Felix Lange

An Online Field Study on Scholarly Journal Annotations

Empirical Evidence and Implications for Software Design in the Digital Humanities

Abstract: Even though there is an abundance of web-based annotation tools that allow users to share their data across the internet, little is known about how these tools are actually used in the daily work routines of scholars in the Humanities. This chapter presents an empirical study on public inline annotations by publishers, article authors and readers in a scholarly open-access journal. The findings of this study are combined with a meta-analysis of the existing empirical literature on marginal annotations in the Humanities and scholars' willingness to share them. The most important conclusion that can be drawn from the empirical data is that the publication of annotations is not a feature that needs to be offered by all types of scholarly annotation software packages.

Keywords: Collaboration, Evaluation, Classifying, Commenting, Form, Tool, Digital Humanities

1 Introduction

In his contribution to this volume, Willard McCarty describes his personal way of writing, storing and processing notes on scholarly texts. An early step in this workflow is to „record ideas, keywords and references to other sources I want to come back to later for more detailed note-making" on separate paper slips (McCarty 2020, 276 ff.). Two aspects of this description, the temporary, transient nature of preliminary comments and the working context in which they are made, can also be taken as distinctive features of individual scholarly annotations written in the margins of texts, henceforth marginalia (cf. Bold and Wagstaff 2017). It has been shown that these textual notes, consisting of a couple of sentences or even only some symbols, have been an important element of academic reading throughout the ages (Agosti et al. 2007; Blair 2004). Less clear is the relationship between this text genre and the scholarly practices of communication and data sharing. In the Digital Humanities (DH), which set out to foster collaboration and information

Felix Lange, Max-Planck-Institut für Wissenschaftsgeschichte

https://doi.org/10.1515/9783110689112-011

sharing, there have been numerous initiatives to provide the means to share or publish this type of data. Connecting primary sources, scholarly literature and annotations on these texts could result in a "Scholarly Web" (Perkel 2015) that crosses the boundaries of disciplines and links once isolated digital collections (Lordick 2015, 2). The individual work of text analysis might be opened up to larger audiences even in the early stages of research (Becker et al. 2016, 10). These ideas are taken to the extreme by Hemminger and TerMaat (2014, 2278): "One can now foresee a time when every scholar's thoughts about a particular article are electronically captured and displayed to other scholars." This vision has already received considerable technological support: The "Web Annotation Data Model" (henceforth WADM) issued by the W3C[1] provides interoperability across software and collections (Hunter et al. 2010), and a number of DH tools offer annotation-sharing capabilities (Müller-Birn et al. 2015, Grassi et al. 2013[2]). At present, though, it is unclear if a Scholarly Web of annotations will come into existence. For other types of annotations such as linguistic tags, established infrastructures for the publication of annotated data exist and are in constant use,[3] but so far no comparable platform has attracted a large number of textual free-form annotations. It seems safe to say that the web-scale publication[4] of marginalia has not become a regular feature of scholarly work in the Humanities so far. The question arises as to whether it should be a feature of annotation environments at all. After all, there have been a number of reports of a general mismatch between user needs and software designs in the Digital Humanities (Juola 2008, 75; Pape et al. 2012, 3 f.). This potential mismatch could reflect a general problem with annotation tools. An annotation feature was devised for one of the first graphical web browsers in the early 1990s (Carpenter 2013), and later Adriano and Ricarte (2012) were able to list eighty different systems in a comparative study of general-purpose annotation software tools. In the (now defunct) DH tool directory DiRT, "Annotation" was among the three functions that were referred to most often (Borek et al. 2016, Par. 9). But web annotation was not included as a feature of later browser generations, whose present-day descendants still do not conform to the WADM (Shaikh-Lesko 2019), and so far no *killer application* (cf. Juola 2008) has been presented for web-based digital annotations. It seems clear that more research is needed on the real potential for annotation-sharing within and outside academia. In light of these problems, this chapter follows Antonijević Ubois' (2016) maxim of "Developing Research Tools via Voices from the

1 https://www.w3.org/TR/annotation-model/ (20.01.2020)

2 See also Hemminger's (2009) conceptual study.

3 See e.g. CLARIN: https://www.clarin.eu/resource-families (20.01.2020)

4 Public annotations as understood here include annotations on targets with copyright restrictions.

Field", gathering empirical evidence on scholarly annotation practices in order to derive ideas for software design. It presents the findings of a study on public inline annotations authored by publishers, article authors and readers in the open-access journal eLife.[5] The findings of this study are combined with a meta-analysis of the existing empirical literature on marginalia in the Humanities and scholars' willingness to share them.

The remainder of this chapter is structured as follows. Section 2 offers an overview and a categorization of annotation systems in the Digital Humanities. Section 3 reviews the existing literature on annotations, and Section 4 presents new data on public scholarly annotations. Sections 5 and 6 set out the findings and derive recommendations for software design.

2 Annotations in the Digital Humanities: Concepts and Systems

At least since the 1990s, shared digital annotation environments have been an active field of study, both within the Digital Humanities and in Computer and Information Sciences (e.g. Ovsiannikov et al. 1999) in general. However, the types of annotations discussed range from marginalia written for private use to digital editions and linguistic markup in text corpora (Hunter 2009, 1). Annotations of the latter two types constitute research findings that are published together with their respective annotation targets (i.e. the objects that annotations are attached to). It is clear that publishing these annotations is normally useful or even necessary. Therefore, annotations need to be categorized to distinguish between different degrees of *a priori* suitability for publication. However, there is no consensus in the literature on a useful typology of digital annotations. In his influential work on "Scholarly Primitives", Unsworth (2000, 1) counted the practice of annotating among the "basic functions common to scholarly activity across disciplines". In a similar manner, the "Taxonomy of Digital Research Activities in the Humanities" project (TaDiRAH, Borek et al. 2016), which draws on Unsworth's work, does not subcategorize "Annotating" any further, but subsumes the practices of "adding, e.g., comments, metadata or keywords"[6] under the entry. In this taxonomy, annotating is a subtype of "Enrichment", as it makes information inherent to the annotation target explicit. Annotating is explicitly contrasted with "Commenting"

5 https://elifesciences.org/ (20.01.2020)

6 http://tadirah.dariah.eu/vocab/?tema=22 (20.01.2020)

Tab. 1: Annotation dimensions

Category	Annotation Author/Reader Scope	Target Type	Annotation Target Granularity	Metadata Depth
Values	1: Individual 2: Collaboratory 3: (Scholary) Public	1: Research Literature 2: Primary Source	1: Publication/File 2: Part of Publication/File	1: Technical/Application-Specific 2: Ad-hoc Semantics 3: Std.-Conformant Metadata (WADM) 4: LOD (Target, Body)

(a subtype of "Dissemination"), an activity that "serves to express some opinion, to add contextual information, or to engage in communication or collaboration.[7]" If these definitions are used to inform software design one-to-one, annotating has to be modeled as one function, and commenting as another. But the distinction between "contextual information" and information which is "inherent" to the annotation target is too subtle for that purpose. Furthermore, there are conceptual doubts about whether a clearcut distinction between the two activities is empirically adequate: Walkowski (2016b, 9 f.) notes that in practice, annotating is most often part of other research activities. And with respect to annotating as a "Primitive", Unsworth later considered the possibility that some of the initial categories might have to be further subcategorized (Unsworth and Tupman 2016, 232). Indeed, it can be shown that a more fine-grained subcategorization of annotating practices is helpful in constructing suitable use cases and, accordingly, functional requirements for software design. The factors presented in Table 1, which have in part been derived from Hunter's comprehensive typology (Hunter 2009, 4–14), form the basis for a tentative subcategorization of annotations and the software systems with which they can be produced.

This choice of features is motivated as follows. Author scope and reader scope[8] indicate whether annotation authoring must be a function presented to all annotation *readers*, or if writing and reading functions can be facilitated by different software modules and interfaces. "Collaboratories" are defined in Cerf et al. (1993, 7 f., cf. Agosti et al. 2004) as networked infrastructures enabling scientific collaboration. They differ from solitary working contexts in that they require networked software for shared annotations. In contrast to web-scale annotations, however, sharing is not necessarily required between applications, text collections and tasks

7 http://tadirah.dariah.eu/vocab/?tema=44 (20.01.2020)

8 The term was coined by Agosti et al. (2004, 245)

or research projects (a factor that limits the interoperability requirements). The target type is important because it helps to distinguish between typical source enrichment annotations (e.g. scholarly editions) and annotations on scholarly literature (Agosti et al. 2004). The annotation target granularity differentiates between regular commenting systems, which are a common feature of online journals,[9] and inline annotation systems, which are offered to journal readers less often. Finally, the metadata depth determines what kind of information is shared across application types, application instances and document collections. For example, typed links are a way to express the semantic relation between the annotation content and its target (Agosti 2005, 95). The WADM defines 13 types for this purpose, including "questioning" and "replying".[10] In contrast, an *ad-hoc semantics* does not conform to a standardized data model and is most likely defined at project level. Technical metadata comprises whatever metadata is needed by the application to store and display annotations. Linked Open Data (LOD[11]) is the metadata concept for Semantic Web applications.

