New Challenges for the Archiving of Digital Writing

Heiko Zimmermann

University of Trier

Follow this and additional works at: <https://docs.lib.purdue.edu/clcweb>

Part of the [American Popular Culture Commons](#), [Art Practice Commons](#), [Comparative Literature Commons](#), [Creative Writing Commons](#), [Data Storage Systems Commons](#), [Digital Humanities Commons](#), [English Language and Literature Commons](#), [European Languages and Societies Commons](#), [Interactive Arts Commons](#), [Library and Information Science Commons](#), [Modern Literature Commons](#), [Other Film and Media Studies Commons](#), [Reading and Language Commons](#), [Rhetoric and Composition Commons](#), and the [Theatre and Performance Studies Commons](#)

Dedicated to the dissemination of scholarly and professional information, [Purdue University Press](#) selects, develops, and distributes quality resources in several key subject areas for which its parent university is famous, including business, technology, health, veterinary medicine, and other selected disciplines in the humanities and sciences.

CLCWeb: Comparative Literature and Culture, the peer-reviewed, full-text, and open-access learned journal in the humanities and social sciences, publishes new scholarship following tenets of the discipline of comparative literature and the field of cultural studies designated as "comparative cultural studies." Publications in the journal are indexed in the Annual Bibliography of English Language and Literature (Chadwyck-Healey), the Arts and Humanities Citation Index (Thomson Reuters ISI), the Humanities Index (Wilson), Humanities International Complete (EBSCO), the International Bibliography of the Modern Language Association of America, and Scopus (Elsevier). The journal is affiliated with the Purdue University Press monograph series of Books in Comparative Cultural Studies. Contact: <clcweb@purdue.edu>

Recommended Citation

Zimmermann, Heiko. "New Challenges for the Archiving of Digital Writing." *CLCWeb: Comparative Literature and Culture* 16.5 (2014): <<https://doi.org/10.7771/1481-4374.2495>>

This text has been double-blind peer reviewed by 2+1 experts in the field.

The above text, published by Purdue University Press ©Purdue University, has been downloaded 386 times as of 11/07/19.

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

This is an Open Access journal. This means that it uses a funding model that does not charge readers or their institutions for access. Readers may freely read, download, copy, distribute, print, search, or link to the full texts of articles. This journal is covered under the [CC BY-NC-ND license](#).

CLCWeb: Comparative Literature and Culture, the peer-reviewed, full-text, and open-access learned journal in the humanities and social sciences, publishes new scholarship following tenets of the discipline of comparative literature and the field of cultural studies designated as "comparative cultural studies." In addition to the publication of articles, the journal publishes review articles of scholarly books and publishes research material in its *Library Series*. Publications in the journal are indexed in the Annual Bibliography of English Language and Literature (Chadwyck-Healey), the Arts and Humanities Citation Index (Thomson Reuters ISI), the Humanities Index (Wilson), Humanities International Complete (EBSCO), the International Bibliography of the Modern Language Association of America, and Scopus (Elsevier). The journal is affiliated with the Purdue University Press monograph series of Books in Comparative Cultural Studies. Contact: <clcweb@purdue.edu>

Volume 16 Issue 5 (December 2014) Article 7
Heiko Zimmermann,
"New Challenges for the Archiving of Digital Writing"
<<http://docs.lib.purdue.edu/clcweb/vol16/iss5/7>>

Contents of **CLCWeb: Comparative Literature and Culture 16.5 (2014)**
Special Issue ***New Work on Electronic Literature and Cyberculture***
Ed. Maya Zalbidea, Mark C. Marino, and Asunción López-Varela
<<http://docs.lib.purdue.edu/clcweb/vol16/iss5/>>

Abstract: In his article "New Challenges for the Archiving of Digital Writing" Heiko Zimmermann discusses the challenges of the preservation of digital texts. In addition to the problems already at the focus of attention of digital archivists, there are elements in digital literature which need to be taken into consideration when trying to archive them. Zimmermann analyses two works of digital literature, the collaborative writing project *A Million Penguins* (2006-2007) and Renée Tuner's *She...* (2008) and shows how the ontology of these texts is bound to elements of performance, to direct social interaction of writers and readers to the uniquely subjective reading process, and to real-time access to data. Zimmermann posits that these features of the digital text pose further challenges for digital archiving and libraries.

