beschreibungskategorien sowie der Typ und das Spektrum der zulässigen Werte, die dem jeweils markierten Objektbereich – einem Morphem, Wort, Satz, Absatz etc. – zugewiesen werden können, sind geregelt – und nicht *ad hoc* erweiter- oder veränderbar. Deklariert wird vielmehr anhand einer vorgegebenen Systematik und Metrik, die zudem gegenüber etwaigen Forschungsfragen agnostisch ist und dies auch sein muss, um überhaupt generalisierbar sein zu können. Ein Textdokument in TEI zu beschreiben oder eine verschriftlichte sprachliche Äußerung zu parsen und mit *POS Tags* zu versehen, läuft so, genau besehen, darauf hinaus, das jeweils selektierte Textelement in eine vorgegebene Tabelle einzusortieren und einen (oder mehrere) Werte aus einem finiten alphanumerischen Wertebereich zuzuordnen.

Zu erinnern ist hier an die von Gadamer getroffene Unterscheidung zwischen dem Verfahren des ‚Beschreibens‘, das die empirischen Wissenschaften anwenden, und jenem des ‚Erklärens‘, das für die Geisteswissenschaften zentral ist. Wittgenstein liefert mit einem Satz in seinen *Philosophischen Untersuchungen* ein gutes Beispiel, was geschieht, wenn man diese Distinktion zur programmatischen Norm erhebt. Dort heißt es zunächst: „Alle Erklärung muß fort, und nur Beschreibung an ihre Stelle treten“;¹⁰ Deskription soll in den Wissenschaften herrschen, nicht Interpretation.¹¹ Aber im Nachsatz räumt Wittgenstein dann ein: „Und diese Beschreibung empfängt ihr Licht, d. i. ihren Zweck, von den philosophischen Problemen“.

Auch wenn man im Grundsatz eine klare und bewusst praktizierte Trennung der beiden Operationen des Beschreibens und des Erklärens auch von der Philologie fordern sollte, muss man zur Kenntnis nehmen, dass in der Praxis der wissenschaftlichen Texthermeneutik die zweckorientierte, philosophisch motivierte ‚Beschreibung‘ des Gegenstandes im emphatischen Sinne immer auf eine ‚Erklärung‘ hinauslaufen wird – mithin auf eine Interpretation, die als Verfahren der Gegenstandserschließung nicht nur auf den Gegenstand selbst gerichtet ist, sondern zugleich reflexiv und dynamisch eben dieses Verfahren der

⁹Julianne Nyhan, „Text encoding and scholarly digital editions“, in: Dies./Claire Warwick/Melissa Terras, *Digital Humanities in Practice. Facet Publishing*, London 2012, 117–138, hier: 123.

¹⁰Ludwig Wittgenstein, *Philosophische Untersuchungen*, Frankfurt a. M. 1975 [1953], 78 f.

¹¹Zur Anwendung Wittgensteins auf die Problematik der Literaturtheorie siehe u. a. Henry Staten, „Wittgenstein and the Intricate Evasions of ‚Is‘“, in: *New Literary History* 19/2 (1988), 281–300.

Erschließung selbst thematisiert. Diese unkalkulierbaren Rückkoppelungseffekte sind nun jedoch genau das, was ein standardisiertes, auf Disambiguierung verpflichtetes Markup auszuschalten versucht.

Der methodologische Reduktionismus, der dem deklarativen ebenso wie dem prozeduralen oder repräsentationellen Markup insofern inhärent ist, hat in den *DH* mehrfach die Forderung nach der Entwicklung eines interpretativen oder hermeneutischen Markupkonzepts ausgelöst. Piez fordert in diesem Sinne prononciert ein „markup that is deliberately interpretive“; diese Art Markup sei dabei „not limited to describing aspects or features of a text that can be formally defined and objectively verified. Instead, it is devoted to recording a scholar’s or analyst’s observations and conjectures in an open-ended way.“¹² Wie schon Caton hervorgehoben hat, ist insofern die Frage der Wahl eines Markupschemas in der hermeneutischen Perspektive eine nachgeordnete; ausschlaggebend ist vielmehr das Textkonzept an sich:

When OHCO-1 encourages encoders to see a written text as a thing, they stay above the content and only drop down to engage with the text as message to identify the occasional editorial object whose nature is not obvious from its appearance. But when encoders see the written text as a communicative act, they must participate in the act: take on the role of hearer, attend to what the text says, and identify the speaker’s intentions not just from the words’ semantics but also from the attitudes conveyed. Metaphorically, encoders must be down at what would be the lowest level of an OHCO tree, completely immersed in the #PCDATA, because content generates interactional encoding far more than any content object. As its practitioners well know, all encoding interprets, all encoding mediates. There is no ‚pure‘ reading experience to sully. We don’t carry messages, we reproduce them – a very different kind of involvement. We are not neutral; by encoding a written text we become part of the communicative act it represents.¹³

Das hier skizzierte Modell eines *Interactional Encodings* zieht einen konzeptionell erheblich weiter gefassten Begriff von ‚Markup‘ nach sich, als er mit einer am Ideal des perfekten *Inter Annotator Agreement* ausgerichteten Auszeichnungspraxis einhergeht – einer Praxis, die aus der Linguistik importiert wurde und damit in der Tradition der Bemühungen um eine Verwissenschaftlichung der Literaturwissenschaften steht, die schon zu Beginn des 20. Jahrhunderts mit dem Russischen Formalismus einsetzte.¹⁴ In den methodengeschichtlich bislang eher desinteressierten *DH* unserer Tage wird dieses Ideal des non-ambigen Enkodierens und Auszeichnens allerdings weniger programmatisch als vielmehr pragmatisch legitimiert: *Gold Standards* spielen für die Automatisierung und Validierung von Markup-Routinen mit Verfahren des ML bekanntlich eine zentrale Rolle.

¹²Wendell Piez, *Towards Hermeneutic Markup. An architectural outline* (2010), <http://dh2010.cch.kcl.ac.uk/academic-programme/abstracts/papers/html/ab-743.html> (letzter Aufruf 9.11.2017).

¹³Paul Caton, „Markup’s Current Imbalance“, in: *Markup Languages 3/1* (2000), 1–13.

¹⁴Ein Faktum, dem interessanterweise Rumelhart als Pionier der textbezogenen AI-Forschung der 1970er Jahre bereits implizit Rechnung getragen hat, indem er wiederholt auf Propps *Morphologie des Zaubermärchens* zurückgriff – wenn auch in einer Lesart, die bezeichnenderweise das Konzept der Morphologie, das Propp mit einem seinem Buch vorangestellten Goethe-Motto geistesgeschichtlich kontextualisiert hatte, auf einen generativen Algorithmus reduziert.

Diese in der Computerlinguistik, der NLP-Forschung und der digitalen Texteditorik verbreitete Konzeption einer am Ideal der Eindeutigkeit orientierten Textauszeichnung wird selbst in jüngeren *DH*-Handbüchern und Einführungen weitgehend alternativlos präsentiert.¹⁵ In dieser Verengung der Perspektive drückt sich eine methodische Problematik aus, die van Zundert jüngst als die der – bislang ungenügend berücksichtigten – *Computationality of Hermeneutics* umreißt.¹⁶ Van Zundert verbindet seinen historischen Abriss der in den *DH* zur Zeit nur noch am Rande geführten Forschungsdebatte zur Frage des methodologisch-philosophischen Verhältnisse zwischen den hermeneutischen Wissenschaften und der Informatik entsprechend mit einem Plädoyer, das hermeneutische Anliegen nicht länger nur „after the algorithmic fact“ zu thematisieren. Gefordert sei vielmehr, hermeneutische Desiderata bereits auf der Ebene der Konzepte, Codes und Modelle zu verhandeln, die für die digitale Repräsentation und Analyse von lebensweltlichen Phänomenen und ästhetischen Artefakten grundlegend sind.

Was aber wäre ein solches der Texthermeneutik wie der Informatik gemeinsames und damit für die *DH* kennzeichnendes hermeneutisches Anliegen ‚before the algorithmic fact‘? Der allen Varianten metasprachlicher Äußerung gemeinsame rhetorische Gestus ist der einer Kommunikation einer (oder mehrerer) Lesart(en) eines Objekttextes – und sei es nur in der Form eines -Tags. Eine typographische Lesarten-Kommunikation mag zwar vordergründig nur noch die an eine Maschine gerichtete Instruktion sein, eine Zeichenfolge zu fetten. Aber selbst diese vermeintlich nur prozessuale Codierung ist ihrerseits Mittel zum Zweck: Auch sie richtet sich in letzter Instanz wieder an einen menschlichen Leser und ist ursprünglich diskurspragmatisch legitimiert.

Die hier vorgeschlagene diskurspragmatische Konzeptualisierung von Annotation als Kommunikation von Lesarten lässt zwei Grenzbedingungen erkennen: Annotation ist einerseits notwendig eine Form von Metatext, eine Rede über die Rede – wo sie sich vom Bezugstext ablöst und autonom wird, nimmt sie ihrerseits objektsprachlichen Charakter an und wird zur selbstständigen Sachverhaltsbehauptung. Andererseits verliert Annotation ihren kennzeichnenden Kommunikationscharakter aber auch dort, wo Metasprache zur reinen Privatsprache wird. Ein Beispiel dafür liefert das jüngst von McGann vertretene Konzept des sog. ‚topologischen‘ Markups.¹⁷ McGann kritisiert das TEI-Modell vor allen Dingen deshalb, weil es die autopoetische Funktion von ästhetischen Texten ausklammere zugunsten einer allopoetischen Funktion, die Bedeutung nur noch als

¹⁵ Siehe Andrea Rapp, „Manuelle und automatische Annotation“, in: Fotis Jannidis/Hubertus Kohle/Malte Rehbein (Hg.), *Digital Humanities. Eine Einführung*, Stuttgart 2017, 253–267, hier: 255 f.; Allen Renear, „Text Encoding“, in: Susan Schreibman/Ray Siemens/John Unsworth (Hg.), *Companion to Digital Humanities*, Oxford 2004.

