From TACT to CATMA or

A mindful approach to text annotation and analysis

I. Thinking about interpretation

In his 2008 "Thinking about interpretation. *Pliny* and Scholarship in the Humanities" John Bradley set out with the sobering observation that irrespective of some 50 years of research into humanities computing,

> (…) our effect on how most scholars work has been very small. Although tremendously innovative techniques have been developed by members of our community, few, if any, scholars from outside the DH community have taken them up.
>
> (Bradley 2008:263)

Another decade later this assessment, unfortunately, still holds. DH may well have turned out the most successful *institutional* venture in the humanities since the millennium–alas, no other recent methodological 'turn' in the Humanities has resulted in a comparable number of dedicated funding lines, the founding of institutional entities such as departments and schools, the establishing of BA, MA and PhD curricula and degrees, and a significant demand for qualified junior academics.[1] But this metric is biased: for DH's *conceptual* role in and for the humanities at large, seen from the perspective of the traditional disciplines, is at best still that of a *Hilfsdisziplin* (ancillary science) and at worst that of a *parvenu* competitor who managed to nail a flimsy humanistic flag to the post of digitization.[2]

Indeed, if one settles for the modest former role the question becomes even more perplexing: why is there so little interest in DH's digital tools and methods among traditionalists? One methodological lacuna which might contribute to the lack of DH uptake was already identified by Bradley (2008)[3] who found existing DH tools to be

---

[1] Da 2019 observes that "(r)esources unimaginable in any other part of the humanities are being redirected toward it (computational literary studies; JCM), and things like positions, hiring and promotion, publishing opportunities, and grant money are all affected." Da finds the trend to be problematic not primarily because it disadvantages traditional literary studies, but rather because of a lack of methodological rigor and quality control in CLS's quantitative research practice.

[2] Underwood's (2019) *Distant Horizons. Digital Evidence and Literary Change* has once again sparked a vivid debate on whether quantitative digital approaches offer added intellectual value to literary history. See e.g. Da 2019 on "The Digital Humanities Debacle" and Underwood 2019b as reply, both of which form part of an ongoing critical discourse in the *The Chronicle of Higher Education* published under the collective title "The Digital Humanities War" (containing i.a. Brennan 2017; Conrad 2014; Uprichard 2014; Fitzpatrick 2011).

[3] In our 2016 DFG (German Research Foundation) grant proposal for the DH dissemination project *forTEXT* we analyzed the DFG project database GEPRIS in order to establish the proportion of successful funding applications in literary studies which had employed DH methods in the widest sense during the 2005 to 2015 period. We found that the database had recorded a total of 2825 research proposals in literary studies; in 49 instances the relevant project abstracts had contained the keywords 'digital' and 'literature'. However, on closer inspection it turned out that only fifteen of these proposals (0.5%) were substantially related

conceptually at odds with many traditional humanists' reliance on associative, at times unstructured and recursive routines of exploring, comparing and mapping source documents, secondary documents, and external references in a cumulative fashion–in other words, on a *modus operandi* that defies computational formalization in terms of linear workflows. Bradley's own development *Pliny* was therefore a conscious attempt to prototype a working environment that would respect and support such exploratory practices.

More recent developments in DH methods have taken a different tack. Computational Literary Studies (CLS) in particular has spearheaded a trend toward quantitative modeling and analysis of literary data, be it primary (actual literary texts) or secondary (e.g., reception data, bibliometric data). The new approach has sparked considerable criticism, a lot of it polemic and ill-informed, yet some also voicing noteworthy concerns about a lack of rigor and transparency in how the respective quantitative methods are being selected and applied, as well as about the tendency of practitioners to present speculative, analogy-based rather than evidence-based justifications for having chosen a quantitative approach toward an object domain which in and by itself is phenomenologically extremely complex, yet at the same time not really a numeric 'big data' phenomenon per se. Or as Da (2019) puts it: "The thing about literature is that there isn't a lot of it, comparatively speaking."

The sparsity of raw digital data can however be compensated for by casting one's net beyond primary texts and other cultural objects and follow the example of the Social Sciences, namely: shift DH's interest from the phenomenology of the object itself to the empirical traces of the social practices around it, and from the unique expression manifested in the form of an individual symbolic artefact to the multitude of manifestations of historical practices motivated by, and at the same time shaping entire classes and genres of artefacts, such as texts, paintings, performances, etc.

This reorientation was proclaimed a future necessity already some forty years ago, i.e. long before Moretti coined the term *distant reading*. However, this initial call was motivated not by pragmatic but by conceptual considerations: In 1978 Susan Wittig found the (then) field of Humanities Computing to be methodologically constrained by its unconscious allegiance to American New Criticism which elevated the artistic object to a self-contained sign system. Influenced by contemporary reader response theory, Wittig argued that one would have to re-think the notion of "text" as such in order for Humanities Computing to become more relevant to textual studies; she concluded:

> I am suggesting that we turn from our analyses of the signal system of the text to a new study of how, and why, and under what conditions, the text is fulfilled with meaning by its readers. (Wittig 1978:214)

Against this backdrop, Bradley's 2008 "Thinking about Interpretation" was published just as DH approached their next cross-roads, the one where data science and statistics would intersect with declarative and taxonomy driven methods of computational literary studies. The express aim of *Pliny* was to serve as a proof of concept, namely that a computational approach is not necessarily deductive, or as Bradley (2008:19) states: "Pliny is meant to support scholarship when it is still "pre-

---

to digital methods: twelve of them were digitization projects, and a mere three (= 0.11%) had actually applied DH methods in practice.

ontological" – before concepts and their relationships to evidence from sources have solidified."[4]

Today's statistical and probabilistic approach toward humanistic objects is by comparison 'post-ontological': it is data driven, no longer theory or taxonomy driven. Or so it seems. Consider for example the modeling of a semantic theme in terms of a statistical TOPIC: the approach is based on the principal assumption that this aspect of human language use, whatever the intentional motives transparent to the speakers, can be adequately modeled by way of a context-blind genetic algorithm. In the specific case of the LDA algorithm of Blei, Ng and Jordan (2003) this assumption lead to the idea of *latent* expression of TOPICS through words. A word is thus not conceptualized as a pre-determined or intentionally selected Saussurean surface-level vector from a *signifiant* to a *signifié*, but rather as a node from which a multitude of stronger and weaker links reach out across the document's network. As a humanist one could argue that LDA thus implicitly acknowledges polysemy–but this is of course pointing out a conceptual 'family resemblance' rather than a logical connection. After all, the algorithm models collocation probability, not semantics, for it is conceptually an import from gene analysis (cf. Pritchard, Stephens, and Donnelly 2000).

Similar conceptual premises do of course also abound in the seemingly 'pre-ontological' practices exercised by the traditional humanities scholars who *Pliny* aimed to support. Yet the hermeneutic circle, when travelled individually, does not necessarily need to be modeled or formalized in order to function. However, if we want to further our critical discourse we will aim for a clearer understanding of how the traditional humanities progress from the unstructured to the structured: exploratory, recursively, in trial and error mode and most of all, by way of critical discourse, variation and continuous approximation. Developing a system like Pliny is therefore not just a matter of providing a handy tool; it is at the same time an exercise in modeling and making more explicit established pre-digital research practices in the humanities.

The digital turn has presented the humanities with a unique opportunity to reconceptualize their objects, and their practices, in terms of a double take on 'structuring the unstructured.' Reconceptualizing our traditional objects of study–texts, paintings, music etc.–has been made easier by technology; we can nowadays almost effortlessly transform the fleeting and continuous sensual phenomena that are presented to us in various modalities into the abstract lingua franca of digital data: into discrete, computable points of observation. Reconceptualizing and explicating in terms of their complex logic and workflows the humanistic practices by which we operate on these objects, whether presented in digital or in 'analogue' form, is a more complicated thing. Indeed, to formalize and model an epistemology as well as an epistemic field of practice–that is, the implicit assumptions, explicit theorems and exploratory, analytical and synthetic methods which a domain specific discipline has developed over time–from the perspective of measurability and computability is a formidable task. Moretti's (2000) "Conjectures on World Literature" presented an attempt to showcase the potential of such an undertaking

[4] In a witty intellectual *myse–en–abyme*, Bradley (2008) himself uses excerpts from Wittig's 1978 article to demonstrate *Pliny's* UI and functionality.

for one such discipline, comparative literature. For Humanities Computing as it relates to the humanities in general this prospect had however already been pointed out some ten years prior by McCarty (1996) who observed that tools, such as the new digital tools,

> (…) are perceptual agents. A new tools is not just a bigger lever and more secure fulcrum, rather a new way of conceptualizing the world, e.g. as something that can be levered.

The digital tool whose conceptual affordances motivated McCarty's reflection in this instance was TACT[5], a suite of text analytical computing tools developed at the University of Toronto. TACT's principal designer was, again, John Bradley, and it is fascinating to re-read his concise description of the program's functionality some thirty years later as an implicit anticipation of its epistemological leveraging potential. Bradley found three functional aspects of TACT to be particularly relevant: interactivity, index based text analysis, and the ability to process text with dense structural markup that may be organized in multiple, parallel hierarchies. Among contemporary readers familiar with the then nascent OHCO debate,[6] the latter feature would indeed have deserved particular attention–but it seems that Bradley was simply too modest. Mentioning this particular aspect only in passing he wrote:

> *TACT* is interactive. It specializes in quickly answering questions related to a work's vocabulary. *TACT* achieves this relatively quick response time by working with a textual database, which contains not only the text, but a complete index of all the word forms in the text, with pointers to their position in the text. (…) *TACT* was designed to support texts with a rich structural Markup. Within *TACT* you can code such things as page numbers, speakers in a play, or other types of structural divisions. (…) Furthermore, the different tags do not need to fit into a single hierarchical structure. Indeed, multiple hierarchical structures can be represented in parallel. (Bradley 1991)

---

[5] The original TACT manual is available at http://projects.chass.utoronto.ca/tact/TACT/tact0.html

[6] A vivid debate about the pros and cons of a hierarchical representation of text was sparked among other by De Rose et al. (1990). Galey (2011:112, note 24) summarizes as follows: "Specifically, DeRose invokes the idea that all texts have an essential structure in the form of an Ordered Hierarchy of Content Objects (OHCO), a tree structure of non-overlapping nodes that conveniently matches the structure of all XML documents. The debate over the OHCO theory of texts divided critics (…) with DeRose, Alan Renear, and their co-authors on the pro-OHCO side, and opposing them McGann, Hayles, and others with links to textual scholarship. From a textual studies perspective, the OHCO thesis lost in theory but won in practice (…) the OHCO model is everywhere in our digital tools, from the structure of XML documents, to the historical core of the TEI guidelines (…)". This debate is far from over as a recent thread on HUMANIST on "the McGann-Renear debate" (see https://dhhumanist.org/volume/32/423/ et passim) demonstrates. In this context Peter Robinson (HUMANIST posting 32.424), after "years of barking up the wrong 'overlapping hierarchies' tree" himself, proposes an alternative model. He invites us to consider "text NOT being a single stream with multiple overlapping hierarchies. Instead, text is better modelled as a set of leaves, with each leaf potentially present in multiple tree-like hierarchies." Bradley's TACT could in fact have supported this model to a degree as it used the COCOA tagging convention which knows no explicit closing tags.

