

# Langzeiterhaltung von Dokumenten aus dem literarischen Schreibprozess

Julia Büchel, Rolf Jufer, Urs Richle, Alexander Wenzel

Trotz der tief greifenden Umwälzung der Schriftkultur durch digitale Textverarbeitung zeigen fast alle Computerbenutzer ein Schreibverhalten, das sich noch am überholten Modell der Schreibmaschine orientiert. Dies gilt insbesondere für Schriftstellerinnen und Schriftsteller. Deren digitale Textzeugen und elektronische Korrespondenzen drohen dabei verloren zu gehen (durch Formatinkompatibilität, Migrationen, Datenverluste usw.). Das Schweizerische Literaturinstitut der HKB und das Departement Technik und Informatik der BFH untersuchen in enger Zusammenarbeit diese Problematik, um ausgehend von einigen Fallstudien eine technische Infrastrukturlösung zu konzipieren und zu realisieren. Sie soll es professionellen Autorinnen und Autoren erlauben, Dokumente aus ihrem Arbeitsprozess langfristig zu archivieren. Diese können so später auch vom Schweizerischen Literaturarchiv übernommen und der Wissenschaft zugänglich gemacht werden.



**Julia Büchel**  
Schweizerisches Literaturinstitut  
HKB BFH  
juliabuechel@yahoo.de



**Prof. Rolf Jufer**  
Departement Technik und Informatik BFH  
rolf.jufer@bfh.ch



**Urs Richle**  
Schweizerisches Literaturinstitut  
HKB BFH  
urs.richle@bfh.ch



**Alexander Wenzel**  
Schweizerisches Literaturinstitut  
HKB BFH  
alexander.wenzel@hkb.bfh.ch

## Von analogen zu binären Aufschreibesystemen

Die Schrift ist eine der ältesten Kulturtechniken der Menschheit. «Dort, wo keine Schrift verwendet wird, glaubt der Europäer «Primitivität» zu erkennen.»<sup>1</sup> Bei genauerer Betrachtung muss jedoch festgestellt werden, dass die Schrift keine Grundbedingung für das Funktionieren einer Gesellschaft darstellt. Bis heute gibt es Gesellschaften, die orale Erzähltechniken und Riten für die Wissensvermittlung und für die Konstruktion eines kollektiven Gedächtnisses benützen. Und gerade die Existenz eines solchen kollektiven kulturellen Gedächtnisses spricht schliesslich gegen die vermeintliche «Primitivität» einer Gesellschaft. In unseren Kulturkreisen bildet die Schrift, indem sie Wissen festhält und von einer Generation zur nächsten weiterreicht und verfügbar macht, einen grossen Teil der kulturellen Identität. Die Tätigkeit des Schreibens kann in diesem Sinn also immer auch als Arbeit an der kulturellen Identität der Gesellschaft, in der wir leben, verstanden werden. Dies gilt ganz besonders für das literarische Schreiben, das immer auch eine reflexive Funktion erfüllt.

In den vergangenen 50 Jahren haben sich die neuen Informations- und Kommunikationstechnologien unmerklich in beinahe alle Handlungen des Alltags eingeschlichen, so auch in das literarische Schreiben. Seit den 1970er-Jahren hat sich neben der primären eine sekundäre Schriftlichkeit entwickelt. Während die primäre Schriftlichkeit den Schriftbesitz seit rund 7000 Jahren in einem ständig wachsenden Bevölkerungsteil förderte, macht die sekundäre Schriftlichkeit eine spezialisierte Informationstechnologie für diejenigen nutzbar, die bereits in einer alphabetisierten Gesellschaft leben.<sup>2</sup> Dass diese neue Technologie das altbewährte Alphabet jedoch in eine Reihe von Kombinationen der Zahlen 0 und 1, also in ein binäres System, verwandelt hat, dessen sind sich die wenigsten Schreibenden bewusst. Viele Schriftstellerinnen und Schriftsteller benützen ihre Schreibprogramme nach wie vor wie eine alte, mechanische Schreibmaschine. Dem steten Wechsel der Werkzeuge und Dokumentenformate sind sie oft hilflos ausgesetzt.<sup>3</sup> Korrespondenzen mit Verlegern und Schriftstellerkollektiven und -kollegen, Recherchedoku-