Heiko ZIMMERMANN**New Challenges for the Archiving of Digital Writing**

The preservation of digital literature poses a number of hitherto unseen challenges for archivists, librarians, readers, writers, publishers, and content providers. In addition to the already established challenges like changing software, there are further obstacles related to the dimensions of performance and social interaction in or via works of digital literature. As Margaret Hedstrom points out in 1995, "recording media for digital materials are vulnerable to deterioration and catastrophic loss, and even under ideal conditions they are short lived relative to traditional format materials" and she also points to the problem of the looming obsolescence in retrieval and playback technology as well as "the absence of established standards, protocols, and proven methods for preserving digital information" (<<http://www.uky.edu/~kiernan/DL/hedstrom.html>>). Microfilming born-digital media seems to be a way of transgressing the temporal limitations of digital data carriers—at least for some texts. However, "many types of digital objects do not have print equivalents and cannot be preserved in non-digital formats" (Hedstrom <<http://www.uky.edu/~kiernan/DL/hedstrom.html>>). For these media, there seem to be number of possible archiving strategies at hand, amongst them are, first, the preservation of the simplest possible digital formats that still retain the intrinsic values of the original. This can, for instance, mean the transcoding of text files into the American Standard Code for Information Interchange (ASCII) or of visually more complex documents into the Tagged Image File Format (TIFF) or the Portable Document Format (PDF). The second strategy is the preservation of the original hardware and software environment. Due to the rapid deterioration of hardware beyond repair, most archivists would dismiss this approach. However, if archivists want to retain publications within their original system or a compatible newer system, they will have to obviate bitrate changes by regularly copying it within the system thus refreshing the data on the respective storage device. They will also have to minimize the risk of fatal data loss by replicating the data in order have backup copies of documents. If they try to keep data within the original family of software environments, they will also have to migrate the archived texts into a version of the respective digital framework which is supported (see, e.g., Edmond <<http://stateofthediscipline.acla.org/entry/archive-now>>). A more promising approach to archiving digital literature is the emulation, that is, the imitation, of possibly outdated hardware systems with the help of emulation software. This would allow archivists to use digital texts within the original applications and operating systems, which are executed within these emulated hardware environments.

One of the most influential proponents of this approach is Jeff Rothenberg, who proposes the archival of digital information in software-independent form using so-called "envelopes" which do not only contain the digital media objects, the original software environment, the original operating system software and emulator specifications, but also metadata which explain and document the content and the appropriate way of retrieving, displaying, and processing the message within the envelope (see Hedstrom <<http://www.uky.edu/~kiernan/DL/hedstrom.html>>; Rothenberg, *Avoiding*, "Ensuring"). In order for this emulation solution to work over longer periods of time, a number of techniques need to be developed (see Rothenberg, *Avoiding*). First, generalizable techniques for the specification of emulators need to be developed, so that they could be run on unknown future computers. These specifications should capture all attributes that are required to recreate the behavior of current and future digital documents (see Rothenberg, *Avoiding*). This first point alone could pose a problem in all archiving endeavors not only because of the demanding task to develop technical protocols but also because of the form of future documents. Creative output in the future can take the shape of entirely new forms that we might not be able to imagine at the moment. The arrival of *Social Web* <<http://www.socialweb.net/>> is one example. A decade ago, hardly anybody was able to predict how the World Wide Web 2.0 would change our communication and social environment. This is not only about Social Networking Sites (SNS) like Facebook or geosocial networking applications like Grindr

that have led to dramatic changes in social dynamics, made away with old communication patterns, and allowed for the creation of new ones. This is also about the effect these channels have on creative processes, enabling fan-fiction to become a world best seller (e.g., *Fifty Shades of Grey*), enabling musicians to become successful without the filter of big labels (e.g., Arctic Monkeys or Justin Bieber), and enabling self-organization of distributed contributions to a major work of art (think of memes or about the pseudo-folklore figure Slenderman which is often called the "first great myth of the web"). As George P. Landow claims, media paradigms become quickly invisible. It is hard for us to understand inter-word spacing, for instance, as a central technique in the age of the book. If our media paradigms and techniques are hard to grasp, if they are invisible to us and taken for granted, how can we even try to imagine what the media paradigms of the future will be?