¹⁶ Joris J. van Zundert, „Screwmenetics and Hermenumericals. The Computationality of Hermeneutics“, in: Susan Schreibman/Ray Siemens/John Unsworth (Hg.), *A New Companion to Digital Humanities*, Chichester, West Sussex, UK, u. a. 2016.

¹⁷ Jerome McGann, Jerome, „Marking texts in many dimensions.“, in: *A new republic of letters: Memory and scholarship in the age of digital reproduction*, Cambridge, MA, 2014, 90–112.

ein dem Text externes Output denkt, nicht aber als das dynamische Produkt eines Text-Leser-Komplexes *in actu*. Sein Gegenentwurf geht deshalb von einer als *Patacriticism* bezeichneten ‚theory of subjective interpretation‘ aus – und läuft in seiner Radikalität dann auf eine Annotation zum Zwecke der Autokommunikation des Annotators hinaus: eine Privatsprache, die keine Diskurspartner mehr sucht und den annotierten Text als Impulsgeber eines Selbstgesprächs begreift.¹⁸

Die von mir vorgeschlagene diskurspragmatische Bestimmung von Annotation als Kommunikation von Lesarten positioniert die Annotation hingegen als eine Art hermeneutischen Mediator, der in der Mitte zwischen dem Objekttext und seinem Gebrauchskontext funktional wird. Dies entspricht der von Boot unter dem Begriff *Mesotext* diskutierten epistemischen Funktion von Annotationen als einem Mittler zwischen dem (Objekt-)Text und dem medial eigenständigen Metatext.¹⁹ Wichtig ist dabei insbesondere Boots Stufung *Textdaten – Mesodaten – Mesotext*, d. h. der Hinweis darauf, dass auch Annotationen immer bereits unter zwei Voraussetzungen generiert und verarbeitet werden: erstens einer der Annotationspraxis inhärenten Texttheorie, die eine deskriptive bzw. explikative Modellierung des Objekts ‚Text‘ überhaupt erst ermöglicht, und zweitens einer Ausrichtung auf ein diskursives ‚Narrativ‘, das die per Annotation formal erfassten einzelnen Beobachtungen, Erläuterungen und lokalen Interpretationen zum Objekttext sinnhaft integriert, mithin aus den isolierten Mesodaten überhaupt erst Information und damit einen Mesotext macht.

Für eine hermeneutisch ambitionierte *DH*-Praxis wäre insofern ein Begriff von ‚Markup‘ zu fordern, der nicht mehr vorwiegend technologisch, sondern vielmehr diskurspragmatisch gedacht ist. Wenn wir Markup als *DH*-spezifische Ausprägung von ‚Annotation‘ begreifen wollen, dann muss das Markup-Konzept auch Praktiken jenseits der bloßen Codierung oder Deklaration von Objektdaten einschließen – denn nur dann können auch die noch nicht auf spezifische informationstechnische Anschlussoperationen hin orientierten philologischen Verfahren der Anreicherung von textuellen Primärdaten mit Zusatzinformationen berücksichtigt werden, die einer hermeneutisch motivierten Kommunikation von Lesarten dienen. Ob und vor dem Hintergrund welcher Modelle diese Lesarten dann ihrerseits evaluiert werden, bleibt dabei eine Frage des spezifischen Anwendungsfalls und seiner Erfordernisse. Widerspruchsfreiheit und Konsistenz sind insofern keine intrinsischen, formalen Kriterien von Annotationen, die als solche normative Gültigkeit beanspruchen könnten – es sind dies vielmehr aus dem Anwendungskontext heraus definierte Anforderungen. Entsprechend wertneutral wäre deshalb auch das Konzept des ‚Markups‘ zu fassen, dessen Maßstab nicht eine abstrakte logische Spezifikation sein kann, sondern nur das aus einem Anwendungskontext heraus formulierte und vor dem Hintergrund einer jeweiligen Texttheorie spezifizierte Markupschema.

¹⁸Das McGanns Artikel zur Illustration beigefügte Beispiel einer autopoetisch inspirierten Annotation und Lektüre liest sich dann auch wie ein *Remake* von Roland Barthes’ *S/Z*.

¹⁹Peter Boot, *Mesotext. Digitised Emblems, Modelled Annotations and Humanities Scholarship*, Utrecht 2009.

3 Funktionen und Formen von Annotationen

Nahezu sämtliche Varianten philologisch fundierter Lesartenhinweise an den Leser können heute als Verfahren digital emuliert und als Mitteilung digital codiert werden. Als wissenschaftliche Operation ist die Textannotation dabei funktional für drei der sog. *Scholarly Primitives* nach Unsworth: Objektkonstitution, Objektbeschreibung und Objektannotation.²⁰ Im engeren Anwendungsbereich der Literaturwissenschaften lassen sich analog drei Komplexitätsebenen unterscheiden:

- *Textauszeichnung*: Auszeichnung von Phänomenen, die keine inhärente oder robust konventionalisierte Semantik aufweisen. Dazu gehören sowohl rein formale Phänomene wie etwa Layout und Typographie als auch strukturell-grammatische Phänomene. Vor der Auszeichnung liegt hier i. d. R. ein eindeutiges Auszeichnungsschema vor (linguistische Taxonomien, TEI etc.).
- *Textannotation*: Auszeichnung lokaler semantischer Phänomene. Diese Variante der Auszeichnung setzt bereits das Verstehen des semantischen Gehalts einer umfangreicheren Textstelle (und damit meistens deren Interpretation im weiteren Sinne²¹) voraus. Diese auf der mittleren semantischen Komplexitätsebene ansetzende Auszeichnung lässt sich im Anschluss an Piez²² als ‚hermeneutische Annotation‘ bezeichnen. Auch wenn in der Praxis der Textannotation qua Verwendung literaturwissenschaftlicher Fachterminologie oftmals eine Art implizite Taxonomie vorausgesetzt ist, wird diese deutlich weniger deterministisch gehandhabt als bei der schemabasierten Textauszeichnung.
- *Textkommentar*: Auszeichnung dezidiert literarischer Bedeutungsphänomene, die sich in der Regel nur durch eine holistische Interpretation des Gesamttextes erschließen. Dazu gehören beispielsweise die allegorische oder parabelhafte Auslegung von Textbedeutungen sowie nur interpretativ lösbare Fragen nach ‚Wahrheit‘ im Kontext fiktionaler Repräsentationen. Für die Auszeichnung derartiger Phänomene ist eine komplexe Interpretationstheorie für literarische Texte erforderlich und, sofern interpretationstheoretisch abgesichert, darüber hinaus auch die Referenz auf kontextuelle (d. h. textexterne) Informationen. Textkommentare sind insofern noch weniger deterministisch und deshalb kaum sinnvoll digital modellierbar.

Diese systematische Differenzierung ist entlang der Achse zunehmender semantischer und damit auch hermeneutischer Komplexität von Annotationen entworfen.

²⁰John Unsworth, *Scholarly Primitives. What methods do humanities researchers have in common, and how might our tools reflect this?*, London 2000.

²¹Siehe Janina Jacke, „Is There a Context-Free Way of Understanding Texts? The Case of Structuralist Narratology“, in: *Journal of Literary Theory* 8/1 (2014), 118–139, hier: 130 f.

²²Piez (Anm. 13).

Eine alternative Systematik schlagen Bauer/Zirker vor,²³ die teils unter dem Gesichtspunkt des Gegenstandsbereiches (Typen A und B) und teils unter dem der explikativen Relation einer Annotation zum Text (C, D, E) fünf Annotationstypen unterscheiden:

- A linguistic
- B formal
- C intratextual
- D contextual
- E interpretive

Beide Systematisierungen ermöglichen es, das gesamte Spektrum an ‚Lesarten-hinweisen‘ zu erfassen, die uns konkret in Form von Hervorhebungen, Unterstreichungen, in den Text direkt eingefügten Anmerkungen, Kommentaren in der Marginalienspalte usw. begegnen.

In einer dritten, auf den ersten Blick rein technischen Perspektive lassen sich Annotationstypen schließlich auch nach der medialen bzw. technischen Distanz einer Annotation zum jeweiligen Referenzbereich unterscheiden. Die im Printmedium geläufigen Formen der typographischen Hervorhebung wie z. B. Fettungen oder Unterstreichungen nutzen dabei das objektsprachliche Zeichen selbst als Träger der metasprachlichen Annotation. In den heute geläufigen SGML-basierten digitalen Varianten treten die objekt- und die metasprachlichen Dimensionen hingegen bereits auf dieser grundlegenden Ebene auseinander.²⁴ Das ist jedoch weder zwingend, noch wird dieses technische Detail in der Regel vom Normalleser eines digitalen Dokuments überhaupt bewusst verarbeitet. In der Perspektive des technisch interessierten Lesers hingegen läuft die in der digitalen Zeichenkette eingebettete Deklaration mittels *Inline Tag* auf eine doppelte Emulation hinaus: Erstens sichert das *Inline Tagging* die Eindeutigkeit der Objektreferenz auf eine quasi-analoge Weise durch die unmittelbare räumliche Nachbarschaft von Objekt- und Metazeichen; zweitens emuliert der digitale Text in den meisten Arbeitsumgebungen dann spätestens auf dem Bildschirm wieder die Verschmelzung der typographischen Annotation mit dem objektsprachlichen Zeichen.

Auf den zweiten Blick allerdings zeigt sich, dass die Gradierung von Annotationen nach ihrer medialen Distanz zum Objektzeichen bereits in der prädigitalen Textpraxis eine funktionale Semantisierung erfahren hat: Je weiter das

²³ Matthias Bauer/Angelika Zirker, „Whipping Boys Explained. Literary Annotation and Digital Humanities“, in: Kenneth M. Price/Ray Siemens (Hg.), *Literary Studies in the Digital Age. An Evolving Anthology*, New York 2013, <https://dlsanthology.mla.hcommons.org/> (letzter Aufruf 3.5.2018).