II. Modeling the cyclical knowledge generation process

In the late 1970s many of the humanities disciplines began to refocus from the investigation of canonized aesthetic artefacts onto the analysis and critique of norms and preferences that manifest itself in how a society defines its canons and how it engages with them in cultural practice. If indeed contemporary DH has begun to follow the same post-structuralist trajectory then ours is nevertheless a somewhat differently motivated *cultural turn*. For its ideological motivation, the enlightened historical-critical interest in the 'slaughterhouse' of the extra-canonical, goes uncannily hand-in-hand with the methodological exigencies of big data centered research.

We should therefore take care not to fall for the empiricist 'data science' narrative and rather consider the range of methodological options in more abstract terms. In dealing with symbolic artefacts and practices computationally DH can

- investigate such symbolic artefacts and practices directly, but restrict the analytical procedures to interpretation free surface level phenomena, objective structural properties and taxonomically robust meta-data; or
- re-define the object domain as such and focus on data-intense second-order phenomena of 'signs in practice' which manifest themselves around defined types of symbolic objects and practices; or
- attempt to model the traditional hermeneutic approach to symbolic artefacts and practices using computational means, and then methodically scale up from scarce data and exemplary exploration to more extensive and robust experimental configurations.

These options are neither mutually exclusive, nor prescriptive: they constitute ideal types that may help us to identify better the nature of our own approach. The one which I will present in the following falls into the third category. Its strategy is to push the limits of the qualitative approach against the backdrop of a more complex, hermeneutic text- and text annotation model. This vision is the underpinning of CATMA (https://catma.de), an open source software and web application for collaborative text annotation and analysis.[7] Its development began in 2008 and, thanks to project grants awarded by various funding agencies and bodies (including the Universität Hamburg, the German Academic Exchange Service DAAD, Google Inc., the German Ministry for Science and Education BMBF and the German Research Foundation DFG) has been ongoing since. CATMA is related to Bradley's TACT not only acronymically (*Computer Assisted Text Markup and Annotation* vs. *Textual Analysis Computing Tools*), but indeed conceptually.[8] And this conceptual affinity can be precisely defined as

[7] CATMA 6 power users can also use the Gitlab REST API and the git protocol to access CATMA as a webservice.

[8] A brief historical *excursus*: TACT was originally designed and developed by Bradley, with the support of Ian Lancashire, as a desktop suite of programs for DOS computers (see https://projects.chass.utoronto.ca/chwp/bradley/ ; Lancashire et.al 1996). In order to apply it in my own research, but even more so in my teaching at Hamburg University from 1994

$$Ix = i\,(Ix{-}\alpha,\ s(Ix{-}\beta,\ t(x{-}\beta)),\ t(x{-}\alpha))$$

This formula is not my invention; it is the brainchild of Manfred Thaller (2018) who uses it to pin down his core concern with the way in which computer science thinks about 'information.' Thaller explains it as follows:

> To be read as: The information available at time x is the result of an interpretative process i() which has interpreted the information available at an earlier point of time x–α over the time span t between x and α, in the context of a knowledge generating process s(). This knowledge generating process in turn has been running over the time span t between x and β, using the available information at the point of the time preceding x by β. (Thaller 2018)

Thaller (2018) concludes that the "implication of the ideas above is, that no such thing as static information exists; 'representing it' just captures a snapshot of a continuously running algorithm." In the annotation model presented below I will build on this observation and refer to the *axis of process*. However, this functional model will be extended further by an *axis of discourse* and an *axis of context*. All three axes need to be taken into consideration in order to conceptualize annotation as an interpretive (rather than merely declarative) activity that contributes to *meaning-making*, albeit on an elementary level.

III. From meaning-making to hermeneutics

Meaning-making as a defining desire and activity in humans was brought to particular attention by the psychiatrist and holocaust survivor Viktor Frankl in his 1946 book "Man's Search for Meaning". Of course, the subjective interpretation of life events and experiences in existential terms is one thing; the interpretation of symbolic phenomena, such as texts, which is motivated by a defined (pragmatic or aesthetic) interest is quite another–and even more so when the latter activity is undertaken in a disciplinary context which stipulates a theoretical and methodological framework. Yet in a structural perspective both are variants of the same semiotic activity: the activation of referential vectors from *signifiant* to *signifié*.

Against this backdrop highly context dependent, unstructured interpretation practices constitute a particular methodological challenge for DH formalization. The more *meaning-making* like, that is, the more subjective, historically contingent and idiosyncratic someone's interpretation of a given text, the less likely we are to capture all the variables and factors that have gone into producing the interpretive output. But this boundary is not incontestable, provided we gain a clearer understanding of what interpretation itself actually *is* or rather, has developed into as a scholarly practice over time.

Today's practice of philological text interpretation is indebted, among other, to the development of the method of explication of textual meaning known as *hermeneutics*. Its theoretical and philosophical reflection as a scholarly method

---

onward I therefore had to use DOS emulators. This alienating effect was in fact a pedagogical advantage: what happened to and what one did with digital text in TACT was completely transparent and the result of a step-by-step interaction between user and machine where the roles were clearly defined. In 2007 the idea of re-implementing TACT as a desktop application for Windows was eventually hatched. For further details on its development history see https://catma.de/documentation/history/.

begins with Schleiermacher and others in the late 18th century.[9] As such it is based on two conceptual tenets: one, the interdependency of analytic and synthetic approaches to text which Friedrich von Ast (1808), the inventor of the term *hermeneutics*, stipulated as follows: "The foundational law of all understanding and knowledge is to find the spirit of the whole through the individual, and through the whole to grasp the individual." Two, the subjectivity, context dependency and hence, historicity of interpretation which therefore cannot be conceptualized as a simple, unilinear transformation of 'text' into 'meaning,' but which must rather be understood as an open ended, recursive process of approximation. These two characteristics contribute to an epistemological condition later to become known as the 'hermeneutic circle.'

Acknowledgment vs. refutation of the historicity of textual meaning, the big *leitmotif* of 19th century Western thought, as well as opposing views on whether it is the author's or the reader's prerogative to determine textual meaning have continued to shape theories and methods of text interpretation from Hegel to Nietzsche. This debate has continued well into the 20th century, from Russian Formalism to Structuralism, Post Structuralism and Deconstruction. But the main *methodological* innovation introduced by 18th century hermeneutics into the practice of textual interpretation has remained uncontested: our exegetic practices are based on the premise that the meaning of a text cannot be fixed dogmatically, but should rather be the product of rational discourse which takes into account textual (linguistic and structural) as well as contextual (historical) evidence. Or to put it differently, hermeneutics has introduced us to the idea that the interpretation of textual meaning is necessarily *parameterized* and *dynamic*.

What, then, is the *methodological* constraint that has to date precluded the hermeneutic activity of parameterized dynamic interpretation from being successfully modelled and supported by 21st century DH?

IV. Reading vs. interpreting

At the core of this problem lies the distinction between first order "Bedeutung" (pragmatic *meaning* as the referential denotation regularly assigned to a given lexical term) and second order "Sinn" (*sense*, the value and subjective importance *for us* that we assign to a word or a phrase as encountered in a given context that is both textual, and existential) which the mathematician, logician and philosopher Frege

[9] As a scholarly method the hermeneutic approach built on a tradition of systematic exegesis of scripture that had evolved in practical theology and religious studies since the Middle Ages, which in turn drew on the teachings of classical rhetoric. Up until the Reformation orthodox biblical exegesis had been based on the claim that correct and true readings of the text can only be determined by the religious authority in power. This is the concept of the authoritative, dogmatic interpretation. Both the religious and the interpretive dogmatic authority were contested among others by Luther who then proclaimed the principle of *sola scriptura*–a call to revisit the original text and rid biblical scripture of the layers of interpretive appropriation by the Roman Catholic church. If Luther's 16th century paradigm shift was essentially one of radical de-contextualization, then the 18th century Romantics counter proposal of hermeneutics was to re-introduce historicity–though this time as a flexible and subjective frame of reference that aims to acknowledge the historicity of the production as well as of the interpretation of a given text.

(1892) has pointed out. First order *meaning* or denotation is relatively easy to look up and deduce; this is the activity which we normally call *reading*. Humans can do it, and machines can do it equally well (if not better and faster) provided the text is grammatical, and the correct grammar and lexicon are available. Of course, for a text to become 'machine readable' in a technical sense some preceding operations will have to be performed, such as the translation of pixels into letters and other typographical information in ASCII or Unicode encoding and then further into a TEI notation. But the principle remains the same: *reading*, whether performed by a machine or by a human being, relies on rule-based transformation, look-up and combination procedures. In other words, it is driven by and can be modeled via formulae.

