

Das Projekt Gutenberg und ein Seitenblick auf weitere Digitalisierungs- projekte im europäischen Raum

In Europa sind derzeit zahlreiche Projekte am Laufen, welche auf eine Digitalisierung und Onlineverfügbarkeit von Beständen und Objekten aus der europäischen Kulturgeschichte abzielen. Während sich das deutschsprachige Projekt Gutenberg noch auf Bücher beschränkt, beziehen andere, gross angelegte Projekte wie etwa das von der Europäischen Kommission angeregte «Europeana» bereits weitere Medien (Ton, Film, Bild) mit ein. Ein Überblick.

Die Erfindung der Druckerpresse vor etwas mehr als 600 Jahren löste eine Entwicklung aus, für die der Begriff «Medienrevolution» von heute aus betrachtet durchaus angemessen erscheint^{Cf. Kasten zu Johannes Gutenberg auf Seite XXX.}. Das Internet ist viel jüngeren Datums. Man vergisst das nur allzu leicht. Rechnet man die Aufschaltung des ARPA-Netzes des amerikanischen Verteidigungsministeriums als Geburtsstunde des Internets, so ist das «Netz der Netze» knapp 40 Jahre alt. Die Adaptierung von TCP/IP erfolgte erst 1982., aber die Umwälzungen, die es mit sich gebracht hat (Stichworte Vervielfältigung resp. Vermassung von Informationen / Zugänglichkeit / Vernetzung), können *mutatis mutandis* durchaus mit jenen der Schwarzen Kunst verglichen werden.

Der Name Gutenberg steht Pate für ein gross angelegtes Projekt im deutschsprachigen Raum, das – verschärft und verkürzt gesagt – die klassische deutschsprachige Literatur im Internet für alle kostenlos zugänglich machen will.

Im März 1994 als kleines Projekt gestartet, hat sich das Projekt Gutenberg in der Zwischenzeit zur grössten Volltextsammlung^{Volltext heisst im Idealfall: In den Texten kann nach Textstellen gesucht werden, es können einzelne Textstellen oder das ganze Buch ausgedruckt und zum privaten, nichtkommerziellen Gebrauch weiterverwendet werden. Andere, ähnlich gelagerte Projekte (etwa «Search inside the book» von Amazon oder Googles «Print») stellen digitalisierte Bildseiten ins Netz, die eine Bearbeitung im Volltext nicht erlauben.}klassischer deutschsprachiger Texte im Internet entwickelt. Übrigens war der erste Text, der vom Projekt Gutenberg ins Netz gestellt wurde, Wilhelm Buschs Bildergeschichte von Max und Moritz.

Seit April 2002 kann das Gutenberg-Projekt auf einen starken Partner zählen: SPIEGEL online stellt dem Projekt einen Rechner zur Verfügung, übernimmt die Kosten für die Internetanbindung und leistet technische Unterstützung. Alle inhaltlichen und redaktionellen Arbeiten leistet das Gutenberg-Team weiterhin eigenverantwortlich und kostenlos.

Angebot

Der Kernpunkt der Site sind natürlich die *Texte*. Diese können auf drei Arten angesteuert werden: über ein alphabetisches *Register aller AutorInnen* von A wie *About*, *Edmond* bis Z wie *Zweig*, *Stefan*, über eine *Liste mit verschiedenen Genres* von A wie *Abhandlung* bis W wie *Witzblatt* sowie über eine Textsuche. Sucht man beispielsweise den Text des von Georg Büchner 1835 geschriebenen Stücks «Dantons Tod», gelangt man zunächst auf eine Seite, die eine Kurzbiographie, ein Werkverzeichnis sowie Angaben darüber, welche Texte von Georg Büchner im Gutenberg-Projekt greifbar sind, enthält. Der Text selbst ist in Akte und Szenen unterteilt, navigiert wird vorwärts und/oder rückwärts über Pfeilsymbole. Über ein separates Informationsfenster können Quellenangaben eingeholt werden.

Die *Volltextsuche* ist fehlerbehaftet. Sucht man beispielsweise die Textstelle «*Das Nichts ist der zu gebärende Weltgott*» (IV, 3), werden zwei Alternativen angeboten (Hugo Ball und Georg Büchner), wobei nicht ersichtlich ist, wer der Urheber des Satzes ist und wer ihn zitiert hat. Steuert man die Alternative «Büchner» an, wird man zudem nicht an die richtige Textstelle verwiesen.