For all dimensions except the target type, values are ordered according to their (probable) implications: an annotation system that allows a single paragraph in a text to be targeted probably also allows an annotation to be made on the text as a whole. Software with LOD metadata produces technical metadata as well.

Most annotation systems presented in DH contexts provide some sort of sharing function. The stand-alone marginalia tool Pliny is a rare exception in this respect. It is explicitly designed to foster "not so much social scholarly interaction but personal research" (Bradley 2008, par. 9) and is able to annotate local PDF files and web-based files (primary sources and articles alike). A much greater number of approaches focus on collaboratories. Bradley and Vetch (2007) discuss a tool that produces TEI-conformant metadata for the Online Chopin Variorum Edition (OCVE).[12] Bauer and Zirker (2015) report on a collaborative tagging task in an undergraduate university seminar in literature studies, where inline tags are defined at project level. De la Flor et al. (2010) test a prototypical Virtual Research Environment for the recognition of linguistic structures on historical inscriptions with a group of classicists. Gius and Jacke (2017) report on a combination of automated and manual annotations of narratological structures in literary texts using the CATMA tool, which allows its TEI-XML-encoded metadata to be exported.[13] Such a hybrid approach is also used by Müller-Birn et al. (2005) in the design of

9 See e.g. DH Quarterly (http://digitalhumanities.org/dhq/about/about.html [20.01.2020])
10 https://www.w3.org/TR/annotation-model/#motivation-and-purpose (20.01.2020)
11 https://en.wikipedia.org/wiki/Linked_data#Linked_open_data (20.01.2020)
12 http://www.chopinonline.ac.uk/ocve/ (20.01.2020)
13 https://catma.de/documentation/tei-export-format/ (20.01.2020)

the semantic annotation suite neonion. E-Carrel, a VRE for collaborative work with primary sources and literature, includes a tagging function and produces RDF-XML (Jones et al. 2010).

Hypothesis[14] is a public annotation tool for authors and readers. It is designed for a wide range of uses, from open-access peer review to the inline annotation of scholarly articles by readers (Staines 2019). Annotations are fully WADM-conformant, but not all metadata specified by the standard are actually written by the system. For example, there is no way to write the WADM motivation tag. Instead, Hypothesis uses a comparably slim semantic model which is largely transparent to the user. In addition to web-based content, it can also be used to annotate local PDF files. Local annotations can be shared between different users across identical file copies. A recent application of this approach in the academic domain that goes beyond collaborative text production and review is the initiative "Annotation for Transparent Inquiry" (Karcher and Weber 2019). In the context of ATI, the Hypothesis function is integrated into an open-access journal and thus allows authors to publish relevant data along with the primary publication. ATI is thus an example of a concept where the scope differs for annotation readers and authors. It resembles the concept of "Living Reviews" developed by Bernhard Schutz and others for "Living Reviews in Relativity", a review journal in the field of physics[15] (Wheary et al. 1998).

The web-based annotation tool Pundit (Grassi et al. 2013) is designed for semantic annotations which contribute to a machine-readable Web of Data (Becker 2016, 14) for public readers and authors alike. Pundit gives the user direct access to semantic web resources and writes LOD-conformant RDF triples. Since the entire data model is based on a Semantic Web Architecture, it is assumed that all annotation targets are also web-based resources. The system is suitable for both primary and scholarly literature, but reports on tests and deployments focus on the enrichment and interpretation of primary sources (e.g. Di Donato et al. 2013). In a laboratory test conducted during the development phase, users presented with sample tasks noted problems with the complexity of the interface, but generally approved of the idea of enriching historical sources with Linked Open Data (Hennicke et al. 2015, 59 ff.). Annotations for authors of scholarly literature are mostly facilitated by page-level commenting systems with shallow metadata (e.g. in the Digital Humanities Quarterly[16]).

14 https://web.hypothes.is/ (20.01.2020)

15 https://www.springer.com/journal/41114 (20.01.2020)

16 See http://www.digitalhumanities.org/dhq/vol/11/4/000340/000340.html (20.01.2020)

Tab. 2: Observed annotation types[17]

Category	Annotation Author/Reader Scope	Target Type	Annotation Target Granularity	Metadata Depth[18]	Supporting Software (Examples)	Public Platforms (Examples)
Private Marginalia	1/1	1,2	1,2	1	Pliny	
Collaboratory/ Edition	2/3	1,2	1,2	4	TEICHI, CATMA, E-Carrel	OCVE
Publisher/Author Journal	1,2/3	1	1,2	3	Hypothesis	ATI, Living, Review
Public Journal	3/3	1	1,2	3	Hypothesis	DHQ
Public Semantic Web	3/3	1,2	2	4	Pundit, neonion	

The tools and infrastructures discussed so far are categorized according to a tentative typology in Table 2. This overview points to a negative correlation between the metadata depth and the scope of annotations: annotations with wider *author* scope tend to have a shallower metadata depth. In fact, we have not found a platform that includes a significant number of public reader-generated semantic annotations. Instead, the published systems and projects mostly focus on collaboratories with an asymmetric author/reader scope. These annotations are in general authored exclusively by members of the collaboratory and may be published as research findings afterwards. To the best of our knowledge, annotations of this type are the only major source of web-scale scholarly semantic annotations in the Humanities.

3 Literature Review

3.1 Forms and Functions of Annotations

In an analysis of handwritten notes in library textbooks, Marshall (1997) shows that the production of marginalia is ubiquitous in academic reading. The move towards digital documents has not changed that, as Liu (2005) shows in a direct

17 The reference numbers refer to values defined in Table 1 on page 224.

18 The values in this category are meant to represent the maximum values which have been observed. For example, not every system in the "Collaboratory" category actually writes Linked Open Data.

comparison of digital and non-digital reader annotations (see also Qayyum 2008). In the large-scale DARIAH-DiMPO survey, which covered the use of digital tools and methods among Humanities scholars, 65.5% of the respondents (N=2,176) stated that they annotated, curated and enriched their objects of research (Dallas et al. 2017).

The function of reader annotations is not necessarily to contribute new information (Marshall 1997); they also add visual structure to texts and function as procedural signals in the context of a specific reading task. These functions correspond with features such as transience versus permanence and semantic explicitness versus implicitness (Marshall 1997, 6; Marshall 2000, 98).

A typical work context for marginalia in the Humanities is writing a publication which references other works of literature and primary sources. O'Hara et al. (2002) describe this type of use for professionals in different domains. From the perspective of the "Distributed Cognition" framework developed in cognitive writing research (Klein and Leacock 2012), they observe that the process or writing an article is structured by "intermediate texts", the production of which serves the purpose of cognitively "offloading" the *reader-writer*'s working memory. Because time and attention are critical factors in the process of annotating, it has to be "minimally disruptive" (O'Hara et al. 2002, 289), and the resulting texts are mostly implicit, transient and informal, in Marshall's terminology. They are only meant to supply information to their author in a specific work context and for a relatively short period of time. Accordingly, formal and semantic features of private annotations are mostly idiosyncratic (cf. Blustein et al. 2011). This fact makes it probable that many, if not all of the marginalia a researcher produces during the compilation of a research article are not suitable for publishing. The transient and informal nature of annotations fits McCarty's description cited in the introduction. In contrast, the findings presented here do not seem to be compatible with the user experience of semantic annotation tools with rich metadata structures and complex user interfaces.