The second technique that needs to be developed in order to allow media texts to be archived in digital envelopes is a way of saving the metadata that is needed to "find, access, and recreate digital documents, so that emulation techniques can be used for preservation" (Rothenberg, *Avoiding* 17). This has to be done in a human-readable form. Although it might sound like a simple task, Rothenberg illustrates the complexity of the issue: "The year is 2045, and my grandchildren (as yet unborn) are exploring the attic of my house (as yet unbought). They find a letter dated 1995 and a CD-ROM (compact disk). The letter claims that the disk contains a document that provides the key to obtaining my fortune (as yet unearned). My grandchildren are understandably excited, but they have never seen a CD before—except in old movies—and even if they can somehow find a suitable disk drive, how will they run the software necessary to interpret the information on the disk? How can they read my obsolete digital document?" (*Avoiding* 1). In this example, the letter added to the CD is a starting point for the possible quest to obtain Rothenberg's fortune. It is human-readable metadata, but it does not suffice to re-enact the desired document. To be able to long-term archive digital media, one will have to write very long letters explaining every detail of a document, or find some alternative feasible technique. However, in order to compose such a message, one will have to understand the technical details of a given product of culture. In the case of a Facebook conversation, for instance, the public neither knows where it is stored nor its format. According to Rothenberg, the third technique that needs to be developed is a way of encapsulating described metadata, data, and emulators in a format and a medium, that ensures their coherence and prevents their corruption. Especially when concatenated in compressed formats, the rate of possible bit rot of digital data is hard to identify and may result in the loss of the whole document. The physical deterioration of data carriers poses yet another challenge for archivists: A printed book might last for about a century or longer, a DVD for less than decade. Despite new developments in the field of optical data carriers (M-Disk), physical longevity remains an issue.

The above-discussed technical challenges are just one aspect of the problems in archiving digital literature. Having solved them might allow future generations to read governmental bulletins in then exotic formats and software environments. It will not allow them to read certain types of digital literature. In the following, I argue that inherent structures of these types of literary texts—detached from their respective technical implementation—can render them "unarchivable." Highlighting two literary texts will prove my thesis, namely *She...* <http://www.fudgethefacts.com/she/_she.html> by Renée Turner and *A Million Penguins* <<http://www.amillionpenguins.com>> written collaboratively by thousands of contributors. In order to describe the structures of these texts, I am using Espen J. Aarseth's texton-scripton model in an adapted form, which renders it productive for the classification of authorship of digital literature (see 62-65). Underlying Aarseth's original model is the idea that there are two sides of digital texts, the "text" and the "script." The elements of the text, the "textons," are projected into "scriptons" during the reading process. The order in which the textons are invoked is determined by the "traversal function," which is an element of the text. This function has seven dimensions with several possible variables, amongst them the so-called "user function." Users can read a work of digital literature either "textonically," adding, modifying, or deleting textons; "configuratively," configuring

the projection of the textons into scriptons; and "exploratively" or "unspecifically," when they are just interpreting the text or sounding it out in order to determine its size and qualities, for example.

Turner's *She...* is a literary mashup, which combines articles from news websites with videos on YouTube and fictional text vignettes. This is achieved by embedding articles and videos into the web-site of *She...* via so-called "inline frames," html elements which allow for the display of content from other web sites as independent media documents within another web site. The core of *She...* is formed by episodes from the lives of seven more or less famous women as well as their proposed self-perception in selected formative moments of their lives. Among these women are Hillary Clinton, Sharon Stone, and Diana, Princess of Wales. The news texts contextualize the given situations of the characters, whereas the YouTube videos and links to the image search engine of Google create a feeling of verisimilitude.

30/05/08: Sharon Stone says earthquake in China could have been karma

Figure 1: Screenshot "Sharon" in Turner's *She...* Copyright does not apply.