Der erste Satz von Max Frischs Roman «*Stiller*» (*Ich bin nicht Stiller!*) bietet 1642 (!) Alternativen, die darüber hinaus nicht alphabetisch geordnet sind.

Mit Goethes Satz (Goethe ist ja nicht gerade niemand in der deutschen Literaturgeschichte ...) «*Man tut nicht gut daran, sich zu lange im Esoterischen auf- zuhalten, das Esoterische trachtet danach, exoterisch zu werden*» wird man an Friedrich Daniel Ernst Schleiermacher und an Max Stirner verwiesen. Für wissenschaftliches Arbeiten ist das Angebot des Gutenberg-Projekts somit unzureichend, um nicht zu sagen unbrauchbar.

Auch der *Ausdruck eines Werkes* kann nicht ohne Hindernisse und Fussangeln bewerkstelligt werden. Im Modus «Druckansicht» kann jeweils nur eine einzige Seite ausgedruckt werden. Für manchen mag da der Gang in eine Bibliothek oder in einen Buchladen einfacher erscheinen – sympathischer sowieso.

Der Vollständigkeit halber sei noch erwähnt, dass auf der Site neben den Texten auch noch weitere Angebote angesteuert werden können, neben weiteren z.B. *Informationen zum Gutenberg-Projekt* oder ein *Link auf einen Shop mit Angeboten des Spiegel-Verlags*.

Unter der Rubrik «Spezial» werden sogenannte «*Handybücher*» angeboten. Das sind Bücher, die auf dem Handy wie eine SMS oder ein Bild gespeichert und ab dem kleinen Bildschirm des Mobiltelefons gelesen werden können. Über den Sinn dieser Einrichtung lässt sich sicher streiten.

Interessant erscheint uns hingegen das Angebot «*Gutenberg-DE Edition 10*». Es handelt sich dabei um 10 DVDs, die das komplette Archiv – Stand September 2007 – des Gutenberg-Projekts enthalten (d.h. gegen 4000 Bücher von über 700 Autoren) und gerade mal € 30.90 kosten.

Weiter im Angebot führt das Projekt Gutenberg eine stattliche Anzahl von *Hörbüchern*¹, darunter etwa die Märchen der Gebrüder Grimm, Gottfried Kellers Novelle «Kleider machen Leute», Oscar Wildes «Bildnis des Dorian Gray» oder Franz Kafkas «Die Verwandlung». Es ist davon auszugehen, dass das Angebot kontinuierlich erweitert wird.

Wie Max und Moritz ins Netz gelangten «Das Projekt Gutenberg-DE ging nach mehrmonatiger Vorbereitungszeit im März 1994 online. Wir haben in alten Archivdateien gesucht und die Logdatei vom 24. März 1994 gefunden, in der die ersten Zugriffe auf das Archiv notiert wurden. Der Server wurde mittags gestartet, um 12.57 Uhr wurde der erste Text abgerufen. Es war die Bildergeschichte Max und Moritz von Wilhelm Busch. In den ersten drei Stunden wurden 10 Dokumente abgerufen. Heute sind es 100 000 in der gleichen Zeit. Im Internet sind 10 Jahre eine Ewigkeit. 1994 gab es nur wenige Nutzer mit privatem Internetanschluss, und die kamen mit Geschwindigkeiten von 2 KB bis 4 KB/Sekunde ins Web, daher mussten wir das Datenvolumen der Seiten klein halten und z.B. die Bildergeschichte von Max und Moritz in Text und Bild aufteilen. Freunde beobachteten uns mit grosser Skepsis und äusserten sich z.B. so: «Wer interessiert sich denn heute noch für alte Bücher, Goethe und so'n Kram, und dann auch noch im WEB, das wird doch nie was, und wie wollt ihr das finanzieren, macht doch lieber was Vernünftiges.» Wir fingen trotzdem voller Optimismus und Begeisterung an: mit einigen Märchen und Bildergeschichten, einem 386-PC, einem selbstgeschriebenen WebServer (WEB- 4HAM, ©: Gunter Hille) und einer analogen Einwahlverbindung, d.h., für jeden Abruf von Gutenberg-Texten mussten wir 18 Pfennig Telefongebühren plus Kosten für das abgerufene Datenvolumen zahlen.»