3.2 Annotation Practices and Scholarly Workflows

The question of to what end scholars write annotations can only be answered satisfactorily against the backdrop of a scholarly workflow model. Empirical workflow studies in the Humanities have been conducted since at least the 1970s with the aim of investigating information behavior patterns and corresponding information needs. Of these studies, Uva's (1977) work is arguably the first to develop and test a sequential workflow model for humanists. Drawing on studies from the Social Sciences, he proposes the workflow phases *Problem Selection, Detailed*

Planning of Data Collection, Data Collection, Analyzing and Interpreting Data and *Presenting Findings*. However, the study concludes that a deterministic sequential process model is not realistic (Uva 1997, 16). While later works have postulated similar workflow *elements*, the concept of a deterministic sequential *ordering* of these steps has not been pursued further (Stone 1980, 1982; Case 1991). While Chu (1999, 259) observed the writing and reading of annotations in almost all research phases (cf. Antonijević and Cahoy 2014; Palmer et al. 2009, 188 f.), it does not seem possible to connect different types of annotations with specific degrees of formality and explicitness to different sequential workflow stages. Neither is there anything resembling a generalizable workflow pattern among humanists. On the contrary, working habits seem to be quite individual and idiosyncratic. For software design, this finding means that it is not feasible to offer different workflow applications with different annotation features for different workflow-related types of annotations, if information fragmentation is to be avoided (cf. Belanger 2010).

3.3 Scholarly Personal Information Management

During the research and writing processes, notes and annotations are generally kept in personal data files, becoming part of a scholar's *Personal Information Management (PIM)*. According to Jones' (2007) definition, PIM encompasses the management of all task-related information activities, including storing, retrieval and curation. In a study of the PIM practices of historians, Case (1991) shows how information that is found in archives and the literature is stored in complex notebooks and on index cards for later retrieval in different contexts (e.g. new research projects). Similarly to scholarly workflows, PIM habits are widely believed to be guided by individual habits.

Antonijević and Cahoy (2014) analyze scholarly PIM habits during all observed workflow phases, from information seeking to information archiving. They postulate two general relevant factors relevant to the way scholars build their *Personal Information Collections (PICs)*. First, interviewees express skepticism about remote storage in academic research data repositories and prefer to use commercial cloud services which allow the user to keep copies on their local hard drives. Second, scholars do not use integrated asset management systems with stand-off metadata, but apply their own principles when ordering their information. The data is primarily stored as PDF files (cf. Cushing and Dumbleton 2017). Even bibliographical reference management systems cannot be considered a standard tool among scholars (cf. Ollé and Borrego 2010, 225 ff.). These findings do not only pertain to older generations: similar observations have been made in a recent study on PhD students (Cushing and Dumbleton 2017, 45 f., see also Given and Willson 2018,

815 f.). In general, PIM techniques are habits that are developed over a long period of time and do not change quickly (Barreau 2008). Since annotation practices are to certain extent PIM practices, that explains why users have been slow to adopt new annotation systems. Furthermore, the fact that users seem to download texts rather than referencing online texts speaks against web-based annotation systems which cannot handle locally stored PDF files.

3.4 Collaborative Working Practices

The ideal case for using web-based annotation systems is collaborative work. Accordingly, in order to assess how useful this centralized architecture is for the scholarly community at large, it is important to assess the role collaborative work plays in modern humanists' working lives.

A bird's-eye view of collaboration in the Humanities is offered by the DiMPO survey. This states that 71% of 2,135 respondents collaborate "often or very often". However, this finding stands in contrast to small-scale workplace and interview studies, which generally assume solitary research to be the norm. A case in point is Stone's (1980, 17 ff.) study. Later works (e.g. Cronin 2003) conclude that scholars contact peers to gather information about research topics and organizational matters but still mostly carry out the actual research on their own.

An often-used quantificational metric for collaboration is co-authorship. This approach has long been criticized, because, for example, colleagues might simply be included on author lists for social and strategical reasons (Katz and Martin 1997). On the other hand, since research in the Humanities does not in general necessitate the use of complex scientific instruments or a high degree of division of labor, it can be argued that co-authorship does at least suggest that the authors are likely to have worked on a research subject together (Burroughs 2017, 511). Judging from publication statistics alone, there can be no doubt that joint authorship is much less common in the Humanities than in STEM fields and in the Social Sciences. This has been shown for European countries (Ossenblok, et al. 2014; Piro, et al. 2013, 309 f.) as well as for China (Ma et al. 2014) and for American universities (Burroughs 2017). If anything, recent bibliometric studies support the thesis that Humanities research still consists mainly of solitary work. Therefore, Bradley's (2008) approach of providing an annotation tool for individual usage seems to match user needs.

3.5 Software Usage

All the annotation environments discussed in Section 2 require users to stop using their stand-alone document viewer. However, the observations regarding scholarly PIM practices have already pointed to the persistence of individual working habits and conventional tool usage. Research findings about scholarly workflows have shown, in fact, that annotations are written and read at virtually any point of the research and publication process; they must be supported by whatever document software the scholar routinely uses, and the annotation must not result in a significant additional workload. In other words, annotation features must be readily available when reading documents. The question is: What software do researchers routinely use for reading and processing research documents? The DiMPO survey states that when it comes to storing research findings, standard office applications are by far the most popular software. Concerning the usage of "digital tools and methods", "annotating, enriching and curating" are the least frequently mentioned activities. It is not entirely clear what kind of annotations are referred to in the survey. However, from the survey findings, the conclusion can be drawn that the respondents do not use DH-specific annotation tools, but prefer annotation features embedded in general-purpose software packages. These findings are in line with a smaller usage study among German scholars (Stiller et al. 2015, 29).

In a medium-scale mail survey (N=123), Müller-Birn et al. (2016, 5 ff.) collect much more detailed usage data. Presented with their categorization in "standard" and "extended" software use use, roughly two thirds of the respondents identify themselves as within the latter category. From information about the specific types of software used, user group clusters are generated. One statistically significant group uses annotation features or software. This group, however, also stands out for using standard office software more often than other groups. There is thus no evidence that this group uses *DH-specific* annotation functions. Indeed, in an earlier interview study conducted by Müller-Birn et al. (2005, 224), all respondents state that they use standard office software for annotating text. This finding is confirmed by Given's recent study, in which a tool inventory compiled from interviews of 20 faculty staff and graduate students in the Humanities reveals the use of DH-specific analysis tools and standard software, but does not include any (DH or general-purpose) annotation software (Given and Willson[19] 2018, 810 ff.).

An anthropological study by Antonijević (2015, 38 ff.) complements these quantitative studies. In qualitative interviews, she shows that the use of standard pro-

19 The study shows, however, the widespread use of XML markup tools.

ductivity tool is prevalent in all phases of the research workflow among Humanist scholars.

In combination, the findings presented here produce a complex picture: scholars use advanced software packages for research-specific purposes, but standard tasks, which surely include document reading and note-making, appear to be handled with standard tools. There is no explicit proof whatsoever that the use of DH-specific annotation tools facilitating annotation sharing is common among scholars.

3.6 Direct empirical Evidence for Annotation Sharing among Humanists

The studies presented so far have shown potential problems for the adoption of existing DH annotation environments, related not only to individual working habits but also to the observed features of marginal annotations. The findings cast some doubt on the idea that users might really want to publish their private annotations.

There are some studies that examine informal textual annotations in online environments. Kopak and Chiang (2007) and Chiang (2010) analyze public online annotations in a laboratory setting with a small number of participants. Among the findings of these studies is strong approval for online annotations as a means of enhancing productivity and as a support for active reading, as well as approval for functions that allow the linking of external content. In a laboratory setting, Hemminger and TerMaat (2014) elicit opinions about sharing annotations. Respondents argue that they would prefer to subject their individual annotations to extensive revision before publishing them, and that they would prefer annotations not to be made anonymously. In a similar setup, Marshall and Brush (2004) asked test participants to choose which of their personal annotations to share, and to decide how to revise them beforehand. In both cases, the revisions involve making the content more explicit and "intelligible" for potential readers, which points to the implicit and idiosyncratic semantics of marginalia observed by Marshall in her earlier studies. Respondents in these studies voice a cautiously positive opinion about annotation-sharing in general, however. This finding mirrors Walkowski's (2016a, 6 f.) small-scale study.

A noteworthy feature of Hemminger and TerMaat's (2014) and Marshall and Brush's (2004) studies is that they describe a two-step process that involves first making a note for personal purposes and, in a later step, publishing it. There is no suggestion that public annotations could have been written as such, and therefore represent an annotation type *sui generis*.

All four studies have the limitation that they elicit data in purely hypothetical usage settings, using test sentences and non-production software. The resulting data is produced mostly through introspection and not collected by observation of daily work practices. It is difficult to evaluate how realistic and reliable the respondents' statements about their hypothetical annotations are. For these reasons, laboratory studies of this kind should be complemented with actual field data to derive empirically grounded software design recommendations.