Figure 1 shows an example of the setup of Turner's text. The scripton "Sharon" contains the heading "Sharon Stone says earthquake in China could have been karma." This scripton is the result not only of the underlying texton on Turner's web site but also of three further textons whose projections are embedded in inline frames within the text. The first frame contains a *Guardian* article by Henry Barnes published on 28 May 2008 about the Chinese reactions to Stone's, presumably inconsiderate, comment about the Sichuan earthquake and the Tibet policy of the People's Republic of China. The article in the *Guardian* refers to a YouTube video that can be found in the second inline frame on the page. Barnes's text also reports about the luxury goods producer Christian Dior, who had removed all advertisements featuring the actress from their billboards in Beijing. Additional information about this can be found in the third inline frame, which contains the article "Dior Drops Sharon Stone in China" published by BBC News on 29 2008. Next to this inline frame, there is the linked text "Oh God, here we go again," which directs the reader to the next scripton—an entirely fictional text vignette about the character. Obviously, the project has little control over the embedded media texts. This is illustrated best using other screenshots (figures 2 and 3):

10/14/07: Diana's private letters to Dodi revealed at inquest

Figure 2: Football manager Calmund in a geo-targeted video advertisement in the scripton "Diana." Copyright does not apply.

To die or
surrender, the
choice would
inevitably arise.

Figure 3: Scripton "Karina" showing a 404 error page. Copyright does not apply.

By embedding articles from news websites, YouTube videos and Google image search results, the text is rendered uncontrollable by the publisher of *She...* Layout, surrounding content elements and displayed advertisements are subject to rapid change, which means that the project becomes unstable and fluent. Moreover, the technical framework is also in a constant process of change. YouTube, for instance, does not allow the embedding of its videos in inline frames any longer and tries to control this by running a JavaScript on their websites. This results in a traversal through *She...* which is marked by being led to YouTube error pages if the reader does not click the stop button of his or her browser quickly enough or has disabled JavaScript from the beginning. All these problems were anticipated by the author. Turner writes in a project description that "it was fascinating to work with pre-existing online material and re-narrate through them. Of course, this also makes the work vulnerable, as webpages can be removed and urls changed ... The piece will collapse irretrievably into a collection of 404 messages" (<http://www.fudgethefacts.com/she/_/she.html>).

This means that the corrosion of the project and the loss of authorial control are conceptual and, thus, an inherent element of the project. Important is not that parts of the scripton's origin come from news websites, but that they are embedded in their live version—not as quotes, screenshots, or off-line files. The power over these scriptons lies in the hands of the respective content providers. The power also lies in the hands of the reader, who can read the text configuratively, configuring the projection of the textons into scriptons, or can even become a "textonic author" (see Zimmermann, *Rekonfigurationen* 21, 148-52). In traditional, easily archivable texts, "authors" are the only agents in the textual action space that act textonically, draft, and write the text. In many works of digital literature, this is different, as recent in-depth analyses of textual action spaces of several works of digital literature have shown (see Zimmermann, *Rekonfigurationen*). Agents who act textonically, the textonic authors, create the text or parts of it rather late in the continuum of production and reception, this leads to a number of problems, amongst them, problems of archivability (248-55). To make this more graspable: The reader could, for example, browse through the material in the inline frames and potentially end up reading the current weather forecast instead of continuing his or her traversal with one of the text vignettes about the women (see Figure 4). Moreover, the dynamic programming of the web sites creates different reading experiences for every reader. During my reading, for instance, I had to watch a video commercial featuring German football manager Reiner Calmund, who recommended me to use his allegedly favorite aviation price comparison web site before watching the last minutes in the life of Lady Di (see Figure 2). At times, this can leave a rather specific impression—in my case a macabre one.

Oh God, here we
go again.

Figure 4: Scripton "Sharon" with current weather forecast for the British Isles. Copyright does not apply.

This influence of content providers and readers leads to a question which is relevant for archiving purposes: where does Turner's mash-up of short prose start and where does it end? What are the textual boundaries? If one wanted to archive this project, notwithstanding the technological hindrances of long-term archiving described above, one could do this using snapshots as is done by the Internet Archive. However, this way of archiving would neglect intrinsic literary qualities of the text. It would lose the dependence of the text on the reader, the content providers, and the time, that is, its momentariness—all these are conceptual parts of *She*...