²⁴ Was durchaus keine technologische Notwendigkeit ist, sondern eine Konsequenz von SGML: Im Prinzip ließen sich mit einem Hexadezimal-Code sämtliche typografischen Varianten eines Zeichens auch einzeln codieren und dennoch zugleich auf ähnliche Weise als funktional ‚familienähnlich‘ behandeln, wie wir dies bei der digitalen Codierung von Farbwerten tun.

objektsprachliche Zeichen und die metasprachliche Annotation auseinandertreten, desto eher wird ein kompetenter Leser die Annotation als interpretationshaltigen Mesotext *sensu* Boot bewerten, der einem eigenständigen Metatext zuarbeitet. Im traditionellen Schrift- und Printmedium beginnt dieses Auseinandertreten mit der Interlinearannotation und reicht über die Formen der Glosse und der Anmerkung in der Marginalienspalte bis hin zur komplexen, kommentierenden Fuß- oder Endnote und möglicherweise gar noch bis zum Apparat. Die mediale und räumliche Distanz zwischen Objekttext und Annotat ist so zum konventionalisierten Indikator diskurspragmatischer bzw. rhetorischer Funktionalität geworden.

4 **Markup als Annotation jenseits des *Document Paradigm***

In der Perspektive der *DH* und mit Blick auf Markup als digitale Variante von Annotation stellt sich an dieser Stelle eine Grundsatzfrage: Alle Markuptechniken, die in der Tradition von SGML stehen, setzen konzeptionell auf die durchgehende Separierung von Objekttext und Metatext. Je deutlicher diese Trennung im Interface und im Workflow einer Markupumgebung ausgeprägt sind, umso sperriger wirkt diese daher für den traditionell arbeitenden Forscher, der Annotation als eine weitgehend unstrukturierte, intuitiv handhabbare Praxis der Anreicherung und Zwischenspeicherung von Anmerkungen, Erläuterungen, Kommentaren und Verweisen schätzt.

Die Suggestivität, Flexibilität und Reichhaltigkeit dieser ‚händischen‘ Annotationspraxis lässt sich zwar mit einigem Aufwand auch digital simulieren, wie John Bradley mit der Entwicklung der Arbeitsumgebung *Pliny* gezeigt hat.²⁵ Aber warum überhaupt an dem prädigitalen Muster festhalten und es zu emulieren suchen? Macht nicht gerade die Entkoppelung des digitalen Markups von der konventionalisierten Semantik ‚händischer‘ Annotation, die eine Anmerkung als umso selbstevidenter und ‚objektiver‘ suggeriert, je geringer die räumliche Distanz zwischen Objekttext und Annotat ist, einen *epistemischen Mehrwert* digitaler Annotationen aus?²⁶ Ist das, was auf den ersten Blick einer informationstechnischen Gegebenheit geschuldet zu sein scheint, nicht eigentlich Ausdruck einer prinzipiell anderen Form von Wissensorganisation und die logische Konsequenz eines neuen Textmodells? Welche neuen Formen, Funktionen und

²⁵ John Bradley, „Thinking about interpretation. Pliny and scholarship in the humanities“, in: *Literary and Linguistic Computing* 23 (2008), 263–279.

²⁶ Interessant ist hier die Frage nach dem funktionalen wie rhetorischen Status eines *inline* codierten Hyperlinks: Wird die Referenz vom Objekttext zum Zieltext damit epistemisch objektiviert und vom Leser als Quasi-Merkmal des Ursprungstextes verarbeitet, weil der Link auf dem Bildschirm mit seinen Trägerzeichen verschmolzen ist – oder wird umgekehrt die Verweisfunktion des Links gerade bewusst als solche wahrgenommen, weil sie erst durch eine Interaktion aktiviert werden muss, die das Trägerzeichen als Link interpretiert?

Praktiken literaturwissenschaftlicher, hermeneutisch ambitionierter Annotation könnten mit diesem Ansatz entstehen?

Eine der zentralen Debatten innerhalb der Markup-Community ist in dieser methodologischen Hinsicht die Kontroverse um die Verwendung von *Inline* vs. *Stand-off Markup*. *Inline Markup* in Form von in den Objekttext eingebetteten Tags ist bei der sequenziellen Verarbeitung von Dokumenten leichter handhabbar, konzeptionell den einfachsten Formen des analogen Verfahrens ähnlich und bietet zudem eine Reihe von Vorteilen bei der Datenverarbeitung (Performanz von Search- und Retrieval-Algorithmen, Stabilität der Referenz u. a. m.). Die Alternative des explizit referenzierenden *Stand-off Markups* bzw. *External Stand-off Markups*, in der das digitale Annotat konsequent vom Bezugstext abgelöst und separat gespeichert wird, kann hingegen besser komplexere Bezüge abbilden, die nicht mehr an die formale Strukturierung des Bezugstextes als lineare Zeichenkette gebunden sind.

Ein wesentliches Argument für die Verwendung von *Stand-off-Markup-Verfahren* ist bekanntlich die Kritik am hierarchischen Textmodell der *Ordered Hierarchy of Content Objects (OHCO)* gewesen, wie es etwa dem TEI-Schema zugrunde liegt. Das OHCO-Modell von ‚Text‘ ist entlang einer hierarchisch geordneten Systematik nach dem Muster ‚Buch (enthält Kapitel (enthält Absatz (enthält Satz (enthält Wort)))‘ entworfen; es lässt sich damit problemlos und performant mit einem *Inline* Markup abbilden, dessen Syntax die entsprechende hierarchische Staffelung der Tags garantiert. Die sog. ‚OHCO-Debatte‘ setzte vor diesem Hintergrund zunächst bei einem technischen Detail und damit *bottom-up* bei einem Praxisproblem an: Wie lässt sich eine sog. *Overlapping Hierarchy* codieren? Ein solcher *Overlap* liegt zum Beispiel bei einem lyrischen Enjambement vor, in dem eine Phrase die Versgrenze überspringt. *Inline* Markup stößt hier an eine Grenze, denn der schließende Tag der hierarchisch untergeordneten syntaktischen Einheit ‚Phrase‘ liegt jenseits des schließenden Tags der übergeordneten Einheit ‚Vers‘.²⁷

Stand-off Markup ist eine von mehreren Möglichkeiten, dieses Detailproblem der Markuppraxis technisch zu lösen.²⁸ Aber das ist eben nur die technische,

²⁷ Zur Problematik der *Overlapping Hierarchies* insbesondere im Kontext von TEI siehe Elena Pierazzo, „Textual Scholarship and Text Encoding“, in: Susan Schreibman/Ray Siemens/John Unsworth (Hg.), *A New Companion to Digital Humanities*, Chichester/West Sussex u. a. 2016, 307–321, hier: 316–319, sowie Andreas Witt, „Multiple hierarchies. New aspects of an old solution“, in: *Proceedings of Extreme Markup Languages (2004)*, <http://conferences.idealliance.org/extreme/html/2004/Witt01/EML2004Witt01.html> (letzter Aufruf 8.1.2017).

²⁸ Eine andere Möglichkeit ist z. B. die Codierung in sog. *Goddags (General Ordered-Descendant Directed Acyclic Graphs)*. Siehe hierzu Yves Marcoux/Michael Sperberg-McQueen/Claus Huitfeldt, „Modeling overlapping structures“, in: *Balisage. The Markup Conference (2013)*, <https://www.balisage.net/Proceedings/vol10/html/Marcoux01/BalisageVol10-Marcoux01.html> (letzter Aufruf 26.4.2018).

problemorientierte Perspektive. Die *konzeptionelle* Relevanz von *Stand-off Markup* für die Praxis digitaler Textwissenschaften ist grundsätzlicherer Natur: Denn überlappende Hierarchien und diskontinuierlich über das Ganze ausgeprägte Phänomene sind nicht nur ein formales Binnenmerkmal des literaturwissenschaftlichen Gegenstandes ‚Text‘ – sie kennzeichnen vielmehr auch den pragmatischen Kontext und die methodischen Praktiken, innerhalb dessen Texte als Objekte der Literaturwissenschaften seit dem *Digital Turn* thematisch werden. Die traditionelle philologische Praxis ließ sich noch denken unter Bezugnahme auf eine hierarchisch-substanzorientierte Vorstellung von Texten (primären Gegenständen und sekundären Informationsquellen), Bibliotheken (Institutionen der Quellen- und Wissensverwaltung) und eine klare Distinktion der Rollen von Autoren (intentional agierenden Produzenten) und Lesern (Adressaten und Interpreten der verschriftlichten Rede). Für eine Philologie, die das Erkenntnispotenzial digitaler Medien, Modelle und Verfahren konzeptionell für sich nutzen will, greift diese aus der Ära des Printmediums stammende Modellierung von Texten und Gebrauchskontexten, die sich u. a. in dem deskriptiv-hierarchischen Markupmodell niederschlägt, jedoch zu kurz.

Die digitale Praxis literaturwissenschaftlicher Textthermeneutik sollte diesen Emulationsmodus deshalb konsequent hinter sich lassen und nicht länger von einer essenzialistischen Vorstellung von Text und traditionellen Rechte-Rollen-Schemata ausgehen, die in der allgemeinen literaturtheoretischen Debatte ohnehin spätestens mit dem Poststrukturalismus infrage gestellt worden sind. Eine Praxis der „literary annotation in the digital age“;²⁹ die sich nicht auf die leichter digital zu operationalisierenden Fragestellungen formal orientierter Arbeitsfelder wie etwa Narratologie oder Editorik beschränkt, wird deshalb vermutlich ein insgesamt eher ereignisorientiertes denn ein substanz- und rollenorientiertes Textmodell entwickeln müssen. Für die digitale Annotation bedeutet dies, den Objekttext wie seine Annotationen mal als Knotenpunkte, mal als Kanten in einem n-dimensionalen, dynamisch rekonfigurierbaren Netzwerk zu denken. Einer der Akteure in diesem Netzwerk ist dabei der Textinterpret: Je nach Erkenntnisinteresse und methodischem Ansatz definiert, systematisiert und exploriert er als hermeneutischer Agent im Zuge der digitalen Annotation die ‚Kanten‘ zu und von den Texten und ihren Annotaten sowie deren mögliche Clusterungen. Ein – wenn nicht gar der zentrale – Vorteil digitaler Repräsentationen von textuellen Objekten und Bezügen gegenüber analogen Vorläufern und Praktiken ist dabei die erheblich einfachere Aggregation, Rekonfiguration und Analyse der textuellen Objekte und Annotate, die jeweils in den Blick genommen werden können. Ein zweiter Vorteil ist die Möglichkeit zum ‚Re-Run‘ einer Forschungsfrage, die digitale Verfahren dank formaler Parametrisierung und computergestützter Abarbeitung repetitiver Routinen eröffnen.

