

Schneider René,

Tags: numérisation e-book,

Virtuelle Eselsohren – u?ber das Lesen im digitalen Zeitalter. Bislang sind weder informationeller Mehrwert noch kulturelle Bedrohung zu erkennen

Das Lesen im digitalen Zeitalter ist so wohl kompliziert als auch tu?ckenreich. Der Weg zu Literatur ist dabei aufgrund der unterschiedlichen Strategien der Suchhilfen, der Vielfalt der digitalen Formate und der uneinheitlichen Funktionalita?ten der Lesesoftware genauso umsta?ndlich wie demoralisierend. Eine Bestandsaufnahme verschiedener Digitalisierungsinitiativen zeigt, dass bislang weder ein informationeller Mehrwert noch eine kulturelle Bedrohung zu erkennen sind.

«In omnibus requiem quaesivi et nusquam inveni nisi in een Hoecken met een Boecken», ist auf einem Portra?t Thomas von Kempens zu lesen: «U?berall suchte ich Ruhe und fand sie nirgends, ausser in einer Ecke mit einem Buch». Ein frommer Wunsch im digitalen Zeitalter, dessen Erfu?llung zunehmend schwierig wird. Damit ist nicht einmal die zunehmende Zahl von Ho?rbu?chern gemeint, die die uralte Tradition des Geschichtenerza?hlens wiederbeleben, gemeint sind eher die u?berdurchschnittlich hohen kognitiven Anforderungen und die Strickfallen, denen sich der Leser digitaler Literatur gegenu?bersieht und die jede Form geistiger Ruhe unterbinden.

Im Artikel werden unterschiedliche Digitalisierungsinitiativen bezu?glich der Komplexita?t des Lesevorgangs beziehungsweise der damit einhergehenden Anspru?chen an das Lesen miteinander verglichen. Die in den teils lautstarken Diskussionen um die aktuellen Digitalisierungsinitiativen ausgetauschten rechtlichen und kulturellen Aspekte sollen dabei ausser Acht gelassen werden.

Im Skriptorium

Eine ganze Reihe von Bibliotheksbeständen wird gegenwärtig von unterschiedlichen Interessengruppen nach dem Motto «all digital, all the time» digitalisiert. Bücher und Zeitschriften erhalten somit ein elektronisches Pendant, das in dafür eingerichteten digitalen Bibliotheken zur Verfügung gestellt wird, und sie werden zu E-Books (kurz für electronic book, geläufig sind auch die Schreibweisen E-book, e-book und ebook), Artikel werden zu E-Papers usw.

Handelt es sich dabei um vormals papiergebundene Texte, werden diese entweder (allerdings in den seltensten Fällen und nur bei Texten sehr kurzer Länge) per Hand über die Tastatur inkl. der möglichen Tippfehler eingetippt oder über spezielle Lesegeräte als Pixel in einem speziellen Format gespeichert und über entsprechende Software zur OCR (Optical Character Recognition) – wiederum inkl. der dabei entstehenden Fehler – in ein elektronisches Zeichenformat überführt.

Kulturhistorisch stehen E-Books von daher in der Tradition der Arbeit mittelalterlicher Kopisten und der damit verbundenen schriftlichen Tradition des kulturellen Gedächtnisses. Möglicherweise erklärt dieser Umstand die Vorliebe für mittelalterliche Manuskripte und deren Digitalisierung bzw. deren Hervorhebung in digitalen Bibliotheken und begründet andererseits die stillschweigende Akzeptanz der doch beträchtlichen Fehler, die mit jedem Kopiervorgang einhergehen, sowie der daraus resultierenden semantischen Instabilität der Dokumente.

Der Wildwuchs der Formate

Betrachtet man die Kopierwut noch etwas genauer, ist man sogar versucht, die Analogie zu den Kopisten der mittelalterlichen Skriptorien soweit zu treiben, dass passend zu der Handvoll existierender Kopien, die von einem Original erstellt und aufbewahrt wurden, für jedes Werk auch im digitalen Zeitalter eine Handvoll Kopien, allerdings in unterschiedlichen Formaten erstellt werden. Man trifft auf den im Softwarebereich üblichen Wildwuchs, der durch eine Vielzahl kryptischer Akronyme wie etwa .txt, .html, PDF, Flip Book, DjVu, .tiff, .jpg, .png gekennzeichnet ist, die in Abhängigkeit zur verwendeten Hard- und Software stehen. Die Frage, ob sich hier jemals ein wirklicher Standard durchsetzen wird, erscheint müßig.