A Million Penguins was a huge collaborative writing effort, started by Penguin Books and De Montfort University, Leicester. It existed in the form of a Wiki to which hundreds of people contributed text within the active project time of about one year. From 1 February 2006 until 7 March 2007, 1476 people registered to edit the wiki novel. They created more than a thousand pages in HyperText Markup Language (html) and edited them. In total, more than 11,000 edits took place within the active project time (see Mason and Thomas 4). The speed and the amount of editing that took place are not the only characteristic features of the novel's design. As a Wiki, it also had history pages, tracking the edits of a given page, talk/comment pages allowing communication about the content and the development of a given page, and the user pages. All of these were used quite actively, so that one cannot say that the front pages, the pages that were to become a novel, are the only focus of activity:

[main page](#) | [discussion](#) | [edit](#) | [history](#)

List Of Characters Here

[Major And Minor Characters And the Menu for this Evening](#)

[Artie](#) - having a whale of a time. Likes bananas. Not spinash. Urk.

[Da Vinci Cod](#)

[The Da Vinci Cod](#) A mysterious, primalt entity - is it fish.. or is it something more??

[Gestalt](#): The person or persons who is all of us, the narrator of the preface.

[Papa Cottonballs](#): Rabbit step father of Mr. Gestalt.

[Idego](#), or [id/ego](#) author of "Towards a Systematic Analysis of the Dominance of Canine over Feline Domestication in Modern Society". Troubled by Ego issues. Likes to mention reviews received in The turdbog gazette and the cat fancier's of 4th street mailer.

[Gina](#): Girlfriend of Mark. Gina gets abducted just like Helena, for the purposes of the evil 'command'. They are used by 'command' as pawns in their smuggling operation, which goes very high into corruption in the government.

[Sara](#): Girlfriend Of Carlo

[Rollerson](#): Agent Who Commissioned Sara About The Mysterious 'Genesis Project'. May be part of the subversive international cell known as "command"

Figure 5: A small part of the table of characters of *A Million Penguins*. Copyright does not apply.

The term "novel" does not seem apt to describe the text, not even when reducing the focus of analytical attention to the structure and content of these pages only. Setting, plot, and characters are so chaotic that they cannot be summarized appropriately. After the closing of the project, the first section of *A Million Penguins* tells the story of the two fictitious authors Tony and Jim, a walrus, Artie (a humpback whale), Fred and Larry, who are mice, and Schrödinger's cat, also called Fluffy, which is both dead and alive at the same time but does not belong to Schrödinger at all but to Mark. There are also Gina, Mary, James, George, Chad, Bella, a couple of penguins, and the minister for environment, fisheries and customs, Sahra Wagenknecht. The diverse settings, for example, an internet café or the ocean, are as confounding as the list of at least 64 characters (see Figure 5). With the exception of a number of rather secluded parts of the text, the project as such does not have a coherent plot. These parts stand out insofar as the traversal function through them uses a personal perspective since they follow a structure which lets readers choose their own adventures (Choose Your Own Adventures or CYOA). In the majority of cases, they were created by a user of the project called Nicholasjh and maintained and edited by the user Nostrum19. Indeed, these CYOA seem to be especially valuable to some readers. The novelist Kate Pullinger, a member of the project team, even goes so far as to call one of them, a feminist science fiction CYOA, "brilliant" (Mason and Thomas 7).

As with *She...*, the first problem for archiving efforts is to define the boundaries of the text. This is a highly problematic task for *A Million Penguins*, especially the distinction between text and paratext. One way of approaching this challenge of classification is by looking closer at the reader's immersion into the text. Hilary P. Dannenberg discusses immersion and suspense in conjunction with the idea of possible fictional worlds which are enacted by the reader in the reading process (36). If the reading process is immersive, the reader transcends his or her ontological boundaries and embarks on a mental journey into the fictional world of the literary text: "A narrative world is thus a mental construction that provides a place of escape and liberty for the human mind to move to beyond the framework of its true ontological spatiotemporal boundaries. The reading of fiction can thus be described as the cognitive simulation of ontological liberation" (21). A prerequisite of this ontological liberation within the immersion process is a belief in the narrative world (Dannenberg 20-21), which explains why "anti-illusionist fictions have never succeeded in displacing immersive fictions; experimental postmodernist texts have provided more excitement for narrative theorists than for the general reader, who can, though, appreciate the thrill of a little 'soft metafiction'" (Dannenberg 22). If *A Million Penguin* is immersive in this sense—and the time users spent browsing and editing the text is an indicator of this—suspense is directed not to the textons on the main pages of the project but rather to the edition process. The reader wants to know what is going to happen next—not in the novel, but within the process of trying to create it. Additionally, if Dannenberg's model of the translocation of the reader into the fictional world is correct, one would have to assume that readers would not understand the characters as fellow humans on the same ontological level but rather as rooted within the fictional or semi-fictional world of the creation of the Wiki.