²⁹ Bauer/Zirker (Anm. 21).

Tab. 1 Prototypische Ausprägungen von digitalen Annotationen als Praktiken der Textmodellierung

	konzeptionelles Modell	logisches Modell
diskursives digitales Annotieren	hermeneutische Annotation	explorativer Lesartenhinweis
deklaratives digitales Annotieren	taxonomische Annotation	formalisierter Lesartenhinweis

Für den engeren Bezugsbereich einer digitalen Texthermeneutik bedeutet dies, dass die einzelne Textinterpretation nicht mehr zwingend als geschlossenes Ganzes gedacht werden muss, sondern dass sie auch als eine unter n möglichen Konfigurationen von Interpretations- und Deklarationsentscheidungen konzeptualisiert werden kann. *Stand-off* Markup unterstützt dieses Verfahrensmodell, weil es konzeptionell nicht mehr ein dokumentenzentriertes Verfahren ist, sondern ein datenbankzentriertes, wie Schloen/Schloen hervorheben:

[...] stand-off markup deviates so much from the original markup metaphor that it no longer belongs within the document paradigm at all and is best implemented within the database paradigm. Stand-off markup involves the digital representation of multiple readings of a text by means of separate data objects, one for each reading, with a system of pointers that explicitly connect the various readings to the text's components. But this amounts to a database solution to the problem. The best way to implement this solution is to abandon the use of a single long character sequence to represent a scholarly text — the document approach — in order to take advantage of the atomized data models and querying languages characteristic of database systems.³⁰

Diese digitale Repräsentation von Annotationen als „multiple readings of a text by means of separate data objects“ setzt allerdings noch mehr voraus als nur den Wechsel vom dokumentenzentrierten zum datenbankzentrierten Modell: Sie bedeutet auch, die Praxis der digitalen Annotation insgesamt als eine des *Data Modeling* zu betrachten. In Anlehnung an Flanders/Jannidis³¹ ließe sich dann die Praxis der Datenmodellierung vorgeschlagene Unterscheidung *Conceptual Model* vs. *Logical Model* und *Curation-Driven Modelers* vs. *Research-Driven Modelers* in abgewandelter Form zu einer Matrix erweitern, mit der sich vier prototypische Ausprägungen von digitalen Annotationen als Praktiken der Textmodellierung unterscheiden lassen (s. Tab. 1).

³⁰David Schloen/Sandra Schloen, „Beyond Gutenberg. Transcending the Document Paradigm in Digital Humanities“, in: *DHQ* 8/4 (2014), <http://www.digitalhumanities.org/dhq/vol/8/4/000196/000196.html> (letzter Aufruf 3.5.2018).

³¹Julia Flanders/Fotis Jannidis, „Data Modeling“, in: Susan Schreibman/Ray Siemens/John Unsworth (Hg.), *A New Companion to Digital Humanities*, Chichester, West Sussex, UK, 2016, 229–237.

Diskursiv orientiertes, hermeneutisch funktionales Annotieren als Exploration und Verschriftlichung individueller ‚Lesarten‘ eines Textes wird einen konzeptionellen Nutzen aus der Digitalisierung des Verfahrens in vielen Fällen zunächst nur mittelbar ziehen können: Je kontextabhängiger die Interpretation einer Textstelle ist, desto schwieriger ist die Formalisierung des Prozesses. Auch für die hermeneutisch ambitionierte Annotationspraxis mag das automatisierte *Pre-Processing* und ggf. auch das automatische Taggen eines Textes zwar durchaus heuristischen Wert haben – die eigentliche Aufgabe der Textinterpretation ist damit jedoch in aller Regel noch nicht erfüllt. Umso wichtiger ist deshalb die Bereitstellung von digitalen Arbeitsumgebungen, mit denen Philologen im ersten Zugriff unstrukturiert oder strukturiert, in jedem Fall aber kollaborativ und damit diskursiv orientiert annotieren, um dann im zweiten Schritt Objekttext wie Annotate mithilfe des Rechners analysieren zu können.

Wichtig ist dabei, den Übergang von der (leichter) automatisierbaren, hermeneutisch jedoch eher niedrigschwelligen formalen Textannotation (z. B. *POS Tagging*) bis hinauf zur nicht mehr algorithmenbasiert reproduzierbaren, dafür jedoch für eine globale Textdeutung entscheidenden semantischen Textannotation als stetig zu denken. Zwischen induktiven und deduktiven, deklarativen und interpretativen Annotationsmodi und automatischen wie intelligenten Verfahren kategorisch zu unterscheiden, hat wenig Sinn, wenn die *DH* mit der traditionellen Texthermeneutik ins Gespräch kommen wollen. Diese grundsätzlichen Kennzeichnungen können jedoch hilfreich sein, um die jeweilige Praxis einer digitalen Textannotation in einem dreidimensionalen Kontinuum zu verorten, dessen Achsen die der Methodik, der Funktion und des Verfahrens sind (s. Abb. 1).

Von einem derart abstrakt formulierten Konzept digitaler Textannotation ausgehend, könnten dann digitale Annotationswerkzeuge und kollaborative Plattformen entwickelt werden, die die konkreten *Use Cases* traditionell arbeitender, hermeneutisch orientierter Forscher aufgreifen, statt wie bisher primär technik- und standardgetrieben voranzuschreiten. Ein strategisch interessanter Ansatz könnte dabei die Formulierung hermeneutischer *Grand Challenges* für die digitale Textannotation und -exploration sein, bei der *DH*-Vertreter und klassisch-hermeneutisch arbeitende Literaturwissenschaftler gemeinsam bewusst Fragestellungen jenseits der linguistisch-textempirischen Phänomene in den Blick nehmen. Von Seiten der Texthermeneutik dürften die formulierten Herausforderungen dabei nur selten darin bestehen, rein technische Routinen und Prozesse der Textannotation als solche zu automatisieren und zu optimieren: Nicht der technologische und informationstheoretische, sondern der epistemologische und konzeptionelle Aspekt dürfte vielmehr bei der Suche nach den wirklichen *Killer Applications*³² im Vordergrund stehen.

³²Patrick Juola, „Killer Applications in Digital Humanities“, in: *Literary and Linguistic Computing* 23/1 (2008), 73–83.

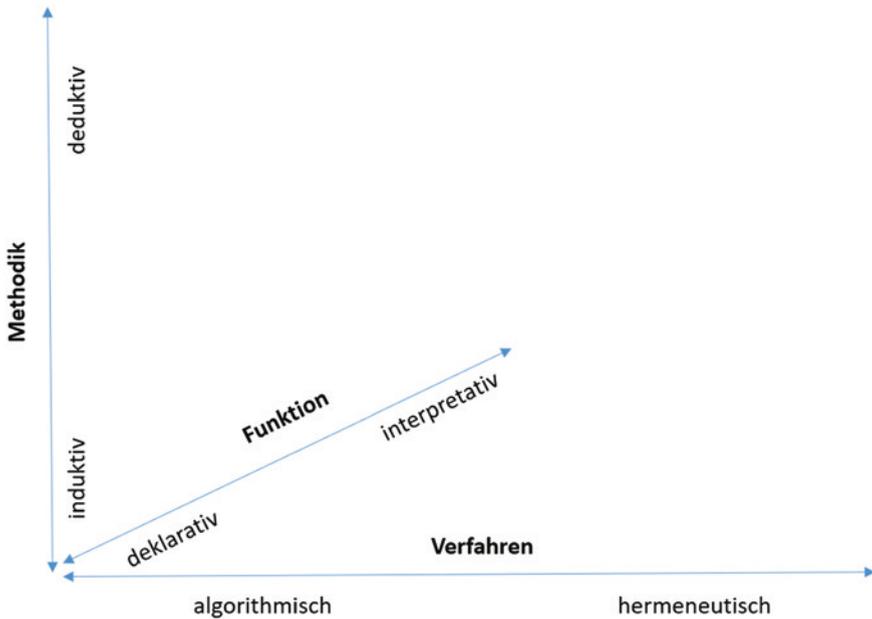


Abb. 1 Dimensionen der digitalen Textannotation

5 Desiderata

Aus den vorangehenden Überlegungen lassen sich drei zentrale Desiderata für die Entwicklung und Dissemination von digitalen Annotationsverfahren und -plattformen ableiten. Anzustreben sind:

- eine umfassende *born-digital*-Modellierung von Textannotation als literaturwissenschaftlichem Verfahren;
- die konzeptionelle und technische Integration der Arbeitsschritte Textauszeichnung – Textannotation – Textanalyse
- eine ‚Naturalisierung‘ digitaler Verfahren in den Philologien.

5.1 *Born-digital*-Modellierung literaturwissenschaftlicher Textannotation

Die digitale Modellierung der traditionellen Praxis non-deterministischer, explorativer Textannotation sollte, wie u. a. Bradley mit Blick auf einen „Richer

Sense of Digital Annotation“³³ hervorgehoben hat, ein erheblich breiteres Spektrum erfassen, als dies mit den gängigen Methoden eines deklarativen, taxonomiegetriebenen Markups möglich ist. Insbesondere die unstrukturierte, intuitive Annotationspraxis muss stärker berücksichtigt werden – was in der traditionellen Praxis bislang auf Papier in Form von Unterstreichungen und Marginalien-Kommentaren zunächst manuell erfasst wurde, um dann allenfalls in einem zweiten Durchgang selektiv in eine digitale Variante kopiert bzw. umgeschrieben zu werden. So sollte also durchgehend eine *born-digital*-Textannotation entstehen können, die dem Nutzer schon in Echtzeit für methodisch anschließende computergestützte Prozesse (Indexierung, *Search and Retrieval*, Schematisierung etc.) zur Verfügung steht.