Allerdings haben sich sowohl in der Vielzahl der Bildformate als auch in den Leseformaten einige Trends herauskristallisiert. Erfolgreich waren und sind dabei sowohl plattform- als auch displayunabhängige Formate, allen voran PDF (Portable Document Format) der Firma Adobe, das durch die daran gekoppelte Lesesoftware «Acrobat Reader» frei verfügbar ist, sowie als Alternative die Derivate der Auszeichnungssprache XML (extensible MarkupLanguage) – ein vom W3-Konsortium spezifizierter Standard –, der teilweise in die E-Book-Software des Marktführers Microsoft aufgenommen wurde.

Gerade die Aktivitäten um XML stellen einen nicht zu vernachlässigenden Ansatz dar, um endlich einen allgemeinen Standard für die Langzeitarchivierung sowie die notwendige und nachhaltige Trennung von Inhalt und Form zu finden.

Die Einzelheiten über die unterschiedlichen Formate sind in einschlägigen Onlineenzyklopädien nachzulesen. Anbietern und Nutzern digitaler Bibliotheken sollte allerdings klar sein, dass – solange kein einheitlicher und sowohl plattform- als auch browserunabhängiger Standard existiert – ein beiderseitiger Erfolg nur dann erzielt werden kann, wenn die digitalisierte Literatur in mehreren Formaten vorliegen wird und Browser oder Off-line-Lesesoftware verwendet wird, die in der Lage sind, mehrere Formate zu verarbeiten.

Vom Umgang mit papiergebundenen und virtuellen Büchern

Vor das Lesen haben die Bibliothekare den Weg zum Buch gesetzt. Dabei ist dieser Weg, der in traditionellen Bibliotheken über den Katalog oder das Durchstöbern der freizugänglichen Regale zum Buch führte, in der digitalen Bibliothek zum Browsen entlang vorgegebener Hierarchien geworden oder findet in der virtuellen Katalogsuche über ein suchmaschinenaähnliches Eingabefeld seine Entsprechung.

Beide Vorgänge stehen im engen Zusammenhang zu den Such- und Findmechanismen des Internets und werfen einige berechtigte Fragen auf: Können Internetbenutzer, insbesondere Anfänger ohne größere Erfahrungen, die gewünschte Information über die Suchmasken finden? Was geschieht, wenn die Katalogsuche (in den Metadaten der Bücher wohlgeordnet) aufgrund des Layouts der Suchmaske als Volltextsuche über eine herkömmliche Internetsuchmaschine (miss-)verstanden wird?

Anders gefragt: Werden digitale Bücher im Internet zu vergleichbarer Information wie jede andere auch oder bleibt ihnen der informationelle Mehrwert papiergebundener Bücher erhalten, der sowohl aus ihrer materiellen Machart als auch aus ihrer charakteristischen Form resultiert?

Treffend beschrieb diesen Zusammenhang unlängst der amerikanische Schriftsteller John Updike¹:

«Books traditionally have edges: some are rough-cut, some are smooth-cut, and a few, at least at my extravagant publishing house, are even top-stained. In the electronic anthill, where are the edges? The book revolution, which, from the Renaissance on, taught men and women to cherish and cultivate their individuality, threatens to end in a sparkling cloud of snippets.»

Es geht von daher auch um die Frage, inwieweit das Web die Besonderheiten des Kulturguts Buch in seiner digitalen Form respektiert und virtuell realisiert oder Bücher lediglich als Objekte bzw. Endknoten in seinen hypertextuellen Kontext vereinnahmt und darin im Seiten- oder Abschnittformat zerstückt.

Was ist also mit jenen besonderen Handlungen, die das Lesen von Büchern impliziert: die Verwendung von Lesezeichen; Markierungen im Text, etwa Unterstreichungen oder in der Form von Glossen am Textrand; das optische Abschaätzen des gelesenen und noch zu lesenden Texts durch einen Blick auf den Buchrand; das Blättern und das Nachschlagen über einen Index?