If the reader's thirst for knowledge was directed towards the editing process, as claimed, this would have two consequences. First, the suspense would fade entirely after the end of the editing process as the text is not being modified any longer. Critical comments on the quality of the text, written after the end of the active project, should, therefore, be devastating. As the official report on the project enumerates, this is, indeed, the case (Mason and Thomas 1). Jordan Jack went so far as to claim that *A Million Penguins* was "the worst book [he had] ever read" and Douglas Rushkoff stated that "*A Million Penguins* looks like fun, but it's still likely to remain more a million penguins than one cohesive or coherent bird" before asserting that a novel definitely needed an author (Jack; Rushkoff qtd. in Mason and Thomas 1). Second, instead of the produced final versions, the genesis of them would entail the readers' experience of suspense. As this is indeed the case, it is plausible that the inherent core of *A Million Penguins* is a dynamic narration about the collaborative writing of a novel and not the result of this effort—the novel. As the "real" text is a dynamic, that is, on-going, narrative performance by

authors or characters about their effort to write a novel, this text cannot be archived. If there ever was the dream of eventually printing the finished novel—as is expressed in some of the readers' comments on the weblog of Penguin and in the project itself—it would have to be discarded after the end of the project for many reasons: First, it would not be commercially exploitable as the situation of the copyright of the texts is unclear. Second, the final, static, and linear text on the main pages is too confusing and lacking structure and thus not attracting a larger reading audience. Third, there is no final linear and complete text as the project has a number of sidelines, for instance the feminist science-fiction CYOA, whose node-and-link structure cannot be properly represented on paper. Fifth and finally, the real text stages its characters like the characters in a film. Printing the text means taking a snapshot and, thereby, disposing of the very core of the text.

Such problems of archiving texts are not exclusive to digital literature. Another field is performance literature. As Christopher Grobe puts it in his "The Breath of the Poem: Confessional Print / Performance circa 1959," "performance practices are notoriously difficult to study. They leave all sorts of evidence in their wake—plans and prompts precede performance; archival and memorial traces trail it—but none of these simply *is* the object of study in the way that the literary text is often presumed to be" (216; emphasis original). Although the only life of performance is, according to Grobe, in the present and although it cannot be "saved, recorded, documented" (216), there is a turn to the archive, to something that can be analyzed thoroughly if there were an appropriate set of vocabulary to describe it. Peggy Phelan is convinced that, "performance is now the dominant language of contemporary life. And yet it is also true that performance is irreducibly expressive, a way of saying. If we lose the intimacy of the connection between literature and performance, we diminish something vital in and between them. Nor is it enough to consider literature as performance, although that task is always welcome. What we need is *and*: close readings of performances and poems, more muscular math for calculating oversound, the thing not in the words, not in the melody, not in the dance, not in the meter" (946; emphasis original). For digital literature like the two texts discussed, this "thing not in the words" is the social text, the interaction, and the performance. However, in order to be allowed to analyze such texts, there has to be some kind of archive, although it may be incapable of catching the performance and interaction aspect. At the moment, attempts at preserving texts seem to be neglected by publishers. *A Million Penguins* is a striking example as it is a publication of the second largest publisher in the world. It was created in collaboration with a university and stirred discussions about the future of publishing and literary collaboration worldwide. However, at the moment, it is inaccessible in its original form as a Wiki. Publishers and editors do not take the responsibility of conserving the text for future generations. The original internet domain <<http://www.amillionpenguins.com>> is now registered to Rethink Networks, an internet marketing company from Pennsylvania that displays advertisements and links to merchants of votive candles and marijuana seeds on the website, that, ignorant of the impact and the structure of *A Million Penguins* now bears the greeting "Welcome to the A Million Penguins Wikipedia."