4 Empirical Study: Public Annotations on eLifeSciences.org

As the preceding sections have shown, annotation-sharing is not yet an established communication channel in the Humanities. There are examples of the successful introduction of both public commenting and annotation systems in other academic disciplines such as the Life Sciences. To the best of our knowledge, however, no study has been carried out systematically examining public scholarly inline annotations in real-life settings, and we were not able to find a data set from a Humanities context large enough to be representative of the field. This study aims to narrow this gap by examining how readers of the open-access Life Sciences magazine eLife use the Hypothesis annotation feature included in the site. With several thousand public annotations (Shaikh-Lesko 2019) made accessible via the Hypothesis API since its introduction in early 2018, it provides a valuable dataset to investigate patterns of use for scholarly annotation-sharing applications.

4.1 Research Design

eLife is an open-access journal that publishes articles classified into 18 Life Sciences research fields such as "Evolutionary Biology" or the "Physics of Living Systems", as well as opinion pieces, announcements (such as calls for papers) and interviews categorized under the heading "Magazine". For every article in the "Research" category, usage metrics (views, downloads, citations) are supplied. The Hypothesis plug-in is available for both categories and for every scientific sub-discipline covered in the journal. Annotation targets can be parts of the text or the text as a whole. Hypothesis replaced a conventional article-commenting system

in January 2018.[20] Earlier comments were imported into Hypothesis and are now presented as page-level annotations (so-called "Page Notes"[21]). The annotation function is not only presented to readers of published articles, but is also used in the review process. Many statements made and discussions held during the review are publicly available as annotations on the website. Furthermore, the publisher uses the feature for post-publication reviews and announcements concerning the article (updated versions, etc.).[22] Annotation authors register at Hypothesis by supplying an ORCID (Open Researcher and Contributor ID), which is, however, not always published on the site.

The data for this analysis were obtained from two sources. First, the Hypothesis API[23] was queried for content relating to elifesciences.org. Second, all articles were crawled that were published between October 15, 2012 and October 31, 2019, listed in searches for the general categories "Magazine" and "Research" and assigned a DOI on the webpage.[24] The resulting analysis sample consists of all the Hypothesis data for each annotation and bibliographical metadata for each published article.

The principal question of this analysis is how the scientific audience of a publication interacts with its authors and publishers through the use of annotations. Therefore, the annotations were assigned to different annotation author roles: *article author*, *publisher/reviewer* and *reader*. These roles correspond to the two annotation types *publisher/author journal* and *public journal* as defined in Section 2. Since not every annotation author account is traceable to a particular person via an ORCID or other unique global identifier, authorship was attributed heuristically according to the following rules:

Article author (Sufficient conditions)

1) Annotation author's first and last names are identical to one article author's first and last names.

20 https://elifesciences.org/for-the-press/81d42f7d/elife-enhances-open-annotation-with-hypothesis-to-promote-scientific-discussion-online (20.01.2020)

21 https://elifesciences.org/labs/3f85f8c2/enabling-scientific-discussion-on-elife-with-hypothesis (20.01.2020), par. 11 ff.

22 Shortly after its introduction in 2018, eLife published an early overview of how the annotation tool was being used by readers, authors and publishers: https://elifesciences.org/inside-elife/7affe390/elife-latest-how-readers-are-using-annotations-on-our-website (20.01.2020)

23 https://hypothes.is/api/, https://h.readthedocs.io/en/latest/api-reference/ (20.01.2020)

24 Search-URL https://elifesciences.org/search (20.01.2020). DOIs were extracted with an XPath expression.

2) Annotation author refers to him- or herself as an/the author in the annotation (e.g. "our publication"[25]).
3) Annotation author's account is connected to a publicly available ORCID account and the account's name and publication list match the article.
4) (Necessary condition) The annotation does not express a question about the article.

Publisher/Reviewer (Sufficient conditions)

1) User account is "eLife Journal".
2) Annotation author refers to him- or herself as a member of the journal staff in the annotation text.[26]
3) Annotation text is an announcement of corrections, new versions or other actions on behalf of the publisher, and the annotation is not an author annotation.

Reader (Necessary conditions)

1) The annotation is by neither an author nor a publisher/reviewer, according to the conditions given above.
2) The annotation is not bot-generated.[27]

This categorization aims to characterize scholarly conversation about an article with respect to which role each annotation author takes on in the discussion. Therefore, the professional affiliation of each author is of secondary importance. For example, the article affiliation of the annotation can be categorized as external if its author is involved in a discussion on the article, rather than referring to the review and publication process, even if the annotation author is a member of the editorial board. For the Article author category, rule 1 technically leaves some ambiguity with respect to namesakes. A reader bearing the same name as one of the authors could be wrongly categorized as the article author. In the analysis sample, however, no cases surfaced where an annotation author with a namesake account voiced a question or critique of the article. Uncertain cases concerning namesake accounts are thus reduced to potential cases where annotations simply add further, publicly available information to an article and where no specific author attribution was possible by applying the above rules. The same applies to

25 https://hyp.is/Qq_ACgXREeiBz9cPIcOU6w/elifesciences.org/articles/19088 (20.01.2020)
26 In some cases the account can also be validated via the ORCID.
27 A bot generated automated annotations, identifiable by a specific account name.

Publisher/Reviewer accounts. Therefore, the uncertainty inherent in this heuristic method of authorship attribution appears to be negligible.

Another feature of the data sources is harder to control for. The fact that annotations can be made on older articles without any time limit introduces a bias into the sample. It is to be expected that the newer an article, the larger the set of annotations that have been and will be added to this article after the data collection has been conducted. It can be assumed that this bias results in lower numbers for reader annotations in relation to annotations made by publishers and reviewers, as well as lower numbers for newer annotations in relation to page-level comment annotations made before Hypothesis was added to eLife in January 2018.[28] Neither relationship is of concern to the present analysis. Nonetheless, to alleviate this bias, while annotations were collected up to and including the publishing date of November 30, 2019, articles were collected up to and including the publishing date of October 31, 2019.[29] Increasing the time distance between the article and annotation publication even further would have diminished the already small sample too much and thereby obviated a quantitative analysis.

4.2 Quantitative Analysis

With the above filters in place, the sample consists of 7,481 articles, of which 6,669 are classified as research articles and 812 belong to the "Magazine" section. The call to the Hypothesis API retrieved 2,000 human-generated annotations linked to articles within the sample. Other annotations were either linked to articles not included in the sample, were not classified as reader annotations or were generated. Of the 891 Reader annotations, 375 belong to the magazine section and 516 are research annotations. Filtering out imported comments produces the core sample of 108 reader research annotations made with Hypothesis. This sample is compared with reader magazine annotations from the same time period and with imported comment annotations. Some baseline counts are given in Table 3 on the next page.

Given this relatively low number, the question seems justified of whether the annotation feature, and the inline annotation feature in particular, are used at all by a relevant percentage of readers. The 108 research reader annotations are distributed among 66 articles and are written by 66 authors. This means that usage

28 The official launch of the feature was in January 2018. But in order to simplify the differentiation between annotations in the old and in the new environment, annotations from that month were cut out of the two samples.

29 Of course, annotations published in November 2019 were only included in the sample if they belong to articles published before November 2019.

Tab. 3: Quantitative results (Reader annotations)

	Total	inline (pct.)	No. of Annotated Articles	No. of Annotated Authors	80th Percentile Text Length (No. of Words)	Ext. Ref. (pct.)
Research Annotations (Hypothesis)	108	95.4	66	66	68	35.2
Magazine Annotations (Hypothesis)	55	92.7	17	31	67	54.5
Comments Research (Before Hypothesis)	391	X	257	352	157	

of the annotation feature is at least not due to an insignificantly small number of early adopters. Of course, however, Hypothesis reader annotations on research articles[30] (mean of 0.02) are much less common than citations (mean of 4.5). The number of comment annotations on pages in the older sample is of course much higher (mean of 0.06), as is the mean number of citations of articles from that period (mean of 30.3).

During the principal observation period of 2018–2019, readers could publish both inline annotations and page comments. The findings in Table 3 make it clear that users preferred the inline feature over Page Notes. Unless the annotation targets were chosen randomly within the annotated text, it is obvious that in most cases users made a conscious decision to direct their argument at a certain portion of the text.