Ein Streifzug

Zur Überprüfung dieser Fragen suchen wir in drei unterschiedlichen Bibliotheken nach Literatur, unbefangenermassen nach «Shakespeare», ferner nach seinem Drama «The Tempest» und – im Wissen darum, dass viele Suchmaschinen digitaler Bibliotheken eine Volltextsuche anbieten – nach einem Zitat aus «The Tempest»: «my library was dukedom large enough». Dabei geht es vorrangig um eine kurze Bestandsaufnahme zur Informationssuche in digitalen Bibliotheken, die nicht den Anspruch einer streng wissenschaftlichen Studie hat, sondern dem virtuellen Streifzug eines herkömmlichen Internetnutzers ähnelt, der die Websites führender Digitalisierungsinitiativen besucht [2](#), 11.8.2006 unter Verwendung des Firefox Browsers durchgeführt.

Google Books Search

In der Betaversion (books.google.com) des Suchmaschinenmarktführers (der die Bücher von einer Handvoll renommierter angelsächsischer Bibliotheken einscannt) ist die Suchmaske vertraut, allerdings wird die Suche hier unterschieden zwischen «All books» und «Full view books». Bei Letzteren handelt es sich um jene, für die kein Copyright mehr besteht; die Suchergebnisse werden auch dementsprechend markiert. Alle anderen Bücher sind nur partiell oder (vergleiche das Updike-Zitat) als Schnipsel anzusehen, gleichfalls verlangt Google für den «Full view» eine Anmeldung des Benutzers. Klickt man auf eines der Ergebnisse (die mit den Buchdeckeln aufgelistet werden), erhält man als Resultat direkt die Seite, in der einer oder sämtliche eingegebenen Suchbegriffe vorkommen, es handelt sich demnach um eine integrierte Katalog- und Volltextsuche.

Die Qualität der Suchmaschine ist in jedem Fall beeindruckend. Wir geben die Suchbegriffe einzeln ein und schliesslich alle in einer Reihe, also Autor, Titel und Zitat, und erhalten direkt die relevante Seite des Werks, allerdings nicht als «Full view». Mit den gleichen Suchbegriffen erhalten wir im «Full view» als Resultat *ein und genau ein* Werk als Resultat, das allerdings den umfangreichen und ominösen Titel «The Family Shakspeare (*Sic!*) in Ten Volumes in which Nothing is Added to the Original Text: But Those Words and Expressions are Omitted which cannot with Propriety be Read Aloud in a Family».

Sämtliche gefundenen Suchbegriffe werden direkt im Text farbig markiert. Es besteht zudem die Möglichkeit, *nach* Buchern und *in* den Buchern zu suchen. An der rechten Seite der einseitig wiedergegebenen Buchseiten sieht man die Textpassagen als Fragmente aufgelistet, in denen die Begriffe vorkommen. Ein Kopieren des Seiteninhalts ist allerdings nicht möglich, es handelt sich in allen Fällen um Scans, also Faksimiles der Buchseiten, die im .png-Format wiedergegeben werden.

Das Blättern im Buch ist relativ einfach, es geschieht entweder über einen Mausklick auf der Seite, ein numerisches Eingabefeld oder zwei Pfeiltasten.

Zu bedauern bleibt allerdings, dass sämtliche ernsthaften und werkgetreuen Ausgaben der Werke Shakespeares nur als «Limited preview» verfügbar sind; es handelt sich demnach letztlich nicht um eine digitale Bibliothek, sondern um nichts anderes als Werbung für die auf allen Seiten genannten Verlage und Buchhändler, bei denen die Bücher erhältlich sind. Die digitale Bibliothek verkommt in diesem Fall zum Verkaufsstand.

Open Content Alliance & Internet Archive

Der grösste angelsächsische Gegenspieler Googles auf dem Gebiet der digitalen Bibliotheken firmiert unter dem Namen Open Content Alliance (www.opencontentalliance.org). Dahinter verbergen sich eine ganze Reihe von Teilnehmern, unter denen allerdings Microsoft und Yahoo! herausragen. Eine digitale Bibliothek wird man auf den Internetseiten allerdings noch vergeblich suchen; es sei denn, man folgt dem Wegweiser auf der Hauptseite: «Open Content Alliance is a trademark of the Internet Archive.» Die Seiten dieses Konsortialpartners (www.archive.org) bieten Information sämtlicher Medien in Huile und Fülle. Wir folgen dem Verweis auf textuelle Information und geben wiederum unsere Suchbegriffe ein.