What clearly sets the human reader apart from the computational is his or her response to an irregular, non-grammatical or innovative case of language use. Unless provided with a choice of grammars and rule sets the machine reader will stop to operate and return an error message. As human readers we tend to react differently–we will try to naturalize, to 'make sense' of the apparent 'error.' The first strategy for doing so is to try and correct the text in order to make it grammatical again. If that attempt fails a second strategy comes into play: We can switch from first order exploration of domain specific denotational meaning to a less formalized, more flexible kind of grammar that enables us to evaluate the statement in terms of Frege's definition of *sense*, of *meaning in existential context*. This seemingly redundant hermeneutic iteration is triggered as soon as we are not satisfied with a mere 'reading' of the passage in question. In such an instance the mere denotational *reading* does not *make sense*, that is to say: it fails to explain the motivational context and backdrop of the utterance. Indeed, very often human readers are not satisfied with finding out *what* has been said anyhow–they also want to know *why* it has been said and what the *relevance* of the utterance is.

Once we find ourselves at the threshold of *meaning-making* in this emphatic sense, things become significantly more complicated as suddenly a multitude of perspectives opens up–for example, *relevance* as perceived by the speaker (narrator), or by the author, or by the reader him- or herself. *Making sense* of statements that cannot simply be 'read' and taken at face (and even less so: at linguistic surface) value is essentially what hermeneutics enables us to do in a controlled fashion. First order grammar, like any formalism, tries to capture the *logic* of the phenomenon (in this case: language use) in an abstract, de-contextualized and generalizable manner. Hermeneutics however adds contingency to logic by considering the criterion of *relevance*. It does so by re-introducing the notion of context dependency into the conceptual model of the linguistic or symbolic phenomena that we encounter. Scholarly hermeneutics in the tradition of Schleiermacher and Gadamer focuses on this second-order functional dimension and stipulates criteria such as plausibility, discursiveness and rationality, salience etc., which one might consider as regulative filters in its processing formalism. Hermeneutics is thus neither 'ungrammatical,' nor in principle impossible to support by computational means–rather, it employs multi-level operating principles of interpretation which are more complex and challenging to express in a well-formed mathematical or logical formalism.

V. Modeling the hermeneutic circle: markup as annotation[10]

How can we tackle this problem in developing an annotation tool? By conceptualizing digital text markup as one specific implementation of a more general, fundamental and richer practice that plays a crucial role in the hermeneutic approach to text as described above: that of text annotation. For in a cultural as well as a methodological perspective, text annotation is not only markup's historical forerunner, but also constitutes a significantly richer and more varied meta-textual practice. One of the most prominent examples in this regard is the technique of interlinear annotation used by monastic scribes. It demonstrates how long before the digital turn different text types and their pragmatic function–such as religious and juridic exegesis of scripture and law–triggered the development of conventions for annotating and referencing source texts. These conventions are the antecedents of today's formal referencing schemata, and markup as a technique of adding declarative metadata to digital source documents is thus merely a recent, technology driven derivate that employs a new set of media-specific conventions.

Markup itself is of course also varied. Coombs et al. (1987) were among the first to propose a systematic, functional differentiation of these variants by distinguishing between *punctuational, presentational, procedural, descriptive, referential*, and *meta-markup*. At the same time, the authors already highlighted descriptive markup as the variant of particular relevance to the human reader.[11] Twenty-five years later Nyhan (2012:123) makes a similar point: she observes that 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."

However, the descriptive markup which Coombs et al. as well as Nyhan refer to is in practice declarative rather than interpretive: The descriptive schema, its categories and the type as well as the range of possible values which can be assigned to a selected character string–a morpheme, word, sentence, paragraph etc.–are in most instances predefined and cannot be extended or modified *ad hoc*. In other words, the 'description' is again a declaration (note Nyhan's qualification "consistently" in the above quote); it is constrained by a defined ontology and metric, both of which remain agnostic to the specific research question and text under investigation. To 'describe' a text document in TEI or to parse it and apply automatic POS tagging are thus operations based on the same deductive approach: in either case the referenced character string–be it a single word, be it an entire document–is conceptually sorted into an abstract table and assigned one or more values therein. This operation proceeds top down, not bottom up, as the table itself remains non-negotiable.

To date only few DH scholars have reacted to the methodological reductionism inherent to all types of digital markup listed by Coombs et al.–declarative, procedural, representational etc. –by explicitly calling for the development of a completely different, namely an interpretive or hermeneutic markup concept. Piez

---

[10] Section V of this paper is a translated, expanded and differently contextualized version of parts of sections 2–4 in my (German language) article J.C.Meister, "Annotation als Mark-Up *avant la lettre*" in Jannidis et.al, 2020 (forthcoming).

[11] The typology was subsequently called into question by Renear (2000) who found that "the descriptive/procedural distinction is flawed."

(2010) for example demands a “markup that is deliberately interpretive;“ such a type of markup would be “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.” But as rightly emphasized by Caton (2000) it is in fact not the choice of markup schema which counts in this hermeneutic perspective; rather it is the underlying concept of text as such. He comments:

> When OHCO 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 (...). 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. (Caton 2000)

This is the model of ‘interactional encoding’–and to implement it in a digital tool we will have to relativize (but not necessarily discard: disagreement on certain textual features expressed via markup can only become productive against the backdrop of conventional ‘ground truths’) the ideal of reaching perfect inter annotator agreement. At the same time, it would be naïve to ignore that the declaration of absolute, objective norms tends to serve a methodological as much as an ideological purpose. In DH the rationale for declaring inter annotator agreement as a normative goal is equally programmatic as much as pragmatic: for example, machine learning, which holds substantial promises for the automation of aspects of humanities research practice, benefits substantially if the machine can be trained on unambiguous ‘gold standard’ annotation data.

Current introductions to DH nevertheless tend to present the ideal of non-ambiguous text markup as an undisputed norm.[12] This technological pragmatism is indicative of a methodological problem which van Zundert (2016) has termed the *Computationality of Hermeneutics*. Van Zundert postulates that hermeneutic considerations should no longer be addressed merely ‘after the algorithmic fact,’ but rather upfront. In other words, hermeneutic desiderata should already inform the computer science aided development of the concepts, codes and models which form the basis for any digital representation and analysis of life world phenomena and aesthetic artefacts.

As far as annotation is concerned, the main question to be considered ‘before the algorithmic fact’ is of a pragmatic order: why do readers actually bother to comment on a text in the first place? All variants of meta-lingual utterances, I would hold, have in common the same rhetorical motivation: to make explicit, document and share one or more observations and understandings of a source text, or of a part thereof.

[12] See e.g. Allen Renear (2004); Andrea Rapp (2017).

Annotation is thus always a type of communication with the next reader–whether it is expressed in the marginal form of a <b> tag or as eloquently as a commentary in an editorial footnote makes no difference. In other words, hermeneutics calls on us to conceptualize annotation (and thus in principle terms also markup) from the point of discourse pragmatics. In this perspective two boundary conditions of annotation become apparent: one, annotation is necessarily a form of meta text relating to an object text. Once annotation loses this nexus and becomes autonomous it turns into an object text itself. [13] Two, annotation is ideally a communication directed at someone other than merely the annotator him or herself. Where it turns into an auto-communication it attains the quality of a *Privatsprache* (Wittgenstein) which may of course still have an aesthetic or mnemonic function, but no longer a discursive one.[14]

A hermeneutically inspired DH practice therefore requires a compatible model of markup which is conceptualized primarily in a discourse pragmatic rather than in a technological perspective. Such a model must be able to capture and represent the logic and workflow of practices that go beyond the base level encoding and declarative explication of object data, in particular the philological and critical practices not yet (or perhaps even not necessarily at all) oriented towards enabling a computer to perform algorithmic DH operations on the source text. *Horribile dictu*: putting this model into practice will also require us to not only tolerate, but in fact *facilitate* via a digital tool the communication, to the 'next reader,' of inter annotator disagreement, ambiguity, and polyvalence, and the provenance and evolution (or 'versioning') of annotations across annotators and annotation iterations. But we can only strive for these goals if we accept the premise that non-contradiction and consistency are neither an intrinsic requirement of annotation, nor of markup. Indeed, both criteria constitute pragmatically and technologically defined constraints which in most instances are 'algorithmic facts' rather than phenomenological essentials. Depending on their pragmatic purpose different types of markup will thus require different types of specification; however, all types of markup must share the fundamental discourse pragmatic, communicative ethos of annotation.

But what exactly do we mean when we refer to 'annotation'? Unsworth (2000) lists *annotating* as one of seven 'scholarly primitives'–*discovering, annotating, comparing, referring, sampling, illustrating, representing*–which represent the fundamental and widely shared epistemic practices of humanistic research at large. In traditional literary studies, for example, this fundamental practice is encountered at three levels of complexity:

1. Base-level markup of linguistic, formal, and structural features of text (layout, typography, grammar, and structural entities such as verse, paragraph,

[13] My definition of annotation as a means to communicate a specific 'reading' positions it, in the function of a hermeneutic mediator, between the source document and this document's potential realm of application and relevance. This corresponds to Boot's (2009) concept of *mesotext* which serves as an epistemic springboard from the source text bound trias of *textdata – mesodata – mesotext* to the fully articulated, medially independent meta text.

[14] As an example for the latter see McGann's (2014) proposal for a so-called 'topological markup' which he bases on the idea of patacriticism, i.e. a 'theory of subjective interpretation' which focuses on the reader's engagement with the autopoetic function of an aesthetic text.

chapter etc.) that in and by themselves are semantically neutral, i.e. do not carry an inherent or conventionalized meaning.

2. Explication of local semantic phenomena–this variant requires the annotator to process and interpret the semantic content of a larger section of text, i.e. one that can be read as a particular statement or proposition about the text's reference domain, or about the text and its functions itself. This is the medium level of complexity which Piez (2010) refers to as 'hermeneutic annotation.' While annotators tend to makes use of disciplinary terminology in order to explicate semantics at this medium level they will generally not take recourse to a specified taxonomy.

3. Relevance and *meaning* oriented text commentary, which explains, contextualizes and cross-references specific features, statements and propositions of a text against the backdrop of a holistic interpretation of the entire document, and with a view to linking it to larger entities–such as an author's work, an epoch, a genre, a critical discourse, a socio-historical trend or an aesthetic program etc. Ideally, this type of philological text commentary should operate within the confines of a fully developed interpretive theory; in reality, however, such theoretical premises are often communicated only implicitly. Because of their complexity, their contextual reach and their exploratory, highly contingent nature text commentaries cannot be modelled and produced digitally.