mente und Versionen von Werken gehen verloren, weil den neuen Technologien oft blind vertraut wird. Kenntnisse, geeignete Werkzeuge und Infrastrukturen für die langfristige Archivierung fehlen. Mit dem Verlust solcher Dokumente geht auch ein Teil des kulturellen Erbes und damit der kulturellen Identität unserer Gesellschaft verloren. Der Erhalt dieses Erbes soll jedoch gewährleistet werden. Bis anhin existiert in der Schweiz keine institutionalisierte Infrastruktur, die es Autorinnen und Autoren ermöglichen würde, Dokumente und Korrespondenzen aus dem literarischen Schaffensprozess langfristig für ihre eigene Arbeit oder für kulturelle und wissenschaftliche Zwecke zu archivieren. Aus diesem Grund erforscht das Projekt «Digitale Literatur» der BFH das Problem der Archivierung literarischer Dateien und entwirft einen konkreten Vorschlag zur Verbesserung der Situation.

## Digitale Archivierung heute

Mit dem Problem des Verlustes von Dokumenten und Korrespondenzen im Zuge der Wandlungen der Informations- und Kommunikationstechnologien stehen die Schriftstellerinnen und Schriftsteller nicht allein. Seit fast zwei Jahrzehnten setzen sich Archivarinnen und Archivare, Bibliotheken und andere Institutionen mit dem Thema der digitalen Archivierung auseinander. Das Bundesamt für Kultur (BAK) stellt in seinem Bericht zur Mopolitik des Bundes<sup>4</sup> fest, dass in der Schweiz ein Defizit an Infrastrukturen zur Erhaltung des «digitalen Gedächtnisses» des Landes besteht. Noch immer haben nicht alle Institutionen den Weg ins digitale Zeitalter geschafft. Mangelnde Kenntnisse, Überforderung mit neuen Technologien und der Anzahl an Dokumenten, unterschiedliche Handhabung des vorhandenen Materials und selbst Uneinigkeit über Sprachgrenzen hinweg führen zu Verzögerungen in der Digitalisierung.

## Das Projekt «Digitale Literatur»

Ganz im Sinne der Mopolitik des BAK erforschen die Projektpartner\* in enger Zusammenarbeit den Einfluss der neuen Informations- und Kommunikationstechnologien auf den literarischen Schaffensprozess mit dem Ziel, eine den Bedürfnissen der Schreibenden angemessene

Lösung zur Verwaltung und Sicherung von Dokumenten und Korrespondenzen aus ihrem Schaffen zu entwickeln. Damit soll das Projekt «Digitale Literatur» einen Beitrag zur Erhaltung des kulturellen Gedächtnisses unseres Landes leisten.

### Resultate einer Umfrage im Rahmen des Projekts

Aus einer Befragung von neun Autorinnen und Autoren (mit qualitativem Fragebogen, halb strukturierten Interviews und Hausbesuchen) zu ihrer schriftstellerischen Arbeit unter den Bedingungen der heutigen technischen Mittel konnten wir Rückschlüsse für eine geeignete Infrastruktur und die Entwicklung von konkreten Werkzeugen ziehen. Während sich demnach die meisten befragten Schriftstellerinnen und Schriftsteller imstande sehen, mehr oder weniger selbstständig mit dem Computer umzugehen, treten doch bei allen eine Reihe von strukturellen Problemen auf, die unter folgende Kategorien fallen:

- Fortlaufender Wechsel der Dokumentenformate durch neue Produkte; die Schreibenden unternehmen meist keine besonderen Schritte, um die Kompatibilität ihrer Dokumente auf Dauer zu gewährleisten (z.B. durch Verwendung standardisierter Formate, Konversion in aktuelle Formate).
- Fortlaufende Überschreibungen von Dokumenten, anstatt Versionierung zur Sicherung der Textentstehung, was zur Verwässerung der Textgenese führt.
- Fortlaufende Vernichtung von Textversionen und Korrespondenzen durch Kurzlebigkeit der Hardware; E-Mails aus der literarischen Korrespondenz werden zumeist auch nicht gesondert abgelegt und gesichert.

Daneben sind weitere kritische Momente des Informationsverlustes hervorgetreten: Beim Entwurf und bei der Korrektur literarischer Texte kann ein – zum Teil iterierter – Wechsel zwischen den Medien Computer und Papier dazu führen, dass literarische Arbeitsschritte nicht explizit in elektronischer Form und zum Teil nicht einmal als eigenständige Dateiversion dokumentiert werden; am deutlichsten wird dies bei der Vorbereitung der Texte zum Druck, bei der zum Beispiel Druckfahnen und spezifische Dateiformate in den Vordergrund treten, sodass die Schriftstellerinnen und Schriftsteller selbst meist keine «Datei letzter Hand» besitzen, die die Druckversion ihrer Texte exakt widerspiegelt.