Die Suchmaschine liefert bei Eingabe der Suchbegriffe «Shakespeare The Tempest» vier Ergebnisse, die direkt zum Werk in voller Länge führen und in drei Fällen als einfache .txt-Datei im Webbrowser zu lesen sind. Die einzelnen Dateien können auch komplett über ftp oder http eingesehen, abgespeichert und demzufolge ausgedruckt werden. Eines der Suchergebnisse liegt im PDF-Format vor, nach Download bleibt das gesamte Werk jedoch aus unerfindlichen Gründen unlesbar.

Sieht man von diesem Ärgernis ab, markiert www.archive.org einen echten Kontrapunkt zu Google, allerdings ist die Lesefreundlichkeit der teilweise unhandlich grossen Textdateien, die im vorliegenden Fall aus dem Gutenbergprojekt (www.gutenberg.org) stammen, als eher gering einzuschätzen.

Reichern wir die Suche durch das Zitat aus «The Tempest» an, bleibt die Recherche erfolglos, wir müssen uns auf die im Webbrowser verfügbaren Suchmechanismen stützen, um das gesuchte Zitat – allerdings in der Schreibweise «librarie» – zu finden. Geben wir lediglich den Begriff «Shakespeare» ein, erhalten wir einen unübersichtlichen Berg von Sekundärliteratur. Einige dieser Ergebnisse erlauben neben dem Download als PDF eine Ansicht des Dokuments in den weniger bekannten Datenformaten .DjVu (sprich: De?ja?-vu) oder als FlipBook.

DjVu kann über ein Applet (also eine dynamisch im Webbrowser geladene Software) genutzt werden. Startet man das Applet, zeigt sich DjVu auf den ersten Blick «als ein zweites PDF», allerdings kann hier durch Klicken auf das Text-Icon zwischen Faksimileansicht und reiner Textansicht hin und her gewechselt werden. Als Beispiel sei auf die Shakespeare Bibliographie unter der URL www.archive.org/details/shakes... ³. Wie wertvoll diese Funktionalität ist, zeigt sich, wenn man sich die eingangs vermerkte teilweise mangelhafte Qualität der OCR vor Augen hält.

Das FlipBook (zu Deutsch: das Daumenkino) stellt durchaus eine weitere Alternative dar ⁴, ⁵. Beidseitig ist das Buch auf dem Bildschirm aufgeschlagen und durchforstbar. Jeder Treffer wird durch ein Lesezeichen markiert. Je nach aufgeschlagener Seite ist auch erkennbar, an welchem Teil des Buchs der Leser sich gerade befindet, wenngleich in einigen Fällen die Auflösung für eine wirkliche Lektüre zu gering ist. Auch das Umschlagen der Blätter wird intuitiv nachvollziehbar durch einfaches Klicken auf die Seiten. Bedauerlicherweise sind nur wenige Bücher in diesem Format vorhanden, andererseits handelt es sich häufig um illustrierte Bücher, bei denen der Mehrwert des FlipBook besonders deutlich wird.

The European Library & BnUE

Die europäische Antwort auf die Digitalisierungsinitiative von Google + Co. lautet «The European Library» und ist unter der Webadresse www.theeuropeanlibrary.org zu finden. Dort findet sich auf der Hauptseite neben der Suchmaske immer ein Verweis auf den neuesten digitalisierten «Schatz»: in der Regel die Digitalisierung eines mittelalterlichen Skripts. Die Suche mit der Eingabe «Shakespeare, The Tempest» bleibt allerdings ohne ein einziges Ergebnis. Gibt man lediglich den Begriff «Shakespeare» ein, liefert die Suchmaschine nahezu 400 Ergebnisse, davon die ersten 20 auf Ungarisch. Aus Sichtweise der europäischen Integration sicherlich eine gelungene Sache, für des Ungarischen nicht mächtige Leser ist der Wert der zurückgelieferten Information jedoch vergleichsweise gering.

Schnell wird allerdings deutlich, dass es sich bei der European Library um eine Metasuchmaschine vieler digitaler Bibliotheken handelt. Durch einen Klick auf die Kollektionen der teilnehmenden Bibliotheken lässt sich die Treffermenge einschränken. Hier reduziert sich die Suche jedoch auf eine einfache Katalogrecherche. Auch die Suchmöglichkeit für Fortgeschrittene liefert keine wirklich brauchbaren Ergebnisse.