Having looked at the examples of Turner's *She...* and Penguin's collaborative prose effort *A Million Penguins* it is obvious that the very ontology of these texts is bound to elements of performance, to direct social interaction of writers and readers, to the uniquely subjective reading process, to the introduction of the readerly technological setup (e.g., browser aesthetics) and to the real-time access to data. How could they ever be archived? If one wants to read *A Million Penguins* today, one has to engage in some sort of textual archaeology, browsing through incomplete snapshots of the media text on the servers of the Internet Archive. This can be a rewarding endeavor. However, it is the reward of the digger, which does not resemble the original reception experience. Using the archive, one can reconstruct the typical battle (see Simanowski 27-34) of the characters (the contributing users) over the grounds of the text which they are in the process of creating collaboratively. This is the first interim solution: Regular snapshots of such media. The problem with this approach, however, lies in the extensive accessibility of material, the capacity of storing devices and the short-term and long-term ac-

cessibility of the archive. Videos on YouTube, user-adapted content, and advertisements, they all belong to the very core of the literary texts in question. There is no feasible way of capturing those features of the text. A workaround is to archive everything published on the world wide web and to make the data available for given points in time—an approach followed by the Internet Archive and one which internet pioneer Ted Nelson had incorporated already into his brainchild and docuverse Project Xanadu whose release has been postponed over and over again. However, how much hard disk space is needed to delta-save the internet regularly, and does this not pose the question of Borges's famous map that becomes so detailed that it eventually covers the landscape it depicts? Even if one could capture all this data, tagged with and sorted according to the time of the snapshot, how could one access a vastly growing number of mirrors of the whole World Wide Web? What technological setup would be necessary to project the menagerie of formats used in such archives?

Aside from static snapshots, the regular staging of important texts could be another solution. One could, for instance, use the log files on the server of the history pages of *A Million Penguins* to create a dynamic copy which develops as the original project had developed. Notwithstanding the missing dedication of the original publishers, such an approach would still fall short of the involvement of every reader in the creation process and, therefore, lack a foundational attraction of the text. In order to retain the intrinsic spirit, the original quality of performative collaborative texts, one could offer new instances of such texts by archiving and re-enacting the design of a given work, either as dynamic copy of the textual system or by archiving the very idea of the text in the form of a project description or a report. This approach resembles Rothenberg's idea of the envelope, as described above, but with the media object itself being void, awaiting user input and interaction. Digital literature and drama share the structural component of the transformation process: In drama, the written script is transformed into a multimedial performance on stage. In digital literature, the textons are projected onto the screen of the respective reader. Using the proposed envelope scheme for the archiving of theatrical performances would entail an inclusion of the setup of the theatre company and the architectonic features of the original theatre in the form of descriptive metadata. Indeed, this is done in scholarly editions of Shakespeare, in which the Elizabethan stage and the specifics of the theatre company are discussed in the paratext. Now and then, theaters like the Globe in London stage Shakespeare's plays in the original pronunciation and the original setup of stage, costumes, props, and actors. Something structurally similar should be possible for works of digital literature in order to enable people to experience a work in its original form. After all, the point of an archive is not to put things into a shelf and forget about them.

One could employ Maurice Halbwachs's idea of collective experience, Jan Assmann's concepts of cultural and communicative memory, or even Edmund Spenser's metaphor of memoria from *The Fairy Queene* with its differentiation of the active and the passive principle—of memory and remembrance (see Zimmermann, "Diverging" 151-54). But in the realm of digital writing, it might be more appropriate to quote Jaron Lanier, who claims that "information is alienated experience. You can think of culturally decodable information as a potential form of experience, very much as you can think of a brick resting on a ledge as storing potential energy. When the brick is prodded to fall, the energy is revealed. That is only possible because it was lifted into place at some point in the past. In the same way, stored information might cause experience to be revealed if it is prodded in the right way ... But if the bits [saved on a hard disk] can potentially mean something to someone, they can only do so if they are experienced. When that happens, a commonality of culture is enacted between the storer and the retriever of the bits. Experience is the only process that can de-alienate information" (28-29). The examples of the two pieces of digital literature shall suffice to show that we still have to find ways of putting the brick onto the ledge to allow future generations to prod it into their own experience. The disappearance of important works of digital literature from the web shortly after their initial publication is an indicator for the urgency to devise strategies and tools for the archiving of such texts.