Hinsichtlich der Einstellung und des Problembewusstseins gegenüber der Archivierung elektronischer Dokumente unternehmen die meisten der befragten

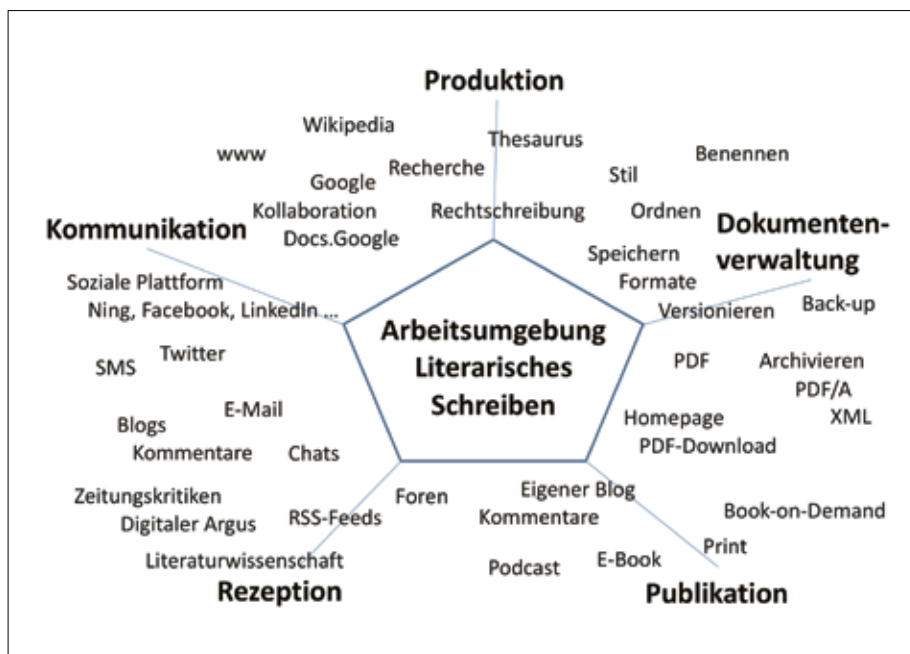


Abbildung 1: Die fünf Bereiche einer digitalen Arbeitsumgebung für das literarische Schreiben

Autorinnen und Autoren denn auch keine besonderen Schritte, um eine langfristige Zugänglichkeit und Kompatibilität ihrer digitalen Texte zu gewährleisten. Die Mehrzahl archiviert ihre Dokumente verteilt auf mehrere Medien (IT/Papier) oder Formate (doc/PDF/Mail), jedoch nicht zwangsläufig redundant. Die meisten Schreibenden haben bei langfristiger Archivierung vor allem den eigenen Zugang zu ihren Texten im Blick, nicht jedoch eine Sekundärnutzung (z.B. durch Archive oder die Literaturwissenschaft), der sie zum Teil skeptisch gegenüberstehen.

### Schlussfolgerungen

Diese Feststellungen und vor allem die Einstellung der meisten Schriftstellerinnen und Schriftsteller zur Archivierungsproblematik haben dazu geführt, dass wir unser Projekt stärker auf deren Nutzen ausrichten haben als geplant. Wir mussten feststellen, dass das Grundproblem nicht in erster Linie die langfristige Archivierung von Dokumenten und Korrespondenzen ist, sondern die adäquate und technikbewusste Verwaltung derselben. Wenn die Autorinnen und Autoren Werkzeuge erhalten, mit denen sie Dokumente kurz- und mittelfristig den technischen Mitteln gemäss verwalten können, dann ist auch die langfristige Archivierung besser gesichert. Eine Archivlösung für literarische Dokumente sollte also als Erstes den Schreibenden selbst dienen. Die Umfrage macht es uns nun möglich, Rückschlüsse auf eine geeignete Infrastruktur zu ziehen und passende Werkzeuge vorzuschlagen.

### Gesamtvision

Anhand der Fallstudien über die Arbeitsweisen von neun professionellen Schriftstellerinnen und Schriftstellern konnten neben der Hauptfunktion der Textproduktion vier weitere Aspekte des heutigen Schreibens und des literarischen Schreibens im Besonderen ans Licht gebracht werden: Kommunikation, Publikation, Rezeption und Dokumentenverwaltung.