Bleibt allein die BnUE (die Bibliothèque Numérique de la Union Européenne, www.bnue.org) als Ausweg, deren Daten bislang jedoch hauptsächlich aus französischen Quellen stammen. So liefert die Suche nach «Shakespeare The Tempest» vier Resultate: drei Bücher von Chateaubriand sowie eines von Balzac, aus deren Titel allerdings kein direkter Zusammenhang zu Shakespeares Werk erkennbar ist. Noch irritierender ist der Umstand, dass sämtliche Versuche, die Bücher anzuschauen, mit einer Fehlermeldung enden.

Die Utopie des idealen BookBrowsers

Sämtliche Digitalisierungsinitiativen vertreten unterschiedliche Ansätze und offenbaren sowohl Stärken als auch Schwächen, die sich ohne weiteres zu einem Ideal formen lassen, wenngleich eine konkrete Aussicht auf deren Realisierung vorerst eine Utopie bleiben muss, da die Interessen, die hinter den jeweiligen Digitalisierungsinitiativen stehen, zu unterschiedlich sind.

Wünschenswert erscheint in jedem Fall eine hochwertige Suchmaschinentechnologie wie die von Google, idealerweise ohne die allgegenwärtige Werbung oder die verschnipselte Auflistung des Buchinhalts, allerdings mit den zwei Suchoptionen: eine der Katalogsuche entsprechende Recherche in den Metadaten der Bücher neben einer direkt in den Daten durchgeführten Volltextsuche. Dabei sollten die Treffer nicht nur im Text, sondern auch optisch durch Lesezeichen am Buchrand markiert sein.

Zudem müsste ähnlich wie in DjVu-Formaten ein Umschalten zwischen zweidimensionalem Faksimile und linearem elektronischen Text möglich sein. Wie im Fall der Flip-Books (vorzugsweise mit variabler Seitenauflosung), sollte sowohl das Blättern als auch das Umschlagen von mehreren Seiten über die Maus möglich sein.

Bedauerlicherweise weist keines der gesichteten E-Book-Formate einen elektronischen Index in Form von E-Referenzen auf, der über eine alphabetische oder thematische Übersicht ein Nachschlagen relevanter Textpassagen ermöglicht.

Auch wenn dies ansatzweise u?ber eine Volltextsuche mo?glich ist, zeigt die ta?gliche Praxis doch, wie hilfreich die Lektu?re im Index fu?r die Texterschliessung sein kann. Schliesslich handelt es sich gerade bei den Links um eine der Sta?rken des World Wide Web, so dass ihr Fehlen auch ein Indiz dafu?r ist, wie schwierig die sachgerechte Integration der linearen *und* nicht linearen Aspekte eines Buchs in den Gesamtkontext des Web ist.

Fazit

Solange die vielfa?ltigen Abha?ngigkeiten zwischen Browsern, Plug-ins und Datenformaten bestehen und der Weg zum Buch so ungewollt finitenreich bleibt wie in den oben beschriebenen Fa?llen, stellen digitale Bibliotheken weder Gewinn noch Bedrohung dar.

Sie verdra?ngen absurderweise und entgegen der urspru?nglichen Absicht die Bibliotheken gegenu?ber dem «herko?mmlichen Web» als Informationsvermittler und unterstreichen gleichzeitig die Notwendigkeit des traditionellen Lesens, zumal «Lesen» in digitalen Bibliotheken derzeit nichts anderes ist als ein Wechselspiel aus Suchbegriffen und Textfragmenten.

Ohne Unterstu?tzung durch eine effiziente Katalog- und Volltextsuchmaschine und ohne eine effektive Visualisierung der Buchinhalte, die den Lesevorgang ada?quat unterstu?tzt, werden digitale Bibliotheken weiter den Pegel der Informationsflut steigen lassen und gegenu?ber handlicher Onlineinformation, die sich ohne weiteres ausdrucken und in einer ruhigen Ecke lesen la?sst, ins Hintertreffen geraten.

¹ New York Times, 25.6.2006.

² Die im Artikel beschriebene Suche wurde am

³ verwiesen.

⁴ Ein Beispiel hierzu findet sich unter der URL www.openlibrary.org/details/tr...

⁵ .



René Schneider

Haute Ecole de Gestion de Gene?ve