Quelle: <http://gutenberg.spiegel.de/in...> (Stand April 2008)

Urheberrechtsfragen

Bei den angebotenen Texten handelt es sich in den allermeisten Fällen um Texte von Autoren, die vor mehr als 70 Jahren verstorben sind. In Deutschland endet der Urheberrechtsschutz 70 Jahre nach dem Tod des Urhebers (und nicht etwa nach dem Erscheinen des Werks)². Nach Ablauf dieser 70 Jahre gelten die betreffenden Werke als sog. *gemeinfrei*. Obwohl die in das Projekt Gutenberg aufgenommenen Werke gemeinfrei sind, erhebt die Firma Hille & Partner einen urheberrechtlichen Anspruch auf sämtliche Inhalte der Gutenberg-Seite. Die Firma Hille & Partner konvertiert die Texte in das im Internet übliche HTML-Format und nimmt die Verlinkungen vor. Aus dieser Konvertierung einen Urheberrechtsanspruch abzuleiten, erscheint aus juristischer Sicht zumindest fragwürdig. Der Entscheid eines zuständigen Gerichts zu dieser Frage ist noch ausstehend.

Was bringt die Zukunft?

Gemäss eigener Aussage möchten die Betreiber des Gutenberg-Projektes mit ihrer Arbeit ohne finalen Zeithorizont weiterfahren, ein eigentlicher Abschluss des Projektes ist nicht geplant. So können bereits heute Texte von Thomas Mann oder Hermann Hesse eingeschickt werden, die erst am 1. Januar 2026 resp. am 1. Januar 2033 gemeinfrei werden und damit ins Netz gestellt werden können. Das zeigt: Die Betreiber haben einen langen Atem und rechnen in grosszügigen Zeitdimensionen.

Sämtliche Texte müssen vor der Aufschaltung gescannt und anschliessend korrektur gelesen werden. V.a. das Korrekturlesen raubt enorm viel Zeit. Um das rasch wachsende Volumen bewältigen zu können, wurde das Projekt GaGa (Gemeinsam an Gutenberg arbeiten) initiiert: Auf einer separaten Internetseite³ wird die Verteilung der Korrekturlese-Arbeiten auf viele Schultern resp. Augen organisiert.

Fazit: Um zu einer veritablen Institution zu werden, einer verlässlichen Anlaufstelle in Sachen klassische deutschsprachige Literatur, müsste die Gutenberg-Seite von der technischen Seite her noch stark verbessert werden (bessere Volltextsuche, logischere Verknüpfungen). Im gegenwärtigen Zustand weist das Projekt einfach noch zu viele Mängel auf.

Das Project Gutenberg und das Projekt Gutenberg

Die virtuelle Bibliothek mit dem verwirrend ähnlichen Namen «Project Gutenberg» ist älter als das Projekt Gutenberg. Das Project Gutenberg wurde bereits 1971 gestartet. Offenbar wollte das Projekt Gutenberg mit der starken Anlehnung an das bereits bestehende Project Gutenberg von dessen bereits breit eingeführtem Namen profitieren. Das Project Gutenberg stellt Texte, für die kein Copyright besteht oder deren Copyright von den jeweiligen Autoren explizit freigegeben wurde, als E-Texte ins Internet. Der Korpus besteht zur überwältigenden Mehrheit aus englischsprachigen Büchern. Auf Deutsch sind nur etwa 200 Bücher greifbar – zu wenig, um eine deutschsprachige Gemeinde nachhaltig binden zu können.

Seitenblick I: Das französische Projekt Gallica – <http://gallica.bnf.fr/>

Gallica wird von der französischen Nationalbibliothek verantwortet und ist weltweit eines der grössten Digitalisierungsprojekte, die es zurzeit gibt. Initiiert wurde das Projekt 1997. Neben einer Sammlung von lateinischen Texten (in der separaten *Analytic Bibliography of Online Neo-Latin Texts*) enthält Gallica eine Unmenge von urheberrechtsfreien Texten in französischer Sprache, aber auch fremdsprachige Literatur, u.a. auch auf Deutsch.

Gallica liegt in einer neuen (Test-)Version vor. Gallica2 wurde notwendig, nachdem das ursprüngliche Projekt Gallica den technischen Anforderungen nicht mehr genügen konnte. Einen Überblick über die Geschichte und den Übergang von Gallica zu Gallica2 findet man auf http://www.bnf.fr/pages/catalog/bibliotheque_numerique.htm (am Schluss, unter dem Titel «Dates clés» [nur auf Französisch] (Stand April Ende 2008)). Die Nationalbibliothek hat sich vorgenommen, ab 2007 jedes Jahr gegen 100 000 neue Bücher in das Netz einzuspeisen – eine gewaltige Aufgabe.

In Gallica2 wurde u.a. die Navigation verbessert und die Interaktivität erhöht.