The text length of annotations can be used as an albeit coarse quantificational measure for their content: very short annotations in the sample tend to be relatively unspecific remarks ("Helpful"[31]), whereas the data also contains many annotations that are in fact lengthy reviews of the annotated articles and do not target a specific part or aspect of its content. As a measure less sensitive to outliers than the mean value, the 80th percentile for the text length (68, mean of 90.3) in the research article sample shows that most annotations are relatively short texts. Annotations with less than 10 words appear 26 times in the sample. There is also a strong correlation between the text length and the structural type of annotation: the 80th percentile is 574 (mean of 528.8) for page comments in contrast to 67 (mean of 69) for inline annotations. The numbers can be seen as evidence that inline annotations constitute another text type compared with the page comments in

30 Articles published during the Hypothesis deployment period 2018–2019. Of course, during that time also older articles could be annotated with the new tool.

31 https://hyp.is/XfTuOIrLEeiB7PtOmyZchA/elifesciences.org/articles/22784 (20.01.2020)

the sample. Comments in the older sample are also longer (mean of 182.7) than inline annotations, but shorter than the page comments in the newer sample. It seems that users employed the commenting function for two types of annotations which are distributed between page comments and inline annotations in the newer dataset (i.e. after the introduction of Hypothesis). One possible conjecture which can be derived from these findings is that most annotations are thematically related to a specific part of the annotated text.

The reply feature in annotation systems facilitates scholarly discussions rather than isolated remarks on the underlying articles. In the analysis sample, it is used in a significant number of cases. Replies and answers within the sample of human-generated annotations of all author affiliations (N=2,000) together form 112 reply chains. However, the dynamics of these conversations are quite limited. In the whole research-article sample (1517 annotations, 86 chains), there are only 15 reply chains of a length greater than 2,[32] and only 10 reader annotations are replies to other reader annotations. This is due to the fact that replies to annotations are mostly author/publisher answers to reader questions.

In earlier empirical works, great emphasis is placed on the linking function of annotations that reference external content via hyperlinks and bibliographical notes (Chiang and Kopak 2007).[33] In the analysis sample, implicit and explicit links to datasets and to the literature occur in more than one third of the research annotations. From these findings it can be hypothesized that annotations mostly express and *link* scholarly arguments rather than making subjective statements.

4.3 Qualitative Analysis

In order to more accurately assess the scientific contribution which the reader-generated annotations in the sample made to scholarly discourse, they were categorized into different content classes. Following Agosti et al. (2004, 5), annotations are modeled as "dialogue acts", in analogy to the Speech Act theory developed in Linguistic Pragmatics (e.g. Austin 1962). This approach helps to identify acts of communication performed by annotations while abstracting from the syntactic form of sentences. The question "How come this paper is neither cited nor discussed in the eLife paper?"[34] is, in this line of thinking, categorized as a *criticism* of the article's Related Work section rather than as a question. On the other hand, "The source of PP2 cannot be found in the methods section. It would be great if the

32 https://hyp.is/ztSf5EgYEemuOm8lZ58aWQ/elifesciences.org/articles/43599 (20.01.2020)
33 https://journals.uic.edu/ojs/index.php/fm/article/view/1961/1838#k3 (20.01.2020)
34 https://hyp.is/5PyWqAXWEeiMW8--MWJ1EA/elifesciences.org/articles/05322 (20.01.2020)

authors comment on this"[35] is a dialogue act that calls on the authors/publishers to supply additional information and can therefore be classified as an *information request* rather than a simple assertion. These dialogue acts are similar to the annotation classes in WADM (cf. Section 2). But some annotations in the sample are texts of considerable length and express more than one dialogue act. This is why the content classification remains coarse and focuses on the contribution the annotation makes to the article's content. Two main classes are of importance in this respect: the contribution of information/content versus a request for information/content. Texts of a more subjective and evaluative nature, as well as unclear cases, have to be differentiated from these two main types. This leads to the following content classification:

1) **Assertion/Contribution**
   Statement about the annotation target that is qualified by either explicit explanations or recourse to external information sources, or an evaluation of a specific property of the annotation target. The statement contributes additional information (resources)
2) **Information Request**
   Request for information (resources) regarding the topic of the annotation target that does not make a contribution as defined in 1)
3) **Unspecified Praise or Criticism**
   Evaluation of the annotation target for which no reason is given explicitly and that does not meet the conditions for 1) or 2)
4) **Unclear**
   Annotations that do not belong in any of the above categories or that can be attributed to both 1) and 2)

The main aim of this categorization is to differentiate between annotations which contribute information in a broader sense and those which request such information. The latter two types are either not explicitly related to an article's content (3) or do not fit into any of the above categories (4). This categorization would be more informative if it differentiated between, say, a question about the meaning of a diagram[36] (i.e. the publication *per se*) and a question concerning certain details of the research design (i.e. the publication's content and scientific contribution). However, more fine-grained differentiations between form/representation and content as *semantic* targets of annotations did not prove to be strictly selective when applied to the analysis sample.

35 https://hyp.is/xeMhkF6xEeia6pN7vEXmyQ/elifesciences.org/articles/20142 (20.01.2020)
36 https://hyp.is/1ipUIgXQEeiZyp_CHrjqCA/elifesciences.org/articles/08347 (20.01.2020)

Tab. 4: Frequencies of annotation content types

	Research Annotations 2012–2019	Hypothesis Research Annotations 2018–2019
1 Assertion/Contribution	65.1%	64.8%
2 Information Request	17.6%	15.7%
3 Unspecified Praise/Criticism	7.9%	4.6%
4 Unclear	9.3%	14.8%

Assertions in the sense described above comprise criticism, e.g. concerning the methodology, corrigendum notices concerning the representation of information in the article, and/or additional information on the research topic. For example, annotations often contain links to related scientific publications. In many such cases, links to other articles corroborate an argument that the annotation author is making. In cases where the annotation author identifies him- or herself as the author of the referenced article, the annotation constitutes a direct exchange of scientific arguments.

Information requests often call for explanations of the article's content or aspects thereof, e.g. the meaning of a diagram. Information is also requested about the article content per se, e.g. details of the research design that are not directly mentioned in the text. As has already been pointed out, requests are not necessarily signaled by the syntactic form of a question.[37] These dialogue acts are intended to help annotators and readers of annotations to *understand* the content of the article, rather than to discuss and evaluate it.

The class "unspecified praise/criticism" comprises annotations that do not supply new information along with the evaluation of the article authors' work. Most often, annotations in this category resemble a written form of applause: "Very helpful work. Thanks."[38] A similar gesture is thanking another annotation author for replying to a question. Annotations in the residual "unclear" category include annotations that contain tags and no text. These tags are used on the Hypothesis platform for searches across tagged webpages and thus serve as a cataloging device.

Table 4 shows that the five content types do indeed classify the analysis sample in an appropriate way. The number of unclear examples in each category is not large enough to distort the relative probabilities in the other three categories. The high percentage of unclear types for the Hypothesis annotations is largely due to test annotations made during the introduction phase. In order to assess the

37 https://hyp.is/xeMhkF6xEeia6pN7vEXmyQ/elifesciences.org/articles/20142 (20.01.2020)
38 https://hyp.is/CF4oOAXREeieiFvGf9taLA/elifesciences.org/articles/04766 (20.01.2020)

reliability of the content type classification done by the author of this study, a student assistant classified a subset of the data. The inter-rater reliability score (Cohen's kappa coefficient) is 0.54.[39]

Summing up, most reader annotations on the journal platform under analysis contribute additional information to the article or offer criticism in the form of short reviews. These contributions are not only comments, but also enrich the information offered by the online article, thus qualifying as annotations according to the TaDiRAH taxonomy. A secondary function is to serve as a "service" communication channel through which authors and publishers help readers to understand the article's content.

5 Discussion

Two limitations of the study presented in the preceding section are worth mentioning. First, the object of study is a Life Science journal and thus does not provide direct empirical evidence for annotation practices in the Humanities. However, the types of dialogue acts described here (methodological questions, criticism and replies) are in no way specific to any specific academic discipline or research method. Therefore, it seems plausible to extend the conclusions drawn here to the Arts and Humanities disciplines. The second limitation is the relatively low number of annotations in the core sample. It is clear that further research needs to be carried out once more annotation data is available. On the other hand, the content analysis has shown clear patterns which are persistent throughout the whole analysis sample.

The scientific value of the contribution which annotations make can of course not be assessed objectively in this analysis. But the relatively low number of annotations that do not supply arguments behind their evaluation of the article content allows the conclusion to be drawn that most annotations are not the spontaneous "scribbles" which have been observed in private marginal annotations, but are instead elaborated, curated statements, which confirms Hemminger and TerMaat's (2014) and Marshall and Brush's (2004) findings.