This three-level distinction differentiates annotation types along the axis of increasing semantic and thus hermeneutic complexity. With regards to technological complexity one might complement this with a second systematic which distinguishes annotation types in terms of the medial distance between an annotation and its reference domain–the annotated string–in the source text. In the print medium certain types of elementary annotation were inscribed directly into the source string, for example by using bold characters. By contrast all SGML based digital texts conceptually 'unflatten' the layers of source text and annotation right from the start.[15] At the same time the preference for inline tags in many markup schemes conceptually nevertheless emulate as closely as possible the spatial proximity between text and text markup which makes the traditional print medium so comfortable to process (and which is then fully emulated on-screen anyhow). More importantly, the pre-digital traditional practice of text studies had already developed a conventionalized implicit semantics of spatial proximity between source document and annotation: the greater the distance between an object text and the meta textual annotation, the more likely a competent reader is to regard it, in the terminology of Boot (2009), as an interpretive *mesotext* which is destined to contribute to the eventual formulation of an independent secondary *metatext*. In traditional manuscripts and print this process of spatial as well as conceptual distancing begins with the progression from underlining to interlinear annotation and continues via the gloss, the margin commentary, the footnote, the endnote, the apparatus etc. Spatial proximity and distance between source document and annotation have thus attained a discourse pragmatic and rhetoric function–they are indicators for the status and ambit of the communicated 'reading.'

[15] Which, though of course very efficient, is by no means a technological must: instead of using a generic SGML <b> –tag one could also define a unique hexadecimal code for every bold letter.

In the digital medium this valuable processing information can easily get lost when on-screen output conveniently hides all markup and 'flattens' the layers. But there are of course ample technological means (e.g., hover effects, pop-ups, integrated interactive data visualizations etc.) which enable us to preserve and express this functional richness as well. Indeed, the digital modeling of this particular aspect–the semantics of spatial proximity–in the relation between a source text and its annotations serves a fundamental conceptual need that goes far beyond the emulation of traditional practices. This is where the controversy about inline vs. standoff markup becomes conceptual, rather than merely a quibble about technological constraints.[16] One of the main arguments pro standoff markup has been the critique of inline markup's implicit OHCO (Ordered Hierarchy of Content Objects) text model. Simply put, SGML inline markup with closing tags cannot handle hierarchical overlaps in texts, such as that of *enjambement* and verse in a poem, whereas standoff markup can. Yet it is not just the source text whose internal organization defies hierarchical modeling and renders inline markup notoriously problematic: overlapping hierarchies, discontinuity and complex multi-dimensional layering are also characteristics of readerly and scholarly practices and operations performed on a text, as Buzetti (2002) has pointed out.

Philological encounters with texts that are of this double 'overlapping' nature can in fact significantly benefit from giving the 'digital turn' another, more ambitious spin. Before the advent of post-structuralism many scholarly practices could still be adequately modelled on the basis of an essentialist-hierarchical concept of text (primary objects and secondary information resources), libraries (institutions for source object and knowledge management) and a clear-cut distinction of the roles of authors (intentionally acting producers of texts), readers (lay recipients and interpreters of texts) and scholars and critics (authoritative instances). 21st century textual practices however are by contrast characterized by interconnectivity, flexibility of roles and competing conceptualizations of text as such. A practice of "literary annotation in the digital age" (Bauer/Zirker 2013, note 21) should accordingly be modelled as one which is no longer oriented towards text objects in an essentialist sense and defined in terms of static roles, but rather as one comprising a range of processes and events of reading, annotating, interpreting, evaluating, arguing, in short: as a discourse.

The practice of digital annotation therefore requires tools that allow us to conceptualize the source text as well as its annotations alternatively as nodes, or as edges in an n-dimensional, dynamically reconfigurable network of textually encoded information. One of the agents in this network is the reader who, depending on his interest and method of choice, will define, systematize and explore edges, nodes, and clusters for hermeneutic purposes.[17] Digital models and technology make it far easier for this agent to recombine, aggregate, reconfigure source and metadata and capture as well as analyze and feedback processing information. For a digital text

[16] On the Overlapping Hierarchies debate with particular regard to TEI see Pierazzo (2016:316–319); for a more general appraisal see Witt (2004), who proposes a "technique of annotating documents in multiple forms" as an alternative to standoff markup.

[17] This proposal, I believe, corresponds to Robinsons (2019) suggestion to consider texts themselves as multi-dimensional networks not just metaphorically: if texts are multi-dimensional networks, as Robinson proposes, then the operations performed on texts, of which annotating is one, should be modelled accordingly.

hermeneutics this high-level model has some fundamental consequences–most importantly, we are no longer required to conceptualize a text's interpretation as a finite whole. Rather, we can think of it as something dynamic, as one instance of 'output' that was generated from n possible configurations of interpretive and declarative decisions made by one or more readers.

Against this backdrop standoff markup proves particularly suitable in that it follows not a document, but a database centered approach, as Schloen/Schloen (2014) point out:

> […] standoff 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. Standoff 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. (Schloen/Schloen 2014)

Yet if we want to employ standoff markup from the perspective of the database paradigm we must obviously also consider annotation itself as a type of data (meta-)modeling. With regard to data modeling practice in general Flanders/Jannidis (2016) have recently suggested to distinguish *conceptual vs. logical model* and *curation-driven vs. research-driven modelers*. Building on their proposal, I would like to propose a matrix of four prototypical variants of digital annotation in which an annotator might 'data-model' a given source text:

	conceptual model	logical model
interpretive digital annotating	hermeneutic „bottom up" annotation	explorative reading
declarative digital annotating	taxonomic „top down" annotation	formal categorization

Fig. 1: Prototypical variants of digital annotation as data modeling

In this matrix the qualifier *descriptive* has been re-labelled into *interpretive* because the latter fits the suggested discourse pragmatic model of annotation. Whereas declarative 'annotation as data modeling' approaches can already in many instances be performed automatically, exploratory bottom-up annotation can only be formalized via iterative approximation–be it 'manually,' i.e. intellectually by human annotators who can abstract from concrete unstructured annotations by drawing on

contextual knowledge and discussion among one another, or be it by machine learning.[18]

Importantly though, the annotation variants which the above matrix juxtaposes categorically do in fact form a continuum: it makes no sense to distinguish dogmatically between inductive and deductive, between declarative and interpretive, and between 'manual' and automatic modes of annotation. One of the core features of a digital tool claiming to support annotation as a discursive practice of knowledge generation must be the ability to facilitate the gradual progression from structured to unstructured annotation and vice versa along the three axes of *Method, Function*, and *Procedure*—in other words, such a tool needs to conceptualize 'annotation' as a vector within a multi-dimensional space which integrates a pragmatic, an epistemological and a technological dimension (see fig.2).

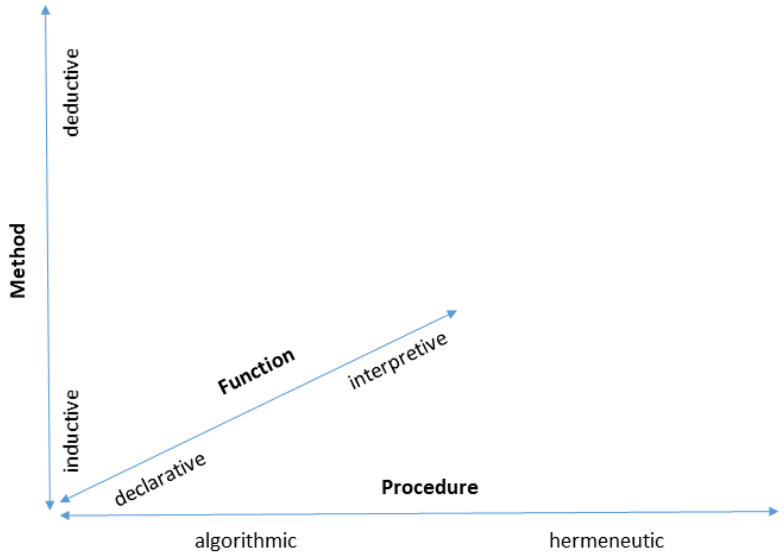

Fig.2: The three axes of digital text annotation

VI. Building CATMA, a web application for collaborative text annotation and analysis

When we set out to build CATMA (*Computer Assisted Text Markup and Analysis*) in 2008 the mission seemed straightforward: re-implement Bradley's DOS based TACT (*Textual Analysis Computing Tools*) as a desktop application for Windows. After consulting TACT's original code, which John Bradley supplied, CATMA's system architecture was developed and then partially implemented by Malte Meister. He tried to emulate the leanness and transparency of TACT's modular architecture and UI in the new architecture and then implemented (in C#) its core function in what later became known as CATMA's *analyzer module*. This is how CATMA's customized,

[18] As an example for this approach see the outcome of the project heureCLÉA in which we used supervised ML to automate narratological high-level annotation of discourse temporality features. See Gius/Jacke 2015.

very powerful query language originated, which is still in use today. By mid-2008 Marco Petris came on board as lead developer, bringing his combined expertise as a commercial systems developer and his keen interest as a scholar of Italian literature and language to the table.