Gallica bietet eine sehr detaillierte thematische Suche (nach Fachgebieten, Personen, Ereignissen, Sprachen etc.), eine erweiterte Suche mit der Möglichkeit, zahlreiche Suchbegriffe aus zahlreichen Themenbereichen einzugeben, sowie eine Suche, in der der Titel eines Buches oder Teile davon, der Name des Autors und eine Volltextsuche miteinander kombiniert werden können.

Von Gottfried Keller z. B. findet man einzig eine französische Übersetzung der Novelle «Romeo und Julia auf dem Dorfe», Robert Walser findet man nur in ein paar verstreuten Werken der Sekundärliteratur. Man verliert sich (gerne) in den Weiten der Suchfunktion und stösst zwischendurch auf «abgelegene» Titel wie etwa «*Zur Anthropologie der Walser des grossen Walsertales im Vorarlberg*» (Romedi Wacker, 1912) oder «*Erinnerungen aus einer Reise durch das südliche Deutschland, Österreich, die Schweiz, in das mittägliche Frankreich und nach Algier*» (Moritz von Haacke, 1840).

Die Bücher werden seitenweise als PDF angezeigt. Teilweise können die Bücher ganz heruntergeladen werden. Die einzelnen Seiten wurden gescannt, oder es wurden Mikrofilme digitalisiert. Ganz allgemein lässt die Qualität zu wünschen übrig, manche Teile sind schlicht unles- und somit unbrauchbar. Wenn Gallica es schafft, eine Qualitätssicherung auf die Beine zu stellen, dürfte es zu einer der ersten Adressen unter den digitalen Bibliotheken werden. Bei 100 000 neuen Titeln pro Jahr ist eine Qualitätssicherung allerdings eine herkulische Aufgabe. Wäre weniger vielleicht mehr?

Johannes Gutenberg Der Namensgeber des Gutenberg-Projekts hiess eigentlich Gensfleisch. Der Erfinder der Schwarzen Kunst wurde zwischen 1387 und 1400 in Mainz geboren (über das genaue Geburtsdatum besteht Unklarheit), am 3. Februar 1468 ist er ebendort gestorben. Seine – technische – Leistung bestand darin, ein neues Druckverfahren mit beweglichen Metalllettern erfunden zu haben. Es wird angenommen, dass die Legierung, die zur Herstellung der Lettern verwendet wurde (Zinn, Blei, Antimon und ein Zusatz von Wismut), ebenfalls eine Erfindung von Gutenberg war, genauso wie die Druckerpresse, die eine grössere, billigere und schnellere Produktion von Druck-Erzeugnissen ermöglichte. Für jede Letter wurde eine erhabene, seitenverkehrte Form (Patrise) angefertigt, die anschliessend in Kupfer eingeschlagen wurde. In die Kupferform wurde dann die negative Form (Matrise) mit der flüssigen Legierung eingegossen. Eine Würdigung der kulturgeschichtlichen Wirkung von Gutenberg ist schwieriger vorzunehmen. Die Erfindung der Buchpresse löste eine eigentliche Medienrevolution aus, das Gedankengut des Humanismus und reformatorische Ideen fanden schneller Verbreitung, zudem führten die massenweise vervielfachten Druck-Erzeugnisse zu einer eigentlichen Alphabetisierungswelle. Jedenfalls wurde die Erfindung des Buchdrucks vom US-amerikanischen Magazin Time Life zur herausragendsten Erfindung des zweiten Jahrtausends gekürt. Ob Gutenberg selbst mit der ihm vom – ebenfalls US-amerikanischen - Netzwerk A & E verpassten, ein wenig marktschreierischen Bezeichnung «Man of the Millenium» einverstanden gewesen wäre, steht auf einem anderen Blatt. Quellen: Meyers grosses Taschenlexikon in 24 Bänden/Wikipedia/Gutenberg-Museum Weiterführende Literatur, Unterlagen und Links: Füssel, Stephan: Johannes Gutenberg. Reinbek b. Hamburg 31999 (Rowohlt Monographie). Ders.: Gutenberg und seine Wirkung. Frankfurt 22004 (Insel Verlag). Schnittstelle Gutenberg – Interaktive CD-ROM – Eine interaktive Entdeckungsreise durch 1000 Jahre Medienkommunikation. Interaktive CD-ROM für MAC und PC mit 32 Themenscreens, interaktiver Druckwerkstatt und Diskussionsforum, halbstündigem Video und einer Stunde MusikMainz: Schmidt, 2000, CD-ROM 29,80 Euro ISBN 3-87