Dies kann sowohl synchron als auch noch während des Annotationsvorgangs selbst geschehen – etwa durch interaktive Auswertung und Visualisierung der Verteilung bereits annotierter Phänomene im Text- oder Korpus-Ganzen. Unterstützt werden sollte dieser Ansatz *idealiter* jedoch zugleich durch eine asynchrone maschinelle Auswertung des Gesamtkomplexes von Textdaten, Annotationsdaten und im System protokollierten Nutzeraktionen. Ziel einer solchen digitalen Metareflexion unter Verwendung von *Machine-Learning*-Verfahren ist dabei die Modellierung und schrittweise Automatisierung von Annotationsproblemen steigender Komplexität. Dabei geht es nicht allein darum, den Nutzer von repetitiven Aufgaben zu befreien. Durch Einbeziehung der informatischen Perspektive auf Problemstellung und Lösungswege sollten digitale Modellierungsangebote vielmehr auch epistemisch funktional gemacht werden: Denn sie emulieren nicht nur bekannte Verfahren, sondern eröffnen auch neue Perspektiven auf den hermeneutischen Gegenstand ‚Text‘, die bislang aufgrund zu hoher Komplexität für den menschlichen Leser und Kommentator nicht greifbar waren. Wesentlich ist dabei insbesondere die Unterstützung kollaborativer Annotation, die eine methodologische Erweiterung philologischer Annotationspraxis bedeutet.

5.2 Konzeptionelle und technische Integration von Textauszeichnung – Textannotation – Textanalyse

In der traditionellen philologischen Praxis ergänzen sich die Textauszeichnung (Beispiel: eine Analyse des Metrums), die Textannotation (Beispiel: die Explikation einer Metapher), die Textanalyse (die systematische Anwendung der beiden ersten Schritte auf den Gesamttext) und der Textkommentar (Beispiel: die

³³ John Bradley, „Towards a Richer Sense of Digital Annotation. Moving Beyond a ‚Media‘ Orientation of the Annotation of Digital Objects“, in: *DHQ* 6/2 (2012), <http://www.digitalhumanities.org/dhq/vol/6/2/000121/000121.html> (letzter Aufruf 26.4.2018).

kritische Würdigung der Form-Inhalt-Relation zwischen Metrum und Metaphorik vor dem Hintergrund eines Werkganzen). Diese methodische Integration muss auch eine digitale Forschungsumgebung für die Philologien zumindest in Hinblick auf die ersten drei Arbeitsschritte leisten⁶³ – und sie sollte dabei zudem berücksichtigen, dass in der literaturwissenschaftlichen Forschungspraxis gerade diese drei Verfahren eigentlich immer schon diskursiv praktiziert wurden. Denn die jeweils erzielten Ergebnisse fließen seit jeher in den fortlaufenden kritischen Fachdiskurs der Literaturwissenschaften ein. Diskursivität als solche ist also kein Novum – der eigentliche konzeptionelle wie methodische Mehrwert liegt vielmehr in der gesteigerten Komplexität, Synchronizität und Nachnutzbarkeit der ‚Arbeit am Text‘, der aus der digitalen Integration der Verfahren resultiert und im kollaborativen Modus nochmals gesteigert werden kann.

Eine virtuelle Textannotationsumgebung sollte deshalb:

- den Dissens über Bedeutungen als produktive Bedingung und *raison d'être* der Textwissenschaften begreifen;
- den Geltungsbereich der normativen Leitidee des *Inter-Annotator Agreements* bewusst relativieren, indem sie im Umgang mit semantisch anspruchsvolleren Textphänomenen auch das explorative Potenzial eines reflektierten *Disagreements* nutzbar macht;
- statt des linearen Voranschreitens von der taxonomischen Textbeschreibung zur finalen Textdeutung den zyklischen Charakter und die Wechselwirkung von Textannotation und Textanalyse als Prozessschema zugrunde legen.

Für den philologischen Nutzer könnten damit auch in der digitalen Umgebung Textauszeichnung, Textannotation und Textanalyse als methodisches Kontinuum handhabbar bleiben. Die angestrebte *User Experience* emuliert damit die etablierten Praktiken des philologischen Arbeitens, statt Workflows und Segmentierungen von Prozessen zu erzwingen, die sich primär an den technischen Anforderungen orientieren.

5.3 ‚Naturalisierung‘ digitaler Verfahren in den Philologien

Diese Anschlussfähigkeit an etablierte ‚analoge‘ Routinen und Gebrauchsmetaphern ist dabei zugleich ein strategisches Desiderat: Denn mit ihr steht und fällt die Akzeptanz einer Plattform. Ihr wäre mit einem mehrstufigen Usability-Konzept Rechnung getragen, das den konzeptionellen Mehrwert eines *Mixed-Methods*-Ansatzes für die Literaturwissenschaften – also: einer Kombination

⁶³ Das Ausformulieren komplexer Textkommentare bleibt dagegen eine intellektuelle Leistung, die von einer Forschungsumgebung nicht konzeptionell, sondern nur instrumentell unterstützt werden kann – etwa durch Bereitstellung von Recherchetools, Datenbankanbindungen, Kollaborationstools, Workflow-Management etc.

manueller und automatischer Routinen – vermitteln kann. Drei Kriterien sind dabei erfolgskritisch:

- Vom Nutzer dürfen keinerlei Vorkenntnisse im Bereich der digitalen Modellierung und Analyse von Texten und Korpora verlangt und damit vom System auch keine implizite Rekonzeptualisierung des Forschungsgegenstandes ‚Text‘ als Datenstruktur erzwungen werden.
- Im Einstiegs-Modus sollte eine Plattform in Bezug auf Oberflächengestaltung, Terminologie und Handhabung bewusst traditionelle Praktiken, Konzepte, Verfahren und Gebrauchsmetaphern der Philologien emulieren, um dann
- im Experten-Modus schrittweise mittels interaktiver Komponenten zwischen dem traditionellen und dem digital modellierten Workflow zu vermitteln. Wesentlich ist hier insbesondere der Schritt zur kollaborativen Arbeit am Text bzw. Textkorpus, die sowohl asynchron wie synchron, also auch in *Real Time* möglich sein und zudem Gelegenheiten zur Einbindung automatisierter Komponenten bieten muss.

6 CATMA als *Proof of Concept*

Der vorangehend skizzierte Anforderungskatalog mag auf den ersten Blick ambitioniert wirken. Tatsächlich gibt es jedoch bereits jetzt eine Reihe von Tools und Plattformen, die ihm in Konzeption und Funktionalität teilweise entsprechen. Als wohl wichtigstes frühes Beispiel zu nennen ist hier die bereits erwähnte, von Bradley entwickelte experimentelle Annotations- und Forschungsumgebung *Pliny*.⁶⁴ Plinys ausschlaggebende konzeptionelle Restriktion ist allerdings, dass das Tool als Desktopapplikation nur den Arbeitsablauf des Einzelnutzers unterstützt, aber keine Kollaborationsfunktionalität beinhaltet.

Die vorangehend skizzierten konzeptionellen Überlegungen und technischen Erfordernisse sind dagegen umfassender umgesetzt worden in dem seit 2008 an der Universität Hamburg entwickelten Textannotationstool CATMA (Computer Assisted Textual Markup and Analysis).⁶⁵ CATMA war zunächst als Re-Implementierung der bereits in der Mitte der 1980er Jahre von Bradley programmierten Programmsuite TACT (Textual Analysis Computing Tools) geplant,⁶⁶ eines der erfolgreichsten und verbreitetsten frühen Textanalysetools. Im Zuge des Architekturdesigns und nach Rücksprache mit Bradley wurde daraus jedoch ein komplexeres Entwicklungsvorhaben. Neben der Annotationsfunktion wurde das Programm in einer ersten Entwicklungsphase bis 2010 um mehrere

⁶⁴ Bradley (Anm. 26); siehe <http://pliny.cch.kcl.ac.uk> (letzter Aufruf 8.1.2016).

⁶⁵ Siehe <http://www.catma.de> (letzter Aufruf 8.1.2016).

⁶⁶ Ian Lancashire, *Using TACT with electronic texts. A guide to text-analysis computing tools, version 2.1 for MS-DOS and PC DOS*, New York 1996.

Funktionskomponenten erweitert (insbes. Textanalysefunktionen und einen *Query Builder*⁶⁷). Ab 2010 wurde CATMA dann als Webapplikation in JAVA implementiert, um webbasierte Kollaborationen in Echtzeit bei der Text-/Textkorpusannotation und -analyse zu unterstützen. Um die notwendige Flexibilität bei gleichzeitiger Standardkonformität sicherzustellen, basieren CATMA-Annotationen auf dem sog. *Feature-Structure-Modell* von TEI.⁶⁸ CATMA-Annotationen sind damit TEI/XML-kompatibel und können in alle gängigen Formate exportiert werden. Umgekehrt erlaubt CATMA den Import aller gängigen Annotationsformate, insbesondere TEI-XML. Objekttexte und Korpora können in diversen Formaten (darunter txt, rtf, doc, docx, pdf) in das CATMA-System geladen oder über eine Schnittstelle direkt aus Repositorien eingelesen werden. Intern verarbeitet CATMA alle Texte in UTF-8 bzw. UTF-16 und unterstützt somit eine Vielzahl UNICODE-basierter Sprachen inkl. einiger rechts-links geschriebener Sprachen wie z. B. Hebräisch und Arabisch. In CATMA 5.0 wurden außerdem prototypisch erstmals *Machine-Learning*-basierte Annotationsroutinen implementiert, mit denen z. B. das in der Narratologie als *Analepse* definierte Phänomen des non-linearen, rückschauenden Erzählens einer Vorgeschichte vom Zeitpunkt der Erzählgegenwart aus in deutschsprachigen Texten automatisch getaggt werden kann. CATMA wird aktuell international in mehr als 120 Forschungsvorhaben eingesetzt und ist bislang die einzige webbasierte Plattform, die

- gezielt non-deterministische, kollaborative Textannotation unterstützt;
- das komplette Funktionsspektrum ‚Textauszeichnung – Textannotation – Textanalyse‘ abbildet;
- Schnittstellen zur Einbindung externer Routinen und Services für automatisierte Textauszeichnung und Korpora besitzt;
- eine systeminterne *Machine-Learning*-Komponente integriert.