Die Vision eines umfassenden Schreibwerkzeugs für Schriftstellerinnen und Schriftsteller integriert alle fünf genannten Aspekte. Was heute mit mehreren Tools und womöglich auf verschiedenen Rechnern erledigt wird, könnte in Zukunft durch ein solches integriertes Schreibwerkzeug zentral geleistet werden.

Im Rahmen dieses Forschungsprojektes behandeln wir nur den Aspekt der «Dokumentenverwaltung».

### Strukturelle Spezifikationen

Wie bei anderen Archivierungslösungen wird die Beschreibung und Indizierung der digitalen Dokumente mit Metadaten geleistet. «Technische» Metadaten, die die elektronischen Dateien als Archivalien beschreiben, können in der Regel automatisch erzeugt werden. Da sie meist nichts über den spezifisch literarischen Inhalt der Dateien aussagen, sind weitere Metadatenkategorien notwendig, die einer intuitiven Nutzung durch Autorinnen und Autoren entgegenkommen oder der literaturwissenschaftlichen Erschließung dienen. Da solche Metadaten zumeist manuelle Pflege erfordern, muss ein Kompromiss zwischen Beschreibungsgrad und Erfassungsaufwand gefunden werden.

Gängige Metadaten konnten aus einschlägigen Standards übernommen werden wie ISAD(G)<sup>5</sup>, ferner ISAAR(CPF)<sup>6</sup> und Dublin Core<sup>7</sup>, die allerdings nur einen allgemeinen Rahmen der Beschreibung von Archivalien bilden, in dem die Metadaten für unsere konkrete Anwendung noch angepasst sowie ergänzt werden mussten, um daraus zweckmässige und aussagekräftige Deskriptoren zu gewinnen. Aus den neun Fallstudien konnten insbesondere spezifische Metadaten erschlossen werden, die weder in den erwähnten Standards noch anderen Datenbanken vorkommen, zum Beispiel vorläufiger und endgültiger Titel von Text oder Werk versus Dateiname oder Entstehungszeitraum des Werks versus Speicherdatum der Datei. Also auch diesbezüglich betritt dieses Projekt Neuland. Spätere Verweise auf entsprechende Einträge zu Schriftstellerinnen und Schriftstellern sowie Werkeinträge im Online-Katalog HELVETICAT der Nationalbibliothek oder in der Archivdatenbank des Schweizerischen Literaturarchivs wären sinnvoll.

### **Technische Spezifizierungen Sicherstellung der Langzeitarchivierung elektronischer Dokumente**

Auch im Projekt «Digitale Literatur» stellt sich die Frage, in welchem Format ein elektronisches Dokument gespeichert werden kann, sodass sein heutiges visuelles Erscheinungsbild und sein Inhalt auch in 10, 30 oder 50 Jahren mit den dann verfügbaren Werkzeugen und Systemen reproduziert werden können. Das von der Standardisierungsorganisation ISO normierte PDF/A<sup>8</sup> gewinnt zunehmend an Bedeutung: Einerseits benötigt das PDF/A-Format typischerweise sehr viel weniger Speicherplatz als andere Formate, und andererseits ist es das erste standardisierte Format, das explizit für die Langzeitarchivierung geschaffen wurde. Es bietet zudem die Möglichkeiten, neben dem reinen Dokumentinhalt auch Metastrukturen und Objektinformationen abzuspeichern; eine Option, die eine rein grafische Darstellung wie zum Beispiel ein TIFF-Rasterbild nicht bietet.

PDF («Portable Document Format») an sich garantiert keine Langzeit-Reproduzierbarkeit, nicht einmal das Prinzip WYSIWYG (what you see is what you get). Damit beides gewährleistet ist, mussten gewisse Einschränkungen und Erweiterungen gegenüber dem existierenden PDF definiert werden. Für die Reproduzierbarkeit ist beispielsweise wesentlich, dass alle notwendigen Informationen im Dokument selbst enthalten sind. Dies umfasst sichtbaren Inhalt wie Texte, Vektorgrafiken, Rasterbilder, Schriftarten, Farbräume

und vieles mehr. Der so entstandene PDF/A-Standard stellt sicher, dass PDF/A-konforme elektronische Dokumente mit «einfachen» Mitteln analysiert und interpretiert werden können, im einfachsten Fall sogar mithilfe eines simplen Texteditors. Somit ist weitgehend sichergestellt, dass PDF/A-Dokumente auch in den oben genannten Zeiträumen noch mit dem zum Entstehungszeitpunkt vorhandenen Erscheinungsbild reproduziert werden können.