In general, public scholarly comments and annotations on research publications still appear to be a rare exception rather than the norm. Some authors have attributed this to an apparent lack of incentives, since annotations are not an established form of micro-publications (cf. Hemminer and TerMaat 2004, 2287; Shaikh-Lesko 2019). Examination of the analysis sample, however, reveals quite a

39 Some raw data is available here: https://doi.org/10.6084/m9.figshare.11872530.v2 (20.01.2020)

number of obvious motivations. Authors carrying out research or publishing on a topic related to the annotated publication insert references to their own work,[40] thus promoting it. Other researchers can also defend differing views on a scientific subject.[41] The general audience, on the other hand, gets the chance to ask and receive answers to questions about the article content.[42] The relatively low annotation rate can be explained by the often-noted restraint that interviewees expressed about making public statements about scientific works (Hemminger and TerMaat 2014).

As noted in Section 2, many annotation tools produce elaborate machine-readable metadata. Since annotation authors upload carefully crafted texts, it is reasonable to assume that they would also make the effort to apply additional metadata to their texts manually. But more empirical work is definitely needed to answer the question of how much work readers of scholarly online journals are willing to invest in order to create semantic annotations.

Concerning the form, function and content of the observed Reader annotations, it is clear that they lack the features which have been identified for individual marginal annotations: implicit semantics, relatively short text lengths and an apparently short attention span on behalf of the annotation author. This was predicted by the laboratory studies discussed in Section 3.6. The question arises of whether public annotations can be thought of as being a revised version of individual marginal annotations or if they represent an annotation type sui generis. From the analysis findings, no clear answer can be given. However, many of the annotations in the sample represent genuine dialog acts and would be useless in PIM contexts. Qualified criticism, on the other hand, makes up a large part of the Contribution annotations in the sample. Such content is likely to also occur in private collections. Given the possibility to annotate PDF files offline, then upload them to the repository of public online annotations, it is plausible to assume that such annotations would be written in private and uploaded after revision. Tracing the origin of annotations is, however, beyond the scope of this chapter. This analysis step would involve the retrieval and analysis of annotations that have been explicitly marked by users as private. Either way, from the perspective of software design, it does not seem useful to treat marginalia and journal annotations as one type of use case. In the hypothetical process of revising (and possibly, rewriting) a marginal annotation for publication, the additional effort required to move the text from one application to another does not seem to make much of a difference.

40 https://hyp.is/-vmaNAXUEei5JyczByagzg/elifesciences.org/articles/00326 (20.01.2020)
41 https://hyp.is/Ug6WZgXREeisju_9ycWXNQ/elifesciences.org/articles/21634 (20.01.2020)
42 https://hyp.is/zp0__AXREeiKVrdk4ckFng/elifesciences.org/articles/27483 (20.01.2020)

Instead, writing a public annotation with the help of marginal annotation software can be modeled in the same way as writing a research paper with the help of private annotations. This finding adds plausibility to the tentative typology presented in Section 2.

6 Summary and Outlook

The most important finding of this study is that Humanities scholars are probably willing to engage in discussions about research findings in online environments if they are presented with an easy-to-use interface. The second finding is that this engagement probably bears no relation to the way their personal data collections are structured and enriched by annotations. Instead, from the perspective of software engineering in the Digital Humanities, at least three types of annotations have to be accounted for: idiosyncratic marginal annotations, enrichments produced in collaborative work environments and annotations "born online" that become elements of asynchronous scholarly dialogs. Thus, it seems that McCarty is right: There will never be a "one-size-fits-all design" (McCarty 2020, 274) and no "killer application" (Juolas 2008) for all relevant cases of annotation use in the Humanities.

Bibliography

Adriano, Christian Medeiros and Ivan Luiz Marques Ricarte. Essential Requirements for Digital Annotation Systems. In: Revista de Sistemas de Informação da FSMA 9–21. 2012, pp. 24–44. URL: http://www.fsma.edu.br/si/edicao9/FSMA_SI_2012_1_Principal_3_en.pdf (20.01.2020).

Agosti, Maristella, Hanne Albrechtsen, Nicola Ferro, Ingo Frommholz, Preben Hansen, Nicola Orio, Emanuele Panizzi, Annelise Mark Pejtersen, and Ulrich Thiel. DiLAS: A Digital Library Annotation Service. In: Jean-François Boujut (Ed.), Proceedings of Annotation for Collaboration – A Workshop on Annotation Models, Tools and Practices. Paris: CNRS - Programme société de l'information. 2005, pp. 91–101.

Agosti, Maristella, Giorgetta Bonfiglio-Dosio, and Nicola Ferro. A Historical and Contemporary Study on Annotations to Derive Key Features for Systems Design. In: International Journal on Digital Libraries 8 (1). Berlin: Springer Verlag. 2007, pp. 1–19. DOI: 10.1007/s00799-007-0010-0.

Agosti, Maristella, Nicola Ferro, Ingo Frommholz, and Ulrich Thiel. Annotations in Digital Libraries and Collaboratories – Facets, Models and Usage. In: Rachel Heery and Liz Lyon (Eds.), Research and Advanced Technology for Digital Libraries (Lecture Notes in Computer Science). Berlin, Heidelberg: Springer. 2004, pp. 244–255. DOI: 10.1007/978-3-540-30230-8_23.

Antonijević, Smiljana. Amongst Digital Humanists – An Ethnographic Study of Digital Knowledge Production. New York: Palgrave Macmillan US. 2015.DOI: 10.1057/9781137484185.

Antonijević, Smiljana. Developing Research Tools via Voices from the Field. In: dh+lib. 2016. URL: https://acrl.ala.org/dh/2016/07/29/developing-research-tools-via-voices-from-the-field/ (20.01.2020).

Antonijević, Smiljana and Ellysa Stern Cahoy. Personal Library Curation: An Ethnographic Study of Scholars' Information Practices. In: Portal. Libraries and the Academy 14 (2). Baltimore, MD: Johns Hopkins University Press. 2014, S. 287–306. DOI: 10.1353/pla.2014.0010.

Austin, John Langshaw. How to Do Things with Words. Oxford: Oxford University Press. 1962.

Bauer, Matthias and Angelika Zirker. Whipping Boys Explained: Literary Annotation and Digital Humanities. In: Kenneth Price and Ray Siemens (Eds.), Literary Studies in the Digital Age. New York: Modern Language Association of America. 2015. DOI: 10.1632/lsda.2015.12. URL: https://dlsanthology.mla.hcommons.org/whipping-boys-explained-literary-annotation-and-digital-humanities/ (20.01.2020).

Barreau, Deborah. The Persistence of Behavior and Form in the Organization of Personal Information. In: Journal of the American Society for Information Science and Technology 59 (2). Hoboken, NJ: Wiley-Blackwell. 2008, pp. 307–317. DOI: 10.1002/asi.20752.

Becker, Rainer, Michael Bender, Luise Borek, Canan Hastik, Thomas Kollatz, Beata Mache, Harald Lordick, and Ruth Reiche. Digitale Annotationen: "Best Practices" und Potentiale (R 6.2.1) Teil II. [Research Report]. DARIAH-DE. 2016. URL: https://wiki.de.dariah.eu/download/attachments/14651583/r621_2.pdf?version=2&modificationDate=1552902388945&api=v2 (20.01.2020).

Belanger, Marie-Eve. Annotations and the Digital Humanities Research Cycle: Implications for Personal Information Management. iConference 2010. Urbana-Champaign, IL, USA. 2010. URL: http://hdl.handle.net/2142/15035 (20.01.2020).

Blair, Ann. Note Taking as an Art of Transmission. Critical Inquiry 31 (1). Chicago, IL: The University of Chicago Press. 2004, pp. 85–107. DOI: 10.1086/427303.

Blustein, James, David Rowe, and Ann-Barbara Graff. Making Sense in the Margins: A Field Study of Annotation. In: Gradmann, Stefan, Francesca Borri, Carlo Meghini, Heiko Schuldt (Eds.), Research and Advanced Technology for Digital Libraries. Berlin, Heidelberg: Springer-Verlag. 2011, pp. 252–259.

Bold, Melanie Ramdarshan and Kiri L. Wagstaff. Marginalia in the Digital Age: Are Digital Reading Devices Meeting the Needs of Today's Readers? In: Library & Information Science Research 39 (1). Amsterdam: Elsevier. 2017, pp. 16–22. DOI: 10.1016/j.lisr.2017.01.004.

Borek, Luise, Quinn Dombrowski, Jody Perkins, and Christof Schöch. TaDiRAH: A Case Study in Pragmatic Classification. In: Digital Humanities Quarterly 010 (1). 2016. URL: https://digitalhumanities.org/dhq/vol/10/1/000235/000235.html (20.01.2020).