CATMA kann damit als *Proof of Concept* gelten, das die konkrete technische Umsetzbarkeit der methodischen Anforderungen demonstriert, die ein reichhaltigeres, stärker auf die texthermeneutischen Erfordernisse eingehendes *DH*-Annotationskonzept nach sich zieht, für das hier plädiert wurde.⁶⁹ CATMA ist zwar bislang die einzige Plattform, die dem entsprechenden technischen Anforderungskatalog gerecht wird.⁷⁰ Aber das ist durchaus keine Konsequenz

⁶⁷ D. i. eine natürlichsprachliche Oberfläche zur Generierung komplexer Anfragen sowie einer Datenvisualisierung.

⁶⁸ Zum *Feature-Structure*-Konzept siehe <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/FS.html> (letzter Aufruf 8.1.2016).

⁶⁹ CATMA wird in dem aktuellen DFG-LIS-Projekt *forTEXT*, <http://www.fortext.net> (letzter Aufruf 8.1.2018) schrittweise um u. a. folgende Komponenten erweitert werden: Skalierbarkeit & Performanz; Rollen & Rechte-Schemata sowie Usability-Anforderungen; Aufbau eines ML-basierten Recommender-Funktionskomplexes; Integration von Visualisierungsfunktionen; UIMA-basierte Einbindung weiterer externer Services; Anbindung an Textrepositorien wie z. B. *TextGrid* oder *CLARIN*-Repositorien.

⁷⁰ Siehe hierzu die Bestandsaufnahme im Anhang, die im Rahmen der Antragstellung zum Projekt *forTEXT* von Marco Petris erarbeitet wurde.

technischer Restriktionen, die auf das Feld der *DH* einwirken. Es ist vielmehr eine Folge der konzeptionellen Selbstbeschränkung der *DH* auf eine Praxis der digitalen Anreicherung von Texten mit ‚Lesartenhinweisen‘, die Markup bleiben. Das aber ist nur die Vorstufe zu dem, was in der Textthermeneutik eigentlich gefordert ist: digitale Annotation.

Anhang

Im Folgenden werden die Ergebnisse der Bestandsaufnahme und Evaluation bestehender Plattformen und Werkzeuge für die Textannotation und -analyse dargestellt, die 2016 für den Antrag zum DFG-geförderten LIS-Projekt *forTEXT* erarbeitet wurden (Autor: Marco Petris). Dabei wurden ausschließlich Plattformen berücksichtigt, die aktiv entwickelt und betreut werden.⁷¹ Die Plattformen wurden bezüglich ihrer Funktionalität, Zielgruppe und Nachhaltigkeit betrachtet. Hinsichtlich der Zielgruppenspezifität wurde insbesondere die erforderliche technische Versiertheit potenzieller Nutzer bewertet. Abhängig von technischem Vorverständnis und Interesse lassen sich dabei drei Nutzertypen unterscheiden:

1. Nutzer ohne Bereitschaft, sich mit technischem Hintergrundwissen auseinanderzusetzen;
2. Nutzer mit der Bereitschaft, sich mit technischem Hintergrundwissen auseinanderzusetzen;
3. Nutzer mit der Bereitschaft, auf Programmcode-Ebene zu arbeiten.

Kriterien für Nachhaltigkeit der untersuchten Plattformen sind i) das Setzen auf offene Standards und ii) die Veröffentlichung des Programmcodes unter einer die Nachnutzbarkeit ermöglichenden Lizenz. Die Ergebnisse der Bestandsaufnahme werden hier tabellarisch dargestellt (s. Tab. A.1). Die Gesamtauswertung ergab, dass 2016 von den bislang entwickelten Werkzeugen keines die in *forTEXT* anvisierten Funktionalitäten im Sinne eines *Hermeneutic-Markup*-Ansatzes abdeckte und dabei

- mit wenig technischem Hintergrundwissen bedienbar ist,
- Funktionen zur Annotation und zur Auswertung integriert,
- kollaboratives Arbeiten ermöglicht,
- die Möglichkeit zur Modellierung komplexerer Taxonomien bereitstellt und
- den Quellcode zur Verfügung stellt.

⁷¹ Projekte wie *Project Bamboo* oder die *Open Corpus Workbench* spielen deshalb als mögliche Komponenten von *forTEXT* keine Rolle, siehe auch: Mark A. Finlayson/Patrick H. Winston, *The Rapid Story Annotation Workbench* (2007), <http://publications.csail.mit.edu/abstracts/abstracts07/markaf/markaf.html> (letzter Aufruf 8.6.2013).

Tab. A.1 Bestandsaufnahme und Evaluation bestehender Plattformen und Werkzeuge für die Textannotation und -analyse

Plattform	Hauptfunktionen	Source Code veröffentlicht? (Lizenz)	Nutzertyp
Apache UIMA ³⁴	<ul style="list-style-type: none"> • Java/C++-basierte Komponenten können über die Common Analysis Structure (CAS) zu Workflows zusammengefasst und über eine Workflow Engine ausgeführt werden • vertikale und horizontale Skalierung • zahlreiche Komponenten für die Arbeit mit Texten • erlaubt die Definition von Typsystemen für die Erzeugung/Verarbeitung von Analysedaten 	ja (ASL) ³⁵	3
DKPro Core ³⁶ & DKPro Lab ³⁷	<p>DKPro Core:</p> <ul style="list-style-type: none"> • Sammlung von Komponenten für NLP auf Basis von Apache UIMA <p>DKPro Lab:</p> <ul style="list-style-type: none"> • Framework zum Erstellen von deklarativ in Java notierten Workflows • Workflows mit unterschiedlichen Parameterbelegungen ausführbar (Fokus auf Effizienz: Ein einzelner Schritt wird nur einmal ausgeführt, sofern sich die Parameter für diesen Schritt nicht ändern.) • speichert Ein- und Ausgabewerte des Workflows, deshalb stehen die Informationen (<i>Data Provenance</i>) auch nach Ausführung noch in vollem Umfang zur Verfügung 	ja (ASL, Teile von DKPro Core in GPL ³⁸)	3
GATE ³⁹	<ul style="list-style-type: none"> • Java-basierte Komponenten können zu Workflows zusammengefasst werden • Komponenten erzeugen bzw. verarbeiten Annotationen mit frei definierbaren Typsystemen 	Ja	3 (GATE Developer: 2)

(Fortsetzung)

³⁴UIMA (Unstructured Information Management Architecture), <https://uima.apache.org/> (letzter Aufruf 8.1.2016). In vorangegangenen Projekten konnten bereits Komponenten zur Integration von CATMA und UIMA erarbeitet werden. Hierzu zählen die Anbindung an die CATMA API zum Import und Export von Texten und Annotationen, der Collection Reader zur Verarbeitung von CATMA Source Documents und die Abbildung von UIMA-Typsystemen auf CATMA-Typsysteme.

³⁵*Apache Software License* (vgl. <http://www.apache.org/licenses/> [letzter Aufruf 8.1.2016]).

³⁶Vgl. Richard E. de Castilho/Iryna Gurevych, „A Broad-Coverage Collection of Portable NLP Components for Building Shareable Analysis Pipelines“, in: *Proceedings of the Workshop on Open Infrastructures and Analysis Frameworks for HLT (2014)*, 1–11.

³⁷Vgl. Richard E. de Castilho/Iryna Gurevych, „A Lightweight Framework for Reproducible Parameter Sweeping in Information Retrieval“, in: *DESIRE (2011)*, 7–10.

³⁸*GNU Public License* (vgl. <http://www.gnu.org/licenses/gpl-3.0.de.html> [letzter Aufruf 8.1.2016]).

³⁹*General Architecture for Text Engineering* (vgl. Hamish Cunningham/Kalina Bontcheva/Valentin Tablan u. a. [Hg.], *Text Processing with GATE. [Version 6]*, Sheffield 2011).

Tab. A.1 (Fortsetzung)

Plattform	Hauptfunktionen	Source Code veröffentlicht? (Lizenz)	Nutzertyp
	<ul style="list-style-type: none"> • umfangreiche Sammlung an Komponenten • GATE-Komponenten können in UIMA-Workflows und UIMA-Komponenten in GATE-Workflows integriert werden • GATE Developer: Desktopwerkzeug zum Erzeugen und Ausführen von Workflows und zum manuellen Annotieren von Texten • GATE Teamware: Webapplikation für Management von Projekten und Teams • GATE Server: ermöglicht kollaboratives Arbeiten mit dem GATE Developer • Lucene⁴²-basierte Suche und JAPE⁴³ • Möglichkeit, über bestehende Annotationen durch reguläre Ausdrücke neue Annotationen zu vergeben 	(LGPL ⁴⁰ , GATE Teamware: AGPL ⁴¹)	
KNIME ⁴⁴ Taverna ⁴⁵ Kepler ⁴⁷	<ul style="list-style-type: none"> • Plattform zum Erstellen und Ausführen von naturwissenschaftlichen Workflows • <i>Rich-Client</i>-Anwendungen mit graphischem Designer zum Erstellen von Workflows aus Komponenten • Komponenten für naturwissenschaftliche Workflows (kleinteiliger, deshalb flexibler nutzbar) 	ja (GPL, ASL, BSD ⁴⁶)	2

(Fortsetzung)

⁴⁰ Lesser GNU Public License (vgl. <http://www.gnu.de/documents/lgpl-3.0.de.html> [letzter Aufruf 8.1.2016]).

⁴¹ Affero GNU Public License (vgl. <http://www.gnu.org/licenses/agpl-3.0.de.html> [letzter Aufruf 8.1.2016]).

⁴² Vgl. <http://lucene.apache.org/> (letzter Aufruf 8.1.2016).