### **Open-Source-Lösung Jackrabbit**

Für die technische Realisierung eines Prototyps haben wir uns für eine Client-Server-Architektur entschlossen. Die Dokumente sollen von den Schriftstellerinnen und Schriftstellern von ihrer privaten Schreibumgebung aus auf einem zentralen Server verwaltet werden. Über ein Login können sich die Benutzenden auf eine Art gesichertes Postfach begeben und darin ihre Dokumente und Korrespondenzen abspeichern, versionieren und archivieren. Für die Entwicklung der Anwendung stützen wir uns auf ein Open-Source-Projekt<sup>\*\*</sup>, welches ein multifunktionales Dokumentenverwaltungssystem bietet, das einige der Grundfunktionen, die wir benötigen, bereits integriert. Im Rahmen unseres Forschungsprojektes erweitern wir diese Open-Source-Lösung im Bereich der Zugriffsrechteverwaltung, der Metadaten und der Workflows (Archivierungs- und Vererbungsverlauf), um auf die spezifischen Bedürfnisse, die wir über die neun Fallstudien definieren konnten, einzugehen.

### **Nutzen und Perspektiven**

Die zukünftigen Benutzer der entwickelten Lösung des Projektes «Digitale Literatur» werden in erster Linie die Autorinnen und Autoren selbst sein. Wenn es uns gelingt, den Erstnutzenden des Projektes eine professionelle Infrastruktur zu bieten, die die fachgerechte Verwaltung der Dokumente aus dem literarischen Schaffensprozess nicht nur ermöglicht, sondern im Idealfall auch fördert, dann sind die Grundvoraussetzungen für eine Langzeitarchivierung gegeben. Die Literaturwissenschaft, Bibliotheken, Archive und Literaturinteressierte werden dadurch auch in Zukunft die Genese eines Werkes, das in der Epoche der digitalen Schreibsysteme entstanden ist, erkunden und erforschen können. Das Schweizerische Literaturarchiv der Nationalbibliothek und der Verband Autorinnen und Autoren der Schweiz stehen uns deshalb für die Realisierung eines ersten Prototyps beratend zur Seite. Die grosse Herausforderung wird es später sein, eine geeignete Institution zu finden, die den Server unserer Anwendung

und damit das literarische Erbe von vielen Schweizer Schriftstellerinnen und Schriftstellern langfristig technisch betreut und aufrecht erhält.

#### **Literaturnachweise**

- 1 Haarmann, H.: Geschichte der Schrift. C.H. Beck Wissen, München, 2007. S. 11.
- 2 Haarmann, H.: Geschichte der Schrift. C.H. Beck Wissen, München, 2007. S. 124.
- 3 Über ihre Erlebnisse im Wandel der Schreibwerkzeuge berichten eindrücklich die Autoren Daniel de Roulet (2005) und Jim Porter (2002): De Roulet, Daniel: La Nouvelle volatilité de l'écrit et les moyens d'y remédier. In: Les archives littéraires, Documents n° 7, Association pour le patrimoine naturel et culturel du canton de Vaud, Lausanne, octobre 2005. Porter, Jim: Why technology matters to writing: A cyberwriter's tale. In: Computers and Composition 20 (2002). S. 375–394, vgl. www.sciencedirect.com.
- 4 Wehrli, M.: Memopolitik. Eine Politik des Bundes zu den Gedächtnissen der Schweiz. Bundesamt für Kultur, Bern, 2008. Online verfügbar unter: <http://www.bak.admin.ch>.
- 5 International Standard Archival Description (General), vgl. International Council on Archives: ISAD(G). General International Standard Archival Description. Second Edition, Ottawa 2000; [http://www.ica.org/sites/default/files/isad\\_g\\_2e.pdf](http://www.ica.org/sites/default/files/isad_g_2e.pdf).
- 6 International Standard Archival Authority Record for Corporate Bodies, Persons and Families, vgl. International Council on Archives: ISAAR(CPF). Internationaler Standard für archivistische Normdaten (Körperschaften, Personen, Familien). Zweite Ausgabe, Wien, 2004; <http://www.ica.org/sites/default/files/ISAAR%20GER.pdf>.
- 7 Vgl. Dublin Core Metadata Initiative (DCMI).
- 8 Vgl. PDF/A: An ISO Standard: <http://www.pdfa.org>.

\* Das Schweizerische Literaturinstitut, ein Fachbereich der HKB, und das Departement Technik und Informatik der BFH.

\*\* Jackrabbit von Apache; <http://jackrabbit.apache.org>.