Petris first augmented the analytic function with a separate (written in Java) but integrated *annotator module*. The two modules combined were launched as CATMA 1.0 in 2009. Petris then gradually migrated the analyzer module to Java as well. We launched CATMA 2.0 soon after–and instantly a flow of feature requests by users started changing the scope of our project dramatically. Non-DH scholars in particular found our markup tool helpful and intuitive, but–they wanted more and different features than we had anticipated. Simply put, they were not content with just *marking up* texts; they also wanted to *annotate* them, discuss the annotated phenomena, interpret these, annotate them again or differently, try new tags, share their various resources, from work in progress to entire tag sets, from source texts to analytical results and visualizations etc. It turned out our straightforward software development project had been sucked into the vortex of what is generally referred to as the hermeneutic circle. And so, after ten years of continuous development, the 2019 version CATMA 6 is a far cry from a mere re-implementation of TACT–not just technologically, but more importantly conceptually: from a standalone desktop tool for single users that focused on text annotation and basic analytical functionality inspired by TACT's USEBASE module it has grown into a web application which

- supports single user as well as collaborative text annotation and analysis undertaken by teams;
- works with any UTF 8 encoded text format in almost any language, including right-to-left written ones like Hebrew;
- allows for the import and/or on-the-fly creation of tagsets, and for the specification of tags via–structured and unstructured–properties;
- organizes all workflows around the core concept of a 'project' and facilitates the sharing of tagsets, source texts and corpora, and of course, annotations and meta-annotations themselves;
- generates XML/TEI compatible external standoff markup using the TEI feature structure module and allows users to export results in Excel and csv format, as well as via an API;
- can ingest documents with e.g. TEI inline markup, which will be converted into (so-called *intrinsic*) standoff markup;
- supports overlapping and discontinuous annotating and is technologically 'undogmatic,' i.e. non-prescriptive with regard to the markup schemata and annotation conventions that users might want to specify;
- allows for the interactive analysis of any combination of source text or source corpora and their respective annotation, up to highly complex and deeply nested queries (which can be formulated either directly in CATMA's query language, or via a widget-like natural-language query builder);
- integrates base level automatic annotation functionality like POS tagging, as well as two use-case specific high level automatic markup options for temporal expressions[19];

[19] The algorithms for these functions are the result of a supervised ML analysis of manual CATMA annotations of a corpus of 100 German 19th century short stories in the project

- contains a set of basic 'off-the-shelf' visualizations for CATMA query output as well as a code 'sandbox' to build highly customizable VEGA visualizations that comply with the standards for *hermeneutic visualizations* developed in the 3DH project[20];
- uses graph database technology and an integrated Gitlab based user, team, project and versioning management functionality.

CATMA's system architecture and functional concepts for version 6 are detailed in appendix 1; for continuously updated information see https://catma.de and https://github.com/mpetris/catma.

Overall, CATMA builds on the foundations established by TACT, resulting in a uniquely 'mindful' markup and text analysis tool–that is, in a tool whose development continues to be inspired and is driven primarily by desiderata of humanities research practice. This overall commitment to an approach that considers hermeneutic desiderata 'before the algorithmic fact' is encapsulated in CATMA's hermeneutic data model, which serves as a high-level conceptual scheme:

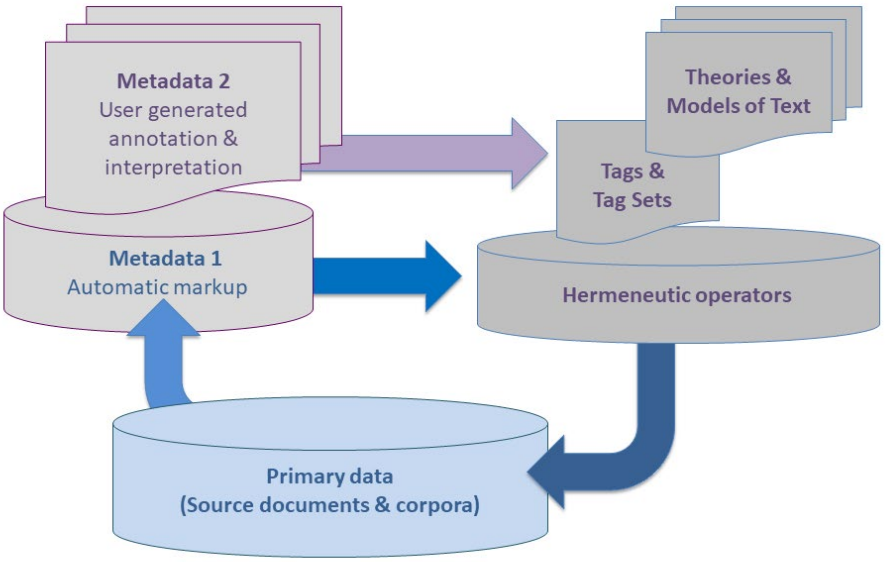

Fig.3: CATMA's hermeneutic data model

VII. Beyond emulation: strengths and weaknesses of standoff markup

A succinct appraisal of the epistemological and cultural relevance of adding meta information to text was formulated by Lou Burnard (2001) who stated: "Text markup is currently the best tool at our disposal for ensuring that the hermeneutic circle continues to turn, that our cultural tradition endures." Thinking about annotation from the database perspective and utilizing the unique flexibility of standoff markup

heureCLÉA (http://heureclea.de). For a discussion of the conceptual approach see Bögel et al. (2015).

[20] See http://threedh.net.

has enabled us to develop CATMA as a digital tool and a working environment for scholars of text and language which not only emulates the traditional disciplines' way of 'keeping the hermeneutic circle turning,' as Burnard called on us to do. In addition to emulation, CATMA also integrates functionalities which were previously not part of the hermeneutic workflow: automatic markup routines, versioning control and management, computational analysis of text and markup, data as well as process visualization, and meta-annotation.

Like annotation itself meta-annotation in CATMA comes in two variants, structured and unstructured. The structured variety enables the user to qualify a selected tag instance in terms of pre-defined attributes and value ranges by assigning a 'property value' to a particular text annotation. The unstructured variant comes in the form of a free-text commentary field into which a user may enter notes, observations, explanations etc. Both variants of meta-annotation can of course be analyzed using CATMA's query language, allowing for complex searches such as (formulated here in semi-natural language, not in CATMA's query language):

"Show me all instances in the corpus where

- a source text string contains a word ending on the string "shire" **and**
- where the word ending on this string has been automatically POS-tagged as SUBJECT **and**
- where the same string was manually annotated by annotator (1) by assigning the tag PROTAGONIST **and**
- where the same annotator (1) instantly qualified this tag instance in terms of the property CERTAINTY **and**
- where the value of this property was set at = 5 **and**
- where two or more annotators **subsequently** added a free-text comment containing one or more strings that possess a SIMILARITY OF >=75% with
- one or more of the phrases {DOUBT, QUESTION, EDITION, CONTEXT} **while**
- DISREGARDING CASE SENSITIVITY in the similarity check.[21]"

The basis for these complex and combined searches across *source text, annotation, meta-annotation* and *annotation timestamp* in CATMA is an implementation of the TEI feature structure tag concept. In the following example I have annotated the phrase

"he felt better than he had for many weeks, a fact (…)"

in Fitzpatrick's *Afternoon of an Author* as a "claim" in terms of a predefined rhetorical tropes tagset. I then added two types of tag instance property information: a structured "plausibility" property whose value I set at "medium," and a free-text commentary intended to remind whoever might want to build on these annotations that this particular qualification needs to be discussed in more general terms because

[21] A query constrained by the conditions specified under the last three bullets effectively analyzes the free text comment as if it were a primary document. This reflexive application of base-level query constraints is already fully implemented for structured CATMA annotations; its extension to the free-text comments is expected for version 6.1. (In case you wonder about the annotated word in the original source document mentioned under bullet 1: it is of course the Cheshire Cat in Carroll's *Alice in Wonderland*).

such opening statements in literary narratives typically aim to condition the reader (the so-called ‘priming effect’).

In an XML export file the relevant section of standoff markup extracted from CATMA’s database takes on this form[22]:

```
<encodingDesc>
   <fsdDecl xml:id="CATMA_08E831DC–EA5F–4367–932E–3A8F2C6D7DA8" n="Rhetorical Tags 2017–08–
08T19:08:12.000+0200">
     <fsDecl xml:id="CATMA_2965A186–18DD–4C47–9C7B–D5061889D7CC" n="2019–06–24T11:16:33.000+0200"
type="CATMA_2965A186–18DD–4C47–9C7B–D5061889D7CC">
       <fsDescr>claim</fsDescr>
       <fDecl xml:id="CATMA_EF8C7681–D140–469F–884A–F4F2905FB79C" name="catma_displaycolor">
        <vRange>
         <vColl>
          <string>–8837951</string>
         </vColl>
        </vRange>
       </fDecl>
       <fDecl xml:id="CATMA_22612A10–E70F–400F–94F1–4D7BB2B6BB12" name="catma_markupauthor">
        <vRange>
         <vColl>
          <string>mail@jcmeister.de</string>
         </vColl>
        </vRange>
       </fDecl>
       <fDecl xml:id="CATMA_32F02054–1031–41DE–B27A–994C1CBA2E4F" name="Plausibility">
        <vRange>
          <vColl>
          <string>low</string>
          <string>medium</string>
           <string>high</string>
         </vColl>
        </vRange>
       </fDecl>
     </fsDecl>
```

and the code section with my two property declarations is

```
<fs xml:id="CATMA_93B49892–DCAB–4785–B1CC–C6CD304E9A4B" type="CATMA_2965A186–18DD–4C47–9C7B–
D5061889D7CC">
   <f name="catma_displaycolor">
    <string>–8837951</string>
   </f>
   <f name="catma_markupauthor">
    <string>mail@jcmeister.de</string>
   </f>
   <f name="Plausibility">
    <vRange>
     <vColl>
      <string>medium</string>
      <string>to be revisited: priming effect</string>
     </vColl>
    </vRange>
   </f>
  </fs>
```

[22] Note that XML is an export/import format only: CATMA 6 stores all tags and annotations internally as JSON data, using the JSON-LD format which is the recommended serialization format for the Web Annotation Data Model on which CATMA annotations are based.