Bradley, John. Thinking about Interpretation: Pliny and Scholarship in the Humanities. In: Literary and Linguistic Computing 23 (3). Oxford: Oxford University Press. 2008, pp. 263–279. DOI: 10.1093/llc/fqn021.

Bradley, John and Paul Vetch. Supporting Annotation as a Scholarly Tool—Experiences From the Online Chopin Variorum Edition. In: Literary and Linguistic Computing 22 (2). Oxford: Oxford University Press. 2007, pp. 225–241. DOI: 10.1093/llc/fqm001.

Burroughs, Jennie M. No Uniform Culture: Patterns of Collaborative Research in the Humanities. In: Portal 17 (3). Baltimore, MD: Johns Hopkins University Press. 2017, pp. 507–527. DOI: 10.1353/pla.2017.0032.

Carpenter, Todd A. iAnnotate — Whatever Happened to the Web as an Annotation System? In: The Scholarly Kitchen. April 30, 2013. URL: https://scholarlykitchen.sspnet.org/2013/04/30/iannotate-whatever-happened-to-the-web-as-an-annotation-system/ (20.01.2020).

Case, Donald Owen. The Collection and Use of Information by Some American Historians: A Study of Motives and Methods. In: The Library Quarterly: Information, Community, Policy 61 (1). Chicago: The University of Chicago Press Books. 1991, pp. 61–82.

Cerf, Vinton G., et al. National Collaboratories: Applying Information Technologies for Scientific Research. Washington, DC: National Academy Press. 1993. URL: https://www.nap.edu/read/2109/chapter/3#7 (20.01.2020).

Chiang, Chia-Ning. A Multi-Dimensional Approach to the Study of Online Annotation. University of British Columbia. [Dissertation]. 2010. DOI: 10.14288/1.0071066.

Chu, Clara M. Literary Critics at Work and Their Information Needs: A Research-Phases Model. In: Library & Information Science Research 21 (2). Amsterdam: Elsevier. 1999, pp. 247–273. DOI: 10.1016/S0740-8188(99)00002-X.

Cronin, Blaise. Scholarly Communication and Epistemic Cultures. In: New Review of Academic Librarianship 9 (1). London: Taylor & Francis. 2003, pp. 1–24. DOI: 10.1080/13614530410001692004.

Cushing, Amber L. and Odile Dumbleton. "We Have to Make an Effort with It": Exploring the Use of Stages to Help Understand the Personal Information Management Needs of Humanities and Social Science Doctoral Students Managing Dissertation Information. In: IFLA Journal 43 (1). London: Sage Publications Sage UK. 2017, pp. 40–50. DOI: 10.1177/0340035216686983.

Dallas, Costis, Claire Clivaz, Nephelie Chatzidiakou, Jurij Hadalin, Elena Gonzalez-Bianco, Beat Immenhauser, and Maciej Maryl. DiMPO – a DARIAH Infrastructure Survey on Digital Practices and Needs of European Scholarship. In: Fabio Ciotti and Gianfranco Crupi (Eds.), AIUCD2017 – Book of Abstracts. Rome: Associazione per l'Informatica Umanistica e la Cultura Digitale. 2017, pp. 197–199. DOI: 10.6092/unibo/amsacta/5885.

Di Donato, Francesca, Christian Morbidoni, Simone Fonda, Alessio Piccioli, Marco Grassi, and Michele Nucci. Semantic Annotation with Pundit: A Case Study and a Practical Demonstration. In: Proceedings of the 1st International Workshop on Collaborative Annotations in Shared Environment: Metadata, Vocabularies and Techniques in the Digital Humanities DH-CASE '13. New York: ACM. 2013, pp. 16:1–16:4. DOI: 10.1145/2517978.2517995.

de la Flor, Grace, Marina Jirotka, Paul Luff, John Pybus, and Ruth Kirkham. Transforming Scholarly Practice: Embedding Technological Interventions to Support the Collaborative Analysis of Ancient Texts. In: Computer Supported Cooperative Work (CSCW) 19 (3). Heidelberg: Springer. 2010, pp. 309–334. DOI: 10.1007/s10606-010-9111-1.

Gius, Evelyn and Janina Jacke. The Hermeneutic Profit of Annotation: On Preventing and Fostering Disagreement in Literary Analysis. In: International Journal of Humanities and Arts Computing 11 (2). Edinburgh: Edinburgh University Press. 2017, pp. 233–254. DOI: 10.3366/ijhac.2017.0194.

Given, Lisa M. and Rebekah Willson. Information Technology and the Humanities Scholar: Documenting Digital Research Practices. In: Journal of the Association for Information Science and Technology 69 (6). Hoboken, NJ: Wiley. 2018, pp. 807–819. DOI: 10.1002/asi.24008.

Grassi, Marco, Christian Morbidoni, Michele Nucci, Simone Fonda, and Francesco Piazza. Pundit: Augmenting Web Contents with Semantics. In: Literary and Linguistic Computing 28 (4). Oxford: Oxford University Press. 2013, pp. 640–659. DOI: 10.1093/llc/fqt060.

Hemminger, Bradley M. NeoNote: Suggestions for a Global Shared Scholarly Annotation System. In: D-Lib Magazine 15 (5/6). 2009. DOI: 10.1045/may2009-hemminger.

Hemminger, Bradley M. and Julia TerMaat. Annotating for the World: Attitudes toward Sharing Scholarly Annotations. In: Journal of the Association for Information Science and Technology 65 (11). Oxford: Oxford University Press. 2014, pp. 2278–2292. DOI: 10.1002/asi.23124.

Hennicke, Steffen, Stefan Gradmann, Kristin Dill, Gerold Tschumpel, Klaus Thoden, Christian Morbindoni, and Alois Pichler. D3.4 – Research Report on DH Scholarly Primitives. 2015. URL: https://dm2e.eu/files/D3.4_2.0_Research_Report_on_DH_Scholarly_Primitives_150402.pdf (20.01.2020).

Hunter, Jane. Collaborative Semantic Tagging and Annotation Systems. In: Annual Review of Information Science and Technology 43 (1). Hoboken, NJ: Wiley. 2009, pp. 1–84. DOI: 10.1002/aris.2009.1440430111.

Hunter, Jane, Tim Cole, Robert Sanderson, and Herbert Van de Sompel. The Open Annotation Collaboration: A Data Model to Support Sharing and Interoperability of Scholarly Annotations. In: Proceedings of the Digital Humanities 2010 Conference. London: Centre for Computing in the Humanities, King's College London. 2010. URL: http://dh2010.cch.kcl.ac.uk/academic-programme/abstracts/papers/html/ab-860.html (20.01.2020).

Jones, Steven E., Peter Shillingsburg, and George K. Thiruvathukal. E-Carrel: An Environment for Collaborative Textual Scholarship. In: Journal of the Chicago Colloquium on Digital Humanities and Computer Science 1 (2). 2010. DOI: 10.6082/M1GB227B.

Jones, William. Personal Information Management. In: Annual Review of Information Science and Technology 41 (1). Hoboken, NJ: Wiley. 2007, pp. 453–504. DOI: 10.1002/aris.2007.1440410117.

Juola, Patrick. Killer Applications in Digital Humanities. In: Literary and Linguistic Computing 23 (1). Oxford: Oxford University Press. 2008, pp. 73–83. DOI: 10.1093/llc/fqm042.

Karcher, Sebastian and Nicholas Weber. Annotation for Transparent Inquiry: Transparent Data and Analysis for Qualitative Research. In: IASSIST Quarterly 43 (2). 2019, pp. 1–9. DOI: 10.29173/iq959.

Katz, J. Sylvan and Ben R. Martin. What Is Research Collaboration? In: Research Policy 26 (1). Amsterdam: Elsevier. 1997, pp. 1–18. DOI: 10.1016/S0048-7333(96)00917-1.

Klein, Perry and Tracey Leacock. Distributed Cognition as a Framework for Understanding Writing. In: V. W. Berninger (Ed.), Past, Present and Future Contributions of Cognitive Writing Research to Cognitive Psychology. New York, London: Psychology Press. 2012, pp. 133–152.

Kopak, Rick and Chia-Ning Chiang. Annotating and Linking in the Open Journal Systems. In: First Monday 12 (10). Chicago, IL: University of Illinois. 2007. DOI: 10.5210/fm.v12i10.1961.