Works Cited

- Aarseth, Espen J. *Cybertext: Perspectives on Ergodic Literature*. Baltimore: The Johns Hopkins UP, 1997.
- A Million Penguins*. Harmondsworth: Penguin, 2006 <<http://www.amillionpenguins.com>>.
- Assmann, Jan. "Collective Memory and Cultural Identity." *New German Critique* 65 (1995): 125-33.
- Dannenberg, Hilary P. *Coincidence and Counterfactuality: Plotting Time and Space in Narrative Fiction*. Lincoln: U of Nebraska P, 2008.
- Edmond, Jacob. "Archive of the Now." *The 2014-2015 Report on the State of the Discipline of Comparative Literature*. Ed. ACLA: American Comparative Literature Association (2014): <<http://stateofthediscipline.acla.org/entry/archive-now>>.
- Grobe, Christopher. "The Breath of the Poem: Confessional Print/Performance circa 1959." *PMLA: Publications of the Modern Language Association of America* 127.2 (2012): 215-30.
- Halbwachs, Maurice. *Collective Memory*. New York: Harper, 1980.
- Hedstrom, Margaret. "Digital Preservation: A Time Bomb for Digital Libraries." *uky.edu* (1995): <<http://www.uky.edu/~kiernan/DL/hedstrom.html>>.
- Landow, George P. *Hypertext 3.0*. Baltimore: The John Hopkins UP, 2006.
- Lanier, Jaron. *You Are Not a Gadget: A Manifesto*. New York: Knopf, 2010.
- Mason, Bruce, and Sue Thomas. *A Million Penguins Research Report*. Leicester: Institute of Creative Technologies, De Montfort U, 2008.
- Phelan, Peggy. "'Just Want to Say': Performance and Literature, Jackson and Poirier." *PMLA: Publications of the Modern Language Association of America* 125.4 (2010): 942-47.
- Rothenberg, Jeff. *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation*. Washington: Council on Library and Information Resources, 1998.
- Rothenberg, Jeff. "Ensuring the Longevity of Digital Information." *Scientific American* 272.1 (1995): 42-47.
- Simanowski, Roberto. *Interfictions. Vom Schreiben im Netz*. Frankfurt: Suhrkamp, 2002.
- Turner, Renée. *She... Fudge the Facts* (2008): <http://www.fudgethefacts.com/she/_she.html>.
- Zimmermann, Heiko. "Diverging Strategies of Remembrance in Traditional and Web-2.0 On-Line Projects." *Cybercultures: Mediations of Community, Culture, Politics*. Ed. Harris Breslow and Aris Mousoutzanis. Amsterdam: Rodopi, 2012. 152-63.
- Zimmermann, Heiko. *Rekonfigurationen des textuellen Handlungsraums digitaler Literatur unter besonderer Berücksichtigung der Autorschaft. Geschichte, theoretische Ansprüche und deren Wechselwirkungen in der digitalen Medienpraxis*. Doctoral Diss. Trier: University of Trier, 2012.

Author's profile: Heiko Zimmermann teaches English literature at the University of Trier. His interests in scholarship include E.M. Forster's work, memory, remembrance on Web 2.0, and hypertext theory. Zimmermann's recent publications include the articles "Diverging Strategies of Remembrance in Traditional and Web-2.0 On-Line Projects," *Cybercultures: Mediations of Community, Culture, Politics* (Ed. Harris Breslow and Aris Mousoutzanis, 2012) and "Teaching E.M. Forster in 2010," *New Aspects of E.M. Forster* (Ed. Krzysztof Fardoński, 2010) and the forthcoming monograph *Autorschaft und Digitale Literatur* (2015). E-mail: <heiko.zimmermann@uni-trier.de>