In reality a CATMA user will of course find it much easier to inspect tags and their properties via the UI, by making use of the hover function and/or by inspecting the detailed description of the selected tag instance:

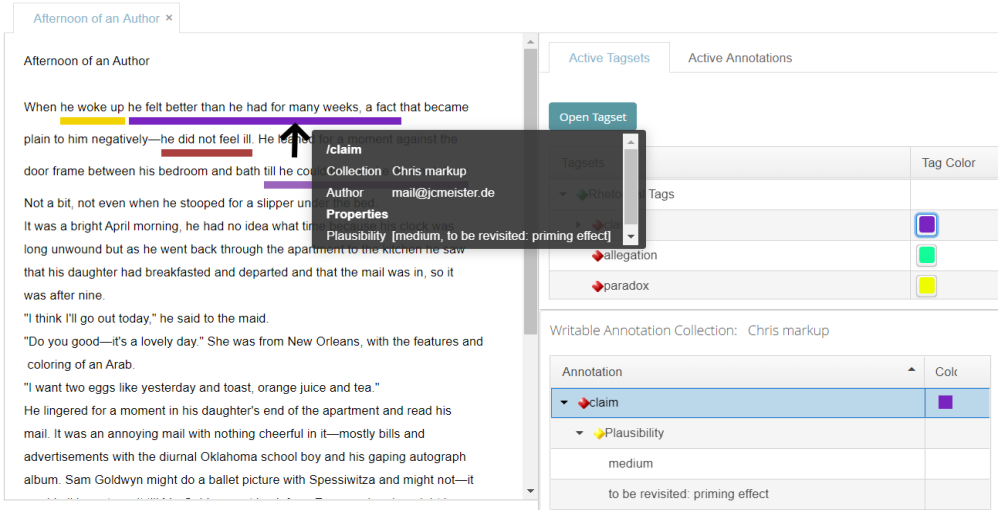

Fig.4: Instance of a rhetorical "claim" tag with "Plausibility" property set to "medium" and a free-text comment by the annotator ("to be revisited: priming effect")

The JSON markup in the CATMA database contains a lot of information, including a unique tag ID, the character standoffs, the definition of tagsets, the assigned tags and properties, their value ranges as well as display features such as color etc. Every entry is also time-stamped, references its annotator as owner as well as the annotation collection to which it belongs.

However, external standoff markup does come with one significant limitation: the annotated source document may not be changed as this would compromise the offsets which reference the character strings. This limitation should not be taken lightly; one of the most prominent feature requests from CATMA users is in fact to facilitate the direct editing of source documents during the annotation and analysis process–not necessarily in an extensive fashion, e.g. by rearranging paragraphs or inserting entire chapters, but at least so that one could attend to trivia such as OCR mistakes, punctuation marks, line breaks etc.

Such an edit feature is difficult to implement in an application that builds on the 'annotation as database'-paradigm and aims to support real-time online collaboration.[23] Recalculating and then re-writing 'on the fly' all subsequent

[23] In a recent publication Kuczera and Neill (2019) claim to have developed „a new approach to the annotation of texts (…) based on standoff properties. These allow for index based multi-dimensional annotations that can be assigned to the relevant users" (my translation; the original reads: "einen neuen Ansatz zur Annotation von Texten (…). Grundlage sind Standoff Properties, die indexbasiert mehrdimensionale Annotationen mit Zuordnung zu den jeweiligen Nutzenden ermöglichen.") Moreover, the authors assert that their markup tool has resolved the problem of editing the source documents in a standoff markup approach. Both claims are a misrepresentation. What the 'new' tool actually does (some ten years after CATMA's first launch as a tool capable of handling inter annotator disagreement, and some three years after we introduced a graph database) is recalculate, subsequent to an on-screen

character offsets across all user specific variants that a team of annotators might have produced for a text corpus using CATMA's collaborative functionality may be possible in theory; in practice it is not feasible. More importantly, taking this route would in fact be paradoxical, for in the 'annotation as database' paradigm the source text is no longer considered as privileged and foundational; rather, it represents one of many nodes in a dynamically evolving network of texts and meta-texts. The logical way to address the problem, then, is to consider an edit operation as a versioning of the source document.[24]

VIII. Integrating visualization as a hermeneutic operator: CATMA and the 3DH project

CATMA's initial hermeneutic data model presented in Fig.3 lists tags and tagsets as well as theories and models of text as 'hermeneutic operators.' But one very important hermeneutic operator is missing in this list: visualization.

In her seminal publication *Graphesis* Drucker (2014) refers to "Visual Forms of Knowledge Production." As far as annotation is concerned, visualization is indeed one of the most powerful and intuitive conceptual enablers which we may use to correlate, investigate and interpret all types of data that are of relevance to the annotation workflow: source documents as well as their annotations and meta-annotations. Bradley's *Pliny* already demonstrated how one might conceptually emulate the logic of the traditional humanists' inductive, explorative workflow using graphical means: rather than resorting to the engineering science's data-driven approach to "visualization-as-product" *Pliny* tried to sketch out the option for a "visualization-as-process" centered approach that uses the desktop metaphor.

CATMA has contained a visualization module since version 3.0. Yet its functionality was limited: the user could plot query results as a line diagram, and from version 5.0

source text edit, the affected standoffs *on the client side*. But this transformation pertains merely to HTML-based screen output and local annotation operations performed in isolation on a client machine; it does not address the complexity of the issue as it presents itself in a collaborative, real-time synchronous online web service like CATMA which relies on a host-side graph database architecture and a Gitlab versioning mechanism.

[24] To be more precise, the source document edit has to be conceptualized as a *multi versioning* problem: each version of a source document belongs to the corresponding versions of its annotations. A change of a source document, i. e. the creation of a new source document version will therefore also imply the creation of new versions of the original annotations. In a multi version setup these version would thus form a meta version. In short, the implementation of a source document editing function in a collaborative working environment and web application like CATMA necessitates a three-dimensional configuration of the version setup across (a) one or more document versions, (b) all their corresponding annotation versions, and (c) all their corresponding tag versions. This can be achieved as long as the altering operations on the source document allow a computation of the impacts on their annotations, just like the changes of a tag allows a computation of the impacts on the annotations in which the tag has been used. In CATMA 6 we already made the first step towards editable source documents: there is a git container that manages all the corresponding versions and that is versioned itself (the meta version). The next step is to compute the impacts of a source document change on its annotations.

onward also in the form of an expandable Double Tree[25] that displays the left and rights contexts of a selected keyword:

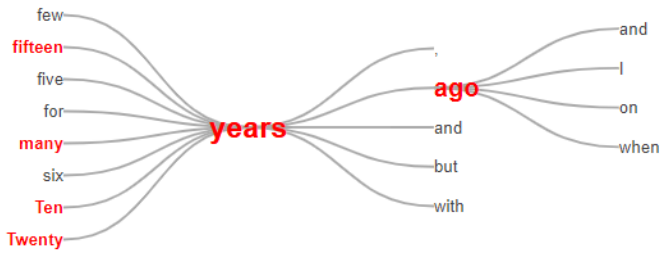

Fig.5: Expandable Double Tree visualization of a keyword in context in CATMA

These visualizations are essentially based on an archetype which, as Drucker (2014:66) has pointed out, we have been culturally conditioned not to 'see' as visualization any longer: the tabular format which maps the conceptual dimensionality of data sorted into columns and lines onto the two axes of spatial dimensionality inherent to all graphic media.

In contemporary DH more fancy and colorful visualizations do of course abound; the choice offered in code libraries such as D3.js is overwhelming. Yet humanists, if anybody, should beware: visualizations can, as Drucker (2011) observes, easily become

> (…) a kind of intellectual Trojan horse, a vehicle through which assumptions about what constitutes information swarm with potent force. These assumptions are cloaked in a rhetoric taken wholesale from the techniques of the empirical sciences that conceals their epistemological biases under a guise of familiarity. So naturalized are the Google maps and bar charts generated from spreadsheets that they pass as unquestioned representations of "what is." This is the hallmark of realist models of knowledge and needs to be subjected to a radical critique to return the humanistic tenets of constructedness and interpretation to the fore. (...) Rendering *observation* (the act of creating a statistical, empirical, or subjective account or image) as if it were *the same as the phenomena observed* collapses the critical distance between the phenomenal world and its interpretation, undoing the basis of interpretation on which humanistic knowledge production is based.

An uncritical, unreflected use of visualisation can thus on the one hand easily result in a conceptual reification of *capta* (Drucker) as *data*. So how can we empower users of DH tools to use this form of "knowledge production" while at the same time bringing to their attention the constructedness of it?

[25] The double-tree visualisation was developed by Chris Culy.

Up to version 5 we tried to achieve this by delegating visualizations to a separate functional module. In CATMA 6 we use a different approach: the UI now seamlessly integrates the epistemic functionality of visualizations with that of query based analyses. But it does so by presenting the user with a choice of four ready-made visualization options on top of its table of query results, and nothing will be rendered before the user has decided to interact with the system and selected one of these options. This emphasis on user activity as a compulsory trigger is the result of a thorough process of theorizing about the requirements of hermeneutic data visualization up front. And users who want to take this critical, reflected approach to visualisation of their data one step further are empowered to do so by a unique technological feature in CATMA: an integrated viewer and editor for the VEGA code which constructs but also deconstructs all visual rendering of output data displayed by the system.

Our more reflected approach to visualisation in CATMA is the fruit of the 3DH project, which we ran in parallel to CATMA's ongoing development, from 2015 to 2017.[26] Our aim in 3DH was to lay the foundations for a 'next-generation,' critical approach to visualization in and for the (digital) humanities: an approach in which the concept of "third dimension" is no longer defined at surface level, i.e. in terms of the traditional z-axis of three-dimensionality that turns the flat image into a mimesis of a physical real-world object. For us the third dimension is that of critical, self-referential reflection which the traditional approaches to data visualization adapted from the empirical science lack. The project's conceptual outcome was therefore the formulation of four postulates for hermeneutic data visualization, summarized in Kleymann (2015) as follows:

1. the "2 way screen postulate" (i.e. an interaction focused approach toward visualisation);
2. the "parallax postulate" (i.e. the idea that visualisation in and for the humanities should not just tolerate, but actively put to use the power of visual multi perspectivity in order to realise epistemic multi perspectivity);
3. the "qualitative postulate" (i.e. the idea that visualisations should not just 'represent' data, but also offer a means to make and exchange qualitative statements about data);
4. the "discursive postulate" (i.e. the idea that visualisations should not just be used to illustrate an already formed argument or line of reasoning, but should also become functional during the preceding/subsequent steps of reasoning, such as exploration of phenomena and data, generation of hypotheses, critique and validation, etc.).