Liu, Ziming. Reading Behavior in the Digital Environment: Changes in Reading Behavior Over the Past Ten Years. In: Journal of Documentation 61 (6). Bingley: Emerald Publishing. 2005, pp. 700–712. DOI: 10.1108/00220410510632040.

Lordick, Harald. Crossing Borders Through (Web) Annotations. [Presentation] May 29, 2015. DOI: 10.5281/zenodo.32509.

Lordick, Harald, Rainer Becker, Michael Bender, Luise Borek, Canan Hastik, Thomas Kollatz, Beata Mache, Andrea Rapp, Ruth Reiche, and Niels-Oliver Walkowski. Digitale Annotationen in der Geisteswissenschaftlichen Praxis. In: Bibliothek Forschung Und Praxis 40 (2). Berlin: De Gruyter. 2016, pp. 186–199. DOI: 10.1515/bfp-2016-0042.

Ma, Feicheng, Yating Li, and Baitong Chen. Study of the Collaboration in the Field of the Chinese Humanities and Social Sciences. In: Scientometrics 100 (2). Berlin: Springer. 2014, pp. 439–458. DOI: 10.1007/s11192-014-1301-z.

Marshall, Catherine C.. Annotation: From Paper Books to the Digital Library. In: Proceedings of the Second ACM International Conference on Digital Libraries. DL '97. New York: ACM. 1997, pp. 131–140. DOI: 10.1145/263690.263806.

Marshall, Catherine C. The Future of Annotation in a Digital (Paper) World. In: Susan Harum and Michael Twidale (Eds.), Successes & Failures of Digital Libraries: [papers presented at the 1998 Clinic on Library Applications of Data Processing, March 22–24, 1998]. Champaign, IL: Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. 2000, pp. 97–117. URL: http://hdl.handle.net/2142/25539 (20.01.2020).

Marshall, Catherine C. and A. J. Bernheim Brush. Exploring the Relationship Between Personal and Public Annotations. In: Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries JCDL '04. New York: ACM. 2004, pp. 349–357. DOI: 10.1145/996350.996432.

McCarty, Willard. Making and Studying Notes. Towards a Cognitive Ecology of Annotation. In: Julia Nantke and Frederik Schlupkothen (Eds.), Annotations in Scholarly Editions and Research. Functions, Differentiation, Systematization. Berlin: De Gruyter. 2020, pp. 271–297.

Müller-Birn, Claudia, Tina Klüwer, André Breitenfeld, Alexa Schlegel, and Lukas Benedix. Neonion: Combining Human and Machine Intelligence. In: Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing CSCW'15. New York, NY: ACM. 2015, pp. 223–226. DOI: 10.1145/2685553.2699012.

Müller-Birn, Claudia, Alexa Schlegel, and Christian Pentzold. Softwarenutzung in der geisteswissenschaftlichen Forschungspraxis. In: Wolfgang Prinz, Jan Borchers, and Matthias Jarke (Eds.), Mensch und Computer 2016. Aachen: Gesellschaft für Informatik e. V. 2016. DOI: 10.18420/muc2016-mci-0114.

O'hara, Kenton P., Alex Taylor, William Newman, and Abigail J. Sellen. Understanding the Materiality of Writing from Multiple Sources. In: International Journal of Human-Computer Studies 56 (3). Amsterdam: Elsevier. 2002, pp. 269–305. DOI: 10.1006/ijhc.2001.0525.

Ollé, Candela and Ángel Borrego. A Qualitative Study of the Impact of Electronic Journals on Scholarly Information Behavior. In: Library & Information Science Research 32 (3). Amsterdam: Elsevier. 2010, pp. 221–228. DOI: 10.1016/j.lisr.2010.02.002.

Ossenblok, Truyken L. B., Frederik T. Verleysen, and Tim C. E. Engels. Coauthorship of Journal Articles and Book Chapters in the Social Sciences and Humanities (2000–2010). In: Journal of the Association for Information Science and Technology 65 (5). Hoboken, NJ: Wiley. 2014, pp. 882–897. DOI: 10.1002/asi.23015.

Ovsiannikov, Ilia A., Michael A. Arbib, and Thomas H. McNeill. Annotation Technology. In: International Journal of Human-Computer Studies 50 (4). Amsterdam: Elsevier. 1999, pp. 329–362. DOI: 10.1006/ijhc.1999.0247.

Palmer, Carole L., Lauren C. Teffeau, and Carrie M. Pirmann. Scholarly Information Practices in the Online Environment: Themes from the Literature and Implications for Library Service Development. Dublin, OH: OCLC Research. 2009. URL: https://www.oclc.org/content/dam/research/publications/library/2009/2009-02.pdf (20.01.2020).

Pape, Sebastian, Christof Schöch, and Lutz Wegner. TEICHI and the Tools Paradox. In: Journal of the Text Encoding Initiative 2. TEI consortium. 2012. DOI: 10.4000/jtei.432.

Perkel, Jeffrey M. Annotating the Scholarly Web. In: Nature 528 (3). Berlin: Springer. 2015, pp. 153–154. DOI: 10.1038/528153a.

Piro, Fredrik Niclas, Dag W. Aksnes, and Kristoffer Rørstad. A Macro Analysis of Productivity Differences across Fields: Challenges in the Measurement of Scientific Publishing. In: Journal of the American Society for Information Science and Technology 64 (2). Hoboken, NJ: Wiley. 2013, pp. 307–320. DOI: 10.1002/asi.22746.

Qayyum, Muhammad Asim. Capturing the Online Academic Reading Process. In: Information Processing & Management 44 (2). Amsterdam: Elsevier. 2008, pp. 581–595. DOI: 10.1016/j.ipm.2007.05.005.

Shaikh-Lesko, Rina. Web Annotation Tool Hypothesis Hits a Milestone. In: Nature 569 (May). Berlin: Springer. 2019, pp. 295–295. DOI: 10.1038/d41586-019-01427-9.

Staines, Heather. Making Peer Review More Transparent with Open Annotation. Hypothesis. 2019. URL: https://web.hypothes.is/blog/transparent-peer-review/ (20.01.2020).

Stiller, Juliane, Klaus Thoden, Oona Leganovic, Christian Heise, Mareike Höckendorff, and Timo Gnadt. Nutzungsverhalten in den Digital Humanities (R 1.2.1/ M 7.6). [Research Report] DARIAH-DE. 2015. URL: https://wiki.de.dariah.eu/download/attachments/14651583/Report1.2.1-final3.pdf?version=1&modificationDate=1426154224304&api=v2 (20.01.2020).

Stone, Sue. CRUS Humanities Research Programme. In: Humanities Information Research: Proceedings of a Seminar. Sheffield: University of Sheffield. 1980, pp. 15–26.

Stone, Sue. Humanities Scholars: Information Needs and Uses. In: Journal of Documentation 38 (4). Bingley: MCB UP Ltd. 1982, pp. 292–313. DOI: 10.1108/eb026734.

Unsworth, John. Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This? London. 2000. URL: http://www.iath.virginia.edu/~jmu2m/Kings.5-00/primitives.html (20.01.2020).

Unsworth, John and Charlotte Tupman. Interview with John Unsworth, Carried out and Transcribed by Charlotte Tupman. In: Collaborative Research in the Digital Humanitär. London: Routledge. 2016, pp. 231–240.

Uva, Peter. Information-Gathering Habits of Academic Historians: Reports of the Pilot Study. [Research Report] Library of the Upstate Medical Center. Washington, DC: ERIC. 1977. URL: https://files.eric.ed.gov/fulltext/ED142483.pdf (20.01.2020).

Walkowski, Niels-Oliver. The Landscape of Digital Annotation and Its Meaning. In: Conference on Language Technologies & Digital Humanities, Ljubljana. 2016a, pp. 6–11. URL: http://www.sdjt.si/wp/wp-content/uploads/2016/09/JTDH-2016_Walkowski_The-Landscape-of-Digital-Annotation.pdf (20.01.2020).

Walkowski, Niels-Oliver. Digitale Annotationen: 'Best Practices' Und Potentiale (R 6.2.1) Part I. [Research Report]. DARIAH-DE. 2016b. URL: https://wiki.de.dariah.eu/download/attachments/14651583/r621_1.pdf?version=2&modificationDate=1552902373909&api=v2 (20.01.2020).

Wheary, Jennifer, Lee Wild, Bernard Schutz, and Christina Weyher. Living Review in Relativity: Thinking and Developing Electronically. In: Journal of Electronic Publishing 4 (2). Ann Arbor, MI: Michigan Publishing. 1998. DOI: 10.3998/3336451.0004.205.