⁴³ Vgl. Hamish Cunningham/Diana Maynard/Valentin Tablan, *JAPE. A Java Annotation Patterns Engine*, o. O. 2000, https://www.researchgate.net/publication/2495768_JAPE_a_Java_Annotation_Patterns_Engine/citation/download (letzter Aufruf 8.1.2016).

⁴⁴ Vgl. Michael R. Berthold/Nicolas Cebron/Fabian Dill u. a., „KNIME. The Konstanz Information Miner“, in: Christine Preisach/Hans Burkhardt/Lars Schmidt-Thieme u. a. (Hg.), *Data Analysis, Machine Learning and Applications*, Berlin/Heidelberg 2008, 319–26, http://link.springer.com/10.1007/978-3-540-78246-9_38, (letzter Aufruf 3.5.2018).

⁴⁵ Vgl. Katherine Wolstencroft/Robert Haines/Donal Fellows u. a., „The Taverna Workflow Suite. Designing and Executing Workflows of Web Services on the Desktop, Web or in the Cloud“, in: *Nucleic Acids Research* 41 (2013), W557–W561.

⁴⁶ Vgl. <https://opensource.org/licenses/BSD-2-Clause> (letzter Aufruf 8.1.2016).

⁴⁷ Vgl. Ilkay Altintas/Chad Berkley/Efrat Jaeger u. a., „Kepler. An extensible system for design and execution of scientific workflows“, in: *Proceedings of the 16th International Conference on Scientific and Statistic Database Management* (2004), 21–23.

Tab. A.1 (Fortsetzung)

Plattform	Hauptfunktionen	Source Code veröffentlicht? (Lizenz)	Nutzertyp
	<ul style="list-style-type: none"> externe Software einbindbar (z. B. R⁴⁸ oder Weka⁴⁹) Workflows können exportiert und über Seiten wie myExperiment⁵⁰ geteilt werden 		
MAXQDA ⁵¹ ATLAS.ti ⁵²	<ul style="list-style-type: none"> <i>Rich-Client</i>-Anwendungen zur qualitativen Datenanalyse Schwerpunkt auf sozialwissenschaftlichen Problemstellungen Modellierung von Taxonomien wegen eindimensionaler Annotationstypen (<i>Codes</i>) 	nein (kommerzielle Lizenzen)	1
TextGrid ⁵³	<ul style="list-style-type: none"> dokumentierte Schnittstellen zur Erweiterung bzw. zur Integration mit anderen Systemen unterstützt kollaboratives Arbeiten in Teams mithilfe einer projektbezogenen Benutzerverwaltung auf Basis von openRBAC⁵⁴ <p>TextGridRepository:</p> <ul style="list-style-type: none"> verteilte Speicherinfrastruktur für Datenhaltung TextGrid Middleware Utilities (kapselt Zugriff auf den Speicher für die Services-Schicht) <p>TextGridLaboratory:</p> <ul style="list-style-type: none"> Services bzw. Komponenten für die Analyse und Verarbeitung von Daten⁵⁵ 	ja (LPGL)	1
	<ul style="list-style-type: none"> graphische Schnittstelle zum Endnutzer basierend auf Eclipse RCP⁵⁶, erweiterbar durch Plugins 		

(Fortsetzung)

⁴⁸Vgl. <https://www.r-project.org/> (letzter Aufruf 8.1.2016).

⁴⁹Vgl. <http://www.cs.waikato.ac.nz/ml/weka/> (letzter Aufruf 8.1.2016).

⁵⁰Vgl. www.myexperiment.org/ (letzter Aufruf 8.1.2016).

⁵¹Vgl. <http://www.maxqda.de/> (letzter Aufruf 8.1.2016).

⁵²Vgl. <http://atlasti.com> (letzter Aufruf 8.1.2016).

⁵³Vgl. Heike Neuroth/Andrea Rapp/Sibylle Söring u. a. (Hg.), *TextGrid. Von der Community – für die Community*, Göttingen 2015.

⁵⁴Vgl. http://www.openrbac.de/en_startup.xml (letzter Aufruf 8.1.2016).

⁵⁵Zurzeit sind für *TextGrid* vergleichsweise wenig Komponenten direkt verfügbar. Anwendungsfälle aus der Editionsphilologie werden durch die bestehenden Komponenten jedoch bereits unterstützt. Die Unterstützung zum Modellieren von Workflows ist noch in einem experimentellen Stadium und in der aktuellen Version (Nightly Build 3.1.0.201507151344) nicht enthalten bzw. wird nicht über den *TextGrid Marketplace* angeboten.

⁵⁶Vgl. <http://www.eclipse.org> (letzter Aufruf 8.1.2016).

Tab. A.1 (Fortsetzung)

Plattform	Hauptfunktionen	Source Code veröffentlicht? (Lizenz)	Nutzertyp
WebAnno ⁵⁷ und ANNIS ⁵⁸	<p>WebAnno:</p> <ul style="list-style-type: none"> • kollaboratives, webbasiertes Annotationswerkzeug • eigene Annotationstypen definierbar (Modellierungsmöglichkeit ist aufgrund der eindimensionalen Typen begrenzt; kleinste annotierbare Zeichenkette ist ein Wort) <p>ANNIS:</p> <ul style="list-style-type: none"> • Analysewerkzeug zur Visualisierung und Auswertung in WebAnno erstellter Annotationen (allerdings kein integrierter Ansatz) 	ja (ASL)	1
WebLicht ⁵⁹	<ul style="list-style-type: none"> • Komponenten auf Basis von REST Web Services können zu Workflows zusammengefasst und ausgeführt werden • eigenes Datenformat: TCF⁶⁰ • umfangreiche Bibliothek mit Komponenten • über Web UI benutzbar und als Web-Service in Plattformen wie GATE und UIMA einbindbar⁶¹ • mit Suchkomponente können Anfragen in CQL⁶² gestellt werden 	nein (steht Forschern und Studierenden zur Verfügung)	2

⁵⁷ Vgl. Seid Muhie Yimam/Iryna Gurevych/Richard E. de Castilho u. a., „A Flexible, Web-based and Visually Supported System for Distributed Annotations“, in: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (2013), 1–6.

⁵⁸ Vgl. Michael Götz/Stefanie Dipper, „ANNIS, Complex Multilevel Annotations in a Linguistic Database (System Demonstration)“, in: *Proceedings of the EACL Workshop on Multidimensional Markup in Natural Language Processing NLPXML* (2006), 61–64.

⁵⁹ Web-Based Linguistic Chaining Tool, vgl. Erhard Hinrichs/Marie Hinrichs/Thomas Zastrow, „WebLicht. Web-based LRT Services for German“, in: *Proceedings of the ACL 2010 System Demonstrations* (2010), 25–29.

⁶⁰ Ulrich Heid/Helmut Schmid/Kerstin Eckart u. a., „A Corpus Representation Format for Linguistic Web Services. The D-SPIN Text Corpus Format and its Relationship with ISO Standards“, in: *Proceedings of the International Conference on Language Resources and Evaluation, LREC* (2010), <http://www.lrec-conf.org/proceedings/lrec2010/summaries/503.html> [letzter Aufruf 3.5.2018].

⁶¹ Das Einbinden von eigenen Komponenten in *WebLicht* ist jedoch vergleichsweise aufwändig. Diese müssen via CMDI (<http://www.clarin.eu/cmdi>) beschrieben und über ein CLARIN-D-Repository, welches Web Services unterstützt (<http://www.clarin-d.de/de/aufbereiten/clarin-zentrum-finden> [letzter Aufruf der Links 8.1.2016]), veröffentlicht werden.

⁶² Oliver Christ, „A Modular and Flexible Architecture for an Integrated Corpus Query System“, in: *Proceedings of COMPLEX, Conference on Computational Lexicography and Text Research* (1994), 23–32.

Literatur

- Altintas, Ilkay/Berkley, Chad/Jaeger, Efrat u.a., „Kepler. An extensible system for design and execution of scientific workflows“, in: *Proceedings of the 16th International Conference on Scientific and Statistic Database Management* (2004), 21–23.
- Bauer, Matthias/Zirker, Angelika, „Whipping Boys Explained. Literary Annotation and Digital Humanities“, in: Kenneth M. Price/Ray Siemens (Hg.), *Literary Studies in the Digital Age. An Evolving Anthology*, New York 2013, <https://dlsanthology.mla.hcommons.org/> (letzter Aufruf 3.5.2018).
- Berthold, Michael R./Cebron, Nicolas/Dill, Fabian u.a., „KNIME. The Konstanz Information Miner“, in: Christine Preisach/Hans Burkhardt/Lars Schmidt-Thieme u.a. (Hg.), *Data Analysis, Machine Learning and Applications*, Berlin/Heidelberg 2008, 319–26, http://link.springer.com/10.1007/978-3-540-78246-9_38, (letzter Aufruf 3.5.2018).
- Boot, Peter, *Mesotext. Digitised Emblems, Modelled Annotations and Humanities Scholarship*, Utrecht 2009.
- Bradley, John, „Thinking about interpretation. Pliny and scholarship in the humanities“, in: *Literary and Linguistic Computing* 23 (2008), 263–279.
- Bradley, John, „Towards a Richer Sense of Digital Annotation. Moving Beyond a ‚Media‘ Orientation of the Annotation of Digital Objects“, in: *DHQ* 6/2 (2012), <http://www.digitalhumanities.org/dhq/vol/6/2/000121/000121.html> (letzter Aufruf 26.4.2018).
- Buzzetti, Dino, „Digital Representation and the Text Model“, in: *New Literary History* 33 (2002), 61–88.
- Buzzetti, Dino, „Text Representation and Textual Models“, in: *1999 ACH-ALLC Conference Program*, <http://www2.iath.virginia.edu/ach-allc.99/proceedings/buzzetti.html> (letzter Aufruf 7.5.2018)
- Castilho, Richard E. de/Gurevych, Iryna, „A Broad-Coverage Collection of Portable NLP Components for Building Shareable Analysis Pipelines“, in: *Proceedings of the Workshop on Open Infrastructures and Analysis Frameworks for HLT* (2014), 1–11.
- Castilho, Richard E. de/Gurevych, Iryna, „A Lightweight Framework for Reproducible Parameter Sweeping in Information Retrieval“, in: *DESIRE* (2011), 7–10.
- Caton, Paul, „Markup’s Current Imbalance“, in: *Markup Languages* 3/1 (2000), 1–13.
- Christ, Oliver, „A Modular and Flexible Architecture for an Integrated Corpus Query System“, in: *Proceedings of COMPLEX, Conference on Computational Lexicography and Text Research* (1994), 23–32.
- Coombs, James H./Renear, Allen H./DeRose, Steven J., „Markup systems and the future of scholarly text processing“, in: *Communications of the ACM* 30 (1987), 933–947.
- Cunningham, Hamish/Bontcheva, Kalina/Tablan, Valentin (Hg.), *Text Processing with GATE. (Version 6)*, Sheffield 2011.
- Cunningham, Hamish/Maynard, Diana/Tablan, Valentin, *JAPE. A Java Annotation Patterns Engine*, o.O. 2000, https://www.researchgate.net/publication/2495768_JAPE_a_Java_Annotation_Patterns_Engine/citation/download (letzter Aufruf 7.5.2018).
- Dombrowski, Quinn, „What Ever Happened to Project Bamboo?“, in: *Literary and Linguistic Computing* 29/3 (2014), 326–339.
- Finlayson, Mark A./Winston, Patrick H., *The Rapid Story Annotation Workbench* (2007), <http://publications.csail.mit.edu/abstracts/abstracts07/markaf/markaf.html> (letzter Aufruf 8.6.2013).
- Flanders, Julia/Jannidis, Fotis, „Data Modeling“, in: Susan Schreibman/Ray Siemens/John Unsworth (Hg.), *A New Companion to Digital Humanities*, Chichester, West Sussex, UK, 2016, 229–237.
- Götze, Michael/Dipper, Stefanie, „ANNIS, Complex Multilevel Annotations in a Linguistic Database (System Demonstration)“, in: *Proceedings of the EACL Workshop on Multi-dimensional Markup in Natural Language Processing NLPXML* (2006), 61–64.

- Heid, Ulrich/Schmid, Helmut/Eckart, Kerstin u.a., „A Corpus Representation Format for Linguistic Web Services. The D-SPIN Text Corpus Format and its Relationship with ISO Standards“, in: *Proceedings of the International Conference on Language Resources and Evaluation, LREC* (2010), <http://www.lrec-conf.org/proceedings/lrec2010/summaries/503.html> (letzter Aufruf 3.5.2018).
- Hinrichs, Erhard/Hinrichs, Marie/Zastrow, Thomas, „WebLicht. Web-based LRT Services for German“, in: *Proceedings of the ACL 2010 System Demonstrations* (2010), 25–29.
- Hockey, Susan/McGann, Jerome J./Renear, Allen, „What is text? A debate on the philosophical and epistemological nature of text in the light of humanities computing research“, in: *1999 ACH-ALLC Conference Program*, <http://www2.iath.virginia.edu/ach-allc.99/proceedings/hockey-renear2.html> (letzter Aufruf 7.5.2018).
- Jacke, Janina, „Is There a Context-Free Way of Understanding Texts? The Case of Structuralist Narratology“, in: *Journal of Literary Theory* 8/1 (2014), 118–139.
- Juola, Patrick, „Killer Applications in Digital Humanities“, in: *Literary and Linguistic Computing* 23/1 (2008), 73–83.
- Lancashire, Ian, *Using TACT with electronic texts. A guide to text-analysis computing tools, version 2.1 for MS-DOS and PC DOS*, New York 1996.
- Marcoux, Yves/Sperberg-McQueen, Michael/Huitfeldt, Claus, „Modeling overlapping structures“, in: *Balisage. The Markup Conference* (2013), <https://www.balisage.net/Proceedings/vol10/html/Marcoux01/BalisageVol10-Marcoux01.html> (letzter Aufruf 26.4.2018).
- Mayer-Schönberger, Viktor/Cukier, Kenneth, *Big data. A revolution that will transform how we live, work, and think*, Boston 2013.
- McGann, Jerome, „Marking texts in many dimensions.“, in: *A new republic of letters: Memory and scholarship in the age of digital reproduction*, Cambridge, MA, 2014, 90–112.
- Meister, Jan C. u. a., *DARIAH-DE Folgeantrag, Arbeitspaket 5.4*, internes Dokument 2013.
- Neuroth, Heike/Rapp, Andrea/Söring, Sibylle u. a. (Hg.), *TextGrid. Von der Community – für die Community*, Göttingen 2015.
- Pierazzo, Elena, „Textual Scholarship and Text Encoding“, in: Susan Schreibman/Ray Siemens/John Unsworth (Hg.), *A New Companion to Digital Humanities*, Chichester, West Sussex u.a. 2016, 307–321.
- Piez, Wendell, *Towards Hermeneutic Markup. An architectural outline* (2010), <http://dh2010.cch.kcl.ac.uk/academic-programme/abstracts/papers/html/ab-743.html> (letzter Aufruf 9.11.2017).
- Rapp, Andrea, „Manuelle und automatische Annotation“, in: Fotis Jannidis/Hubertus Kohle/Malte Rehbein (Hg.), *Digital Humanities. Eine Einführung*, Stuttgart 2017, 253–267.
- Renear, Allen, „Text Encoding“, in: Susan Schreibman/Ray Siemens/John Unsworth (Hg.), *Companion to Digital Humanities*, Oxford 2004.
- Renear, Allen, „The descriptive/procedural distinction is flawed“, in: *Markup Languages* 2/4 (2000), 411–420.
- Schloen, David/Schloen, Sandra, „Beyond Gutenberg. Transcending the Document Paradigm in Digital Humanities“, in: *DHQ* 8/4 (2014), <http://www.digitalhumanities.org/dhq/vol/8/4/000196/000196.html> (letzter Aufruf 3.5.2018).
- Staten, Henry, „Wittgenstein and the Intricate Evasions of ‚Is‘“, in: *New Literary History* 19/2 (1988), 281–300.
- Unsworth, John, *Scholarly Primitives. What methods do humanities researchers have in common, and how might our tools reflect this?*, London 2000.
- Nyhan, Julianne, „Text encoding and scholarly digital editions“, in: Dies./Claire Warwick/Melissa Terras, *Digital Humanities in Practice. Facet Publishing*, London 2012, 117–138.
- Witt, Andreas, „Multiple hierarchies. New aspects of an old solution“, in: *Proceedings of Extreme Markup Languages* (2004), <http://conferences.idealliance.org/extreme/html/2004/Witt01/EML2004Witt01.html> (letzter Aufruf 8.1.2017).
- Wittgenstein, Ludwig, *Philosophische Untersuchungen*, Frankfurt a. M. 1975 [1953].

- Wolstencroft, Katherine/Haines, Robert/Fellows, Donal u.a., „The Taverna Workflow Suite. Designing and Executing Workflows of Web Services on the Desktop, Web or in the Cloud“, in: *Nucleic Acids Research* 41 (2013), W557–W561.
- Yimam, Seid Muhie/Gurevyeh, Iryna/de Castilho, Richard E. u. a., „A Flexible, Web-based and Visually Supported System for Distributed Annotations“, in: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (2013), 1–6.
- Zundert, Joris J. van, „Screwmenetics and Hermenumerals. The Computationality of Hermeneutics“, in: Susan Schreibman/Ray Siemens/John Unsworth (Hg.), *A New Companion to Digital Humanities*, Chichester, West Sussex, UK, u. a. 2016.

Online-Ressourcen⁷²

- 3DH, Jan C. Meister, <http://jcmeister.de/projects/3dh/>.
- Apache Licenses, <http://www.apache.org/licenses/>
- Apache Lucene, <http://lucene.apache.org/>.
- Apache UIMA, <https://uima.apache.org/>.
- ATLAS.ti, <http://atlasti.com>.
- CATMA 5.0, <http://www.digitalhumanities.it/catma>.
- CATMA, <http://www.catma.de>.
- CLARIN-D, Zentren, <http://www.clarin-d.de/de/aufbereiten/clarin-zentrum-finden>.
- Database System Research, Michael Gertz, <http://dbs.ifi.uni-heidelberg.de/index.php?id=129>.
- DHd – Digital Humanities im deutschsprachigen Raum, www.dig-hum.de.
- Eclipse, <http://www.eclipse.org>.
- forTEXT, <http://www.fortext.net>.
- GitHub, mpetris [d.i. Marco Petris], CATMA, <https://github.com/mpetris/catma/issues?q=is%3Aopen+is%3Aissue+label%3Aenhancement>.
- GNU, Affero General Public License, <http://www.gnu.org/licenses/agpl-3.0.de.html>.
- GNU, General Public License 3, EN, <http://www.gnu.org/licenses/gpl-3.0.de.html>.
- GNU, Lesser General Public License 3, DE, <http://www.gnu.de/documents/lgpl-3.0.de.html>.
- heureCLÉA, <http://heureclea.de>.
- MAXQDA, <http://www.maxqda.de/>.
- MyExperiment, <http://www.myexperiment.org/>.
- NINES, Nineteenth-Century Scholarship Online, <http://www.nines.org>.
- Open RBAC, http://www.openrbac.de/en_startup.xml.
- Open Source Initiative, The 2-Clause BSD License, <https://opensource.org/licenses/BSD-2-Clause>.
- Pliny, <http://pliny.cch.kcl.ac.uk>.
- Project Bamboo, <http://www.projectbamboo.org/>.
- Text Encoding Initiative (TEI), <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/FS.html>.
- The R Project for Statistical Computing, <https://www.r-project.org/>.
- WebLicht, FAQ, http://weblicht.sfs.uni-tuebingen.de/weblichtwiki/index.php/FAQ#Can_I_use_WebLicht_even_if_my_institute_is_not_on_the_provider_li st.3F.
- Weka 3, Machine Learning Software in Java, <http://www.cs.waikato.ac.nz/ml/weka/>.

⁷² Sämtliche zitierten Online-Ressourcen wurden letztmalig am 8.1.2016 eingesehen.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