The ready-made visualization options for query output in CATMA 6 from which a user can choose are KWIC (keyword in context), Distribution Graph, Word Cloud, Double Tree (and potentially also Network.) These four types meet the postulated requirements of hermeneutic data visualization in varying degrees: all use a 'linked

[26] The core contributors to this project were members of the original CATMA and heureCLÉA teams (Evelyn Gius, Janina Jacke, Jan Christoph Meister, Marco Petris), visualization experts (Johanna Drucker, Geoffrey Rockwell, Marian Dörk) and 3DH's own Rabea Kleymann and Jan-Erik Stange. In addition, during the summer semester 2016 we gained valuable input from numerous international visualization experts who addressed particular aspects during a 3DH lecture series programmatically titled "A word says more than a thousand pictures"; for details see http://threedh.net.

screen' approach that allows the user to jump from any point in, say, a line chart directly to the relevant string in the source document (first postulate); some of them also enable the user to express and explore multi perspectivity and add to the database qualitative statements (postulates two and three); none of them can however be directly integrated into a discursive argument (postulate four)–for CATMA, unlike Bradley's conceptually more ambitious *Pliny*, is not (yet) a tool or a working environment in which one can in fact *formulate* a coherent meta-text.

However, the 3DH project also resulted in a software prototype that demonstrates the more ambitious, argument centered use of visualisation that we will aim for in the next development phase: *Stereoscope* (http://www.stereoscope.threedh.net) can ingest a CATMA source text and its annotations and meta-annotations. In this prototype all three types of data are automatically visualized. The user can then qualify, discuss, cross-link them; the various annotations generated on a canvas during this process can also be saved as a so-called 'views,' i.e. as a visual snapshot which can be annotated, commented on, and combined with other such views. Visualization is thus considered equally from the perspective of process and output and consequently contributes directly toward the formulation of an elaborate visual-textual argumentation.

Fig.6: *Stereoscope*, a 3DH-compliant prototype that supports the generation, critique and discursive organization of CATMA generated annotation and meta-annotation data

While *Sterescope* is presently not yet integrated into CATMA's productive version, another feature already mentioned enables the user to generate, critique, interact with, and manipulate in principle any type of data visualization that one might choose to explore and interact with CATMA's database via visualizations: All CATMA visualizations are coded in VEGA, a high level visualization language based on Wilkinson's generic "Grammar of Graphics." This underlying VEGA code can be accessed and manipulated directly from within the CATMA UI. The *techne* of visualization is thus no longer a hard-coded wizardry that remains opaque to the user, but is rather made transparent in terms of its definitions, parameter settings,

data flows, interactions and algorithmic transformations that can be inspected and critiqued–just as one inspects, discusses, and critiques the annotations themselves in CATMA.

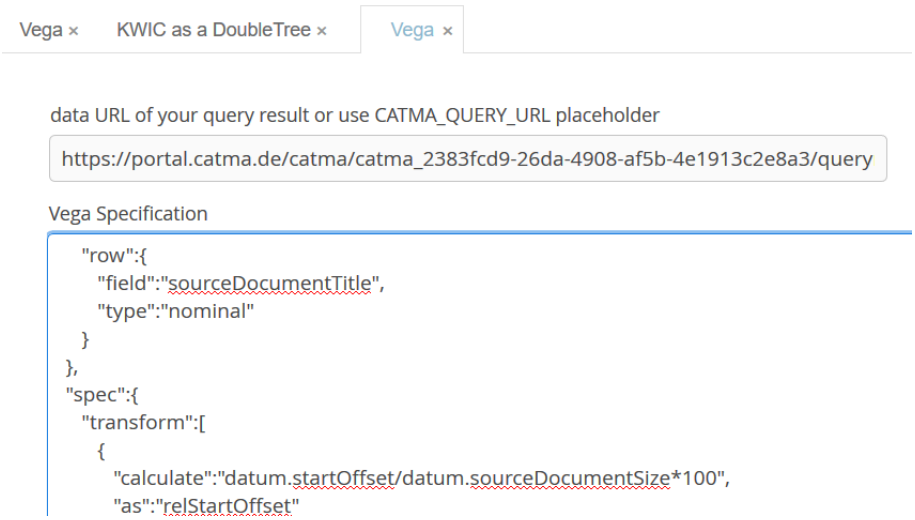

Fig.7: VEGA code editor in CATMA 6 (prototype)

IX. Consensus ex machina? Consensus *qua* machina!

Bradley's TACT and *Pliny* stand out as tools–one robust and applied in numerous text analytical projects in the day and age of Humanities Computing, the other a prototype which brushes algorithmic logic almost against the grain–that demonstrated in exemplary fashion the significant conceptual dimension of software development in and for the humanities which Ramsay and Rockwell (2012) and others have repeatedly pointed out. However, the painstaking manual interaction with source data in close reading mode is no longer the primary focus in contemporary cutting-edge DH research. The methods which at present enjoy attention are those that can contribute to the automated analysis of large corpora, and they obviously come with different constraints and pitfalls than manual annotation. Yet once they have become more robust and reliable we might choose to integrate one or the other not necessarily technically, but certainly conceptually with semi-automatic and manual environments for text annotation and analysis like CATMA. Such an undertaking at 'mixing methods' will of course come with a new risk: that of reifying no longer the data as such, but rather the second order data patterns and structures which approaches such as topic modeling, word2vec, stylometry etc. may generate. However, this problem is a praxeological one and thus a matter of investing equally into 'doing' DH, and into reflecting on how we do what we do in DH. But more importantly, it is a matter of reminding ourselves of the type of phenomenon that we aim to engage with as humanists.

This is a matter of choice, not of dogma. As for me, the primary object domain of the humanities is that of symbolic artefacts, of man-made signs and meaning bearing systems that constantly change, adapt, and impact on us, their observers. This is a field that presents us with a unique methodological challenge because it defies objectivist empiricism. Dynamic feedback between observer and observed may of course nowadays be considered a fundamental epistemological principle of all knowledge generation–but if indeed it is, it does not manifest itself in the same way across epistemic domains. In the humanities it is real and measurable on an everyday basis, and the digital humanities are therefore called upon to become more 'realistic' in their critical, self-reflective approach to knowledge. At the same time digital tools like TACT have helped us to re-conceptualize what was once considered to be a matter of "two worlds" as an epistemological continuum that extends between phenomenological and formal approaches to an object domain. In the practice of textual and language studies this continuum takes on the shape of a methodological triad: that of annotation, analysis, and synthesis. CATMA aims to support scholars of text to explore this epistemological continuum, and to practice this triad of methods in an undogmatic, discursive, and collaborative manner. To this end our web application employs state of the art computational concepts and technology. Conceptually though it remains indebted to Bradley's TACT and *Pliny*.

And so am I. My first contribution to the (then nascent) field of Humanities Computing was a critical intervention against what I considered to be a naïve programmatic vision, namely that of a *consensus ex machina*–this was the title of the 1994 ACH/ALLC conference in Paris–which we could reach by empiricising the humanities, and in particular by employing digital tools.[27] Inspired by TACT and my own fledgling attempts at modeling narrated events in PROLOG in terms of what I might nowadays call "latent actions" I argued that we should rather strive for a *consensus* ***qua*** *machina*: digital approaches in the humanities presented us with a unique opportunity to make explicit and transparent, via annotation as well as formal modeling, many of the premises and assumptions that the traditional humanities had hitherto been able to avoid addressing. For the *machina* of the computer is on the one hand uncompromising in its insistence on explication, yet at the same time always willing to engage in a new, slightly differently parameterized iteration and recombination of computationally operationalized concepts and ideas. Humanists, I argued, should therefore begin to use this cognitive *machina* to continuously approximate, question and revisit knowledge rather than as a means to generate automatically finite results. As computing humanists we should always be mindful of the specificity of our object domain, and of our ultimate goal: to understand how humans construct, communicate and interpret meaning using symbolic artefacts and practices.

Twenty-five years later my current contribution is essentially an attempt to reiterate this point. In taking up Bradley's suggestion to "think about interpretation" I have outlined the trajectory from TACT to *Pliny* to CATMA, from text analysis to annotation and (visual as well as argumentative) synthesis–and then ultimately toward (theoretically infinite) re-interpretation. And so it might seem that we have

[27] Jan Christoph Meister, "Consensus ex machina? Consensus qua machina!"; in: Literary and Linguistic Computing. Journal of the Association for Literary and Linguistic Computing; Oxford/UK. 1995, 10:4, 263–270.

merely gone full circle. But the hermeneutic circle is a spiral: We progress by consciously revisiting where we've already been–be it in theorizing, in modeling, or in building tools.

Literature

Bauer, Matthias, and Angelika Zirker. 2013. 'Whipping Boys Explained: Literary Annotation and Digital Humanities'. In Literary Studies in the Digital Age, edited by Kenneth M. Price and Ray Siemens. Modern Language Association of America. https://doi.org/10.1632/lsda.2015.12.

Bögel, Thomas, Michael Gertz, Evelyn Gius, Janina Jacke, Jan Christoph Meister, Marco Petris, and Jannik Strötgen. 2015. 'Collaborative Text Annotation Meets Machine Learning: HeureCLÉA, a Digital Heuristic of Narrative'. DHCommons Journal. 2015. http://dhcommons.org/journal/issue-1/collaborative-text-annotation-meets-machine-learning-heurecl%C3%A9-digital-heuristic.

Boot, P. 2009. 'Mesotext. Digitised Emblems, Modelled Annotations and Humanities Scholarship'. Dissertation. 20 November 2009. http://dspace.library.uu.nl/handle/1874/36539.

Bradley, J. 2008. 'Thinking about Interpretation: Pliny and Scholarship in the Humanities'. Literary and Linguistic Computing 23 (3): 263–79. https://doi.org/10.1093/llc/fqn021.

Bradley, John. 1991. 'TACT Design'. CCH Working Papers, Vol. 1 (1991). 1991. http://projects.chass.utoronto.ca/chwp/bradley/.

———. 2003. 'Finding a Middle Ground between "Determinism" and "Aesthetic Indeterminacy": A Model for Text Analysis Tools'. Literary and Linguistic Computing 18 (2): 185–207. https://doi.org/10.1093/llc/18.2.185.

———. 2012. 'Towards a Richer Sense of Digital Annotation: Moving Beyond a "Media" Orientation of the Annotation of Digital Objects'. Digital Humanities Quarterly 006 (2).

Buzzetti, Dino. 2002. 'Digital Representation and the Text Model'. New Literary History 33: 61–88.

Caton, Paul. 2000. 'Markup's Current Imbalance'. Markup Lang. 3 (1): 1–13. https://doi.org/10.1162/109966201753537123.

Coombs, James H., Allen H. Renear, and Steven J. DeRose. 1987. 'Markup Systems and the Future of Scholarly Text Processing'. Communications of the ACM 30 (11): 933–47. https://doi.org/10.1145/32206.32209.

Da, Nan Z. 2019. 'The Digital Humanities Debacle. Computational Methods Repeatedly Come up Short'. The Chronicle of Higher Education, 27

March 2019. https://www.chronicle.com/article/The-Digital-Humanities-Debacle/245986.

Gius, Evelyn, and Janina Jacke. 2015. ‘Informatik und Hermeneutik. Zum Mehrwert interdisziplinärer Textanalyse’. http://www.zfdg.de/sb001_006.

Kleymann, Rabea. 2015. ‘Visualisation of Literary Narratives: How to Support Text Analysis with Visualisations? – Creating a Narratological Use Case’. 3DH (blog). 15 May 2015. http://threedh.net.

Kuczera, Andreas, Thorsten Wübbena, and Thomas Kollatz. 2019. ‘Die Modellierung Des Zweifels – Schlüsselideen und -Konzepte zur graphbasierten Modellierung von Unsicherheiten.’ Hg. von Andreas Kuczera / Thorsten Wübbena / Thomas Kollatz. Wolfenbüttel 2019. (= Zeitschrift Für Digitale Geisteswissenschaften - ZfdG / Sonderbände, 4)-. https://doi.org/10.17175/sb004_013.

Lancashire, Ian, et.al. n.d. ‘Lancashire, Ian, in Collaboration with John Bradley, Willard McCarty, Michael Stairs, and T. R. Wooldridge (1996). Using TACT with Electronic Texts: A Guide to Text-Analysis Computing Tools, Version 2.1 for MS-DOS and PC DOS. New York: Modern Language Association of America.’

Marcoux, Yves, Michael Sperberg-McQueen, and Claus Huitfeldt. n.d. ‘Modeling Overlapping Structures’. Accessed 12 September 2017. https://doi.org/10.4242/balisagevol10.marcoux01.

McCarty, Willard. 1996. ‘Finding Implicit Patterns in Ovid’s Metamorphoses with Tact’. Digital Studies/Le Champ Numérique 0 (2). https://doi.org/10.16995/dscn.227.

Meister, Jan Christoph. 2012. ‘Crowdsourcing “True Meaning”: A Collaborative Markup Approach to Textual Interpretation’. In Collaborative Research in the Digital Humanities, edited by Marilyn Deegan, 105–22. Ashgate Pub.

Moretti, Franco. 2000. ’Conjectures on World Literature’. In New Left Review, Vol.1. https://newleftreview.org/issues/II1/articles/franco-moretti-conjectures-on-world-literature.

Nyhan, Julianne. 2012. ‘Text Encoding and Scholarly Digital Editions’. In Digital Humanities in Practice, edited by Claire Warwick, Melissa M. Terras, and Julianne Nyhan, 117–38. London: Facet Publishing in association with UCL Centre for Digital Humanities.

Pierazzo, Elena. 2016. ‘Textual Scholarship and Text Encoding’. In A New Companion to Digital Humanities, 307–21. http://eu.wiley.com/WileyCDA/WileyTitle/productCd-1118680596.html.

Piez, Wendell. n.d. ‘Towards Hermeneutic Markup: An Architectural Outline’. Accessed 11 September 2017. http://dh2010.cch.kcl.ac.uk/academic-programme/abstracts/papers/html/ab-743.html.

Ramsay, Stephen, and Geoffrey Rockwell. 2012. ‘Developing Things: Notes toward an Epistemology of Building in the Digital Humanities’. In Debates in the Digital Humanities, edited by Matthew K. Gold, 75–84.

University of Minnesota Press. https://doi.org/10.5749/minnesota/9780816677948.003.0010.

Rapp, Andrea. 2017. 'Manuelle und automatische Annotation,'. In Fotis Jannidis/Hubertus Kohle/Malte Rehbein (eds.), Digital Humanities. Eine Einführung, Stuttgart 2017, p. 255f.

Renear, Allan. 2004. 'Text Encoding'. In Companion to Digital Humanities (Blackwell Companions to Literature and Culture), Hardcover. Blackwell Companions to Literature and Culture. Oxford: Blackwell Publishing Professional. http://www.digitalhumanities.org/companion/.

Renear, Allen. 2000. 'The Descriptive/Procedural Distinction Is Flawed'. Markup Languages: Theory and Practice 2 (4): 411–20. https://doi.org/10.1162/109966200750428722.

Schloen, David, and Sandra Schloen. 2014. 'Beyond Gutenberg: Transcending the Document Paradigm in Digital Humanities'. Digital Humanities Quarterly 008 (4).

Witt, Andreas. 2004. 'Multiple Hierarchies: New Aspects of an Old Solution'. Extreme Markup Languages 2004. 2004. http://conferences.idealliance.org/extreme/html/2004/Witt01/EML2004Witt01.html.

Wittig, Susan. 1977. 'The Computer and the Concept of Text'. Computers and the Humanities 11 (4): 211–15.

Zundert, Joris van. 2016. 'Screwmeneutics and Hermenumericals. The Computationality of Hermeneutics'. In A New Companion to Digital Humanities, edited by Susan Schreibman, Raymond George Siemens, and John Unsworth. Chichester, West Sussex, UK: John Wiley & Sons Inc. http://eu.wiley.com/WileyCDA/WileyTitle/productCd-1118680596.html.

Software

CATMA. Computer Assisted Textual Markup and Analysis - https://catma.de

Meister, Jan-Christoph, Petris, Marco, Gius, Evelyn, Jacke, Janina, Horstmann, Jan, & Bruck, Christian. (2018, June 18). CATMA (Version v5.2). Zenodo. http://doi.org/10.5281/zenodo.1470119

Pliny. A note manager - http://pliny.cch.kcl.ac.uk/

Bradley, John (2017)

TACT - http://projects.chass.utoronto.ca/tact/index.html (1996)

Appendix I: The CATMA 6 System Architecture

Marco Petris, Lead Developer CATMA

CATMA 6 (release date: October 2019; for the code and technical documentation see https://github.com/mpetris/catma-core) consists of two main components: a JAVA based servlet web application with a GitLab installation as its backend.

The central organizational unit in CATMA is a **Project**. A Project consists of the following resources: Documents, Annotation Collections and Tagsets. A Project also has a team of one or more users.

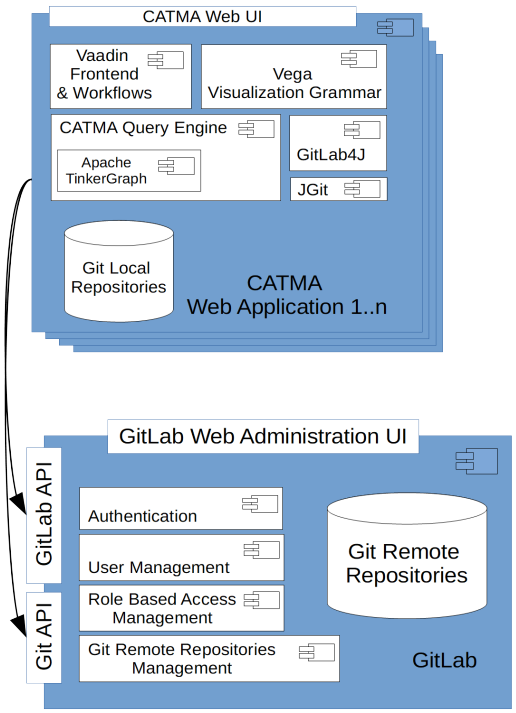

Fig. 1: CATMA 6 System Architecture

A **Document** is the primary object of investigation in CATMA. Each Document can have zero or more associated **Annotation Collections**. The Document cannot and must not be altered after having been uploaded to the system[28].

A **Tagset** is a set of zero or more **Tags**. Tags form a single rooted tree where a Tag has one or no parent and zero or more children.

A Tag has a name, a color and an author. Each Tag can have zero or more user defined **Properties** each with a name and a list of zero or more values to be proposed upon application.

An **Annotation Collection** is a collection of Annotations associated with a Document. Each Annotation Collection belongs to exactly one Document and can contain zero or more Annotations. An Annotation is always applied to one or more possibly discontinuous segments of text with character start and character end offsets. It is typed by its Tag and gets zero or more user defined Properties from its Tag. These Properties then receive an annotation-specific configuration of values. The values can

[28] This might change in the future as we plan to add an algorithm that can adjust associated Annotations based upon a versioned journal of (minor!) changes applied to the Document.

be from the list of values proposed by the Tag but they are not limited to those values. An Annotation has an author and a timestamp.

Each Project has an **owner**. Other than the 'owner' role there are the four other GitLab roles (maintainer, developer, reporter and guest) that drive the permissions on the Project and its resources.

Each resource is managed as a **git repository**. In order to manage the versions of the participating resources on the Project level, there is a container git repository that contains all resource git repositories as git submodules.

All Annotations and meta data are stored in JSON format. The Annotations are modeled according to the Web Annotation Data Model.

The GitLab backend provides user management and role-based access control. The GitLab equivalent of a Project is a Group. The Group is the namespace of all resource git repositories. This allows the reuse of resources in different Projects by forking the git repositories into a new Group, i. e. into a new namespace with a fresh setting of users, roles and permissions. GitLab enables collaborative work on reusable shared resources by ensuring role-based access and by exchanging resources. The heavy work instead is done by the CATMA Web Application.

The CATMA Web Application is able to scale horizontally and in principle also to run as a local desktop client. It talks to the GitLab backend via the GitLab REST API and the GitLab Git API for authentication and collaboration. The Projects are managed via JGit as local git projects, cloned from the GitLab remotes. On opening a Project it gets loaded into an in-memory TinkerGraph graph database that drives the CATMA Query Engine and the Vaadin UI's data models and workflows. The graph database provides a balance between fast indexing and fast retrieval and goes beyond the capabilities of token based indexers such as Lucene/Elastic Search.

GitLab and the CATMA Web Application can run on a single machine or on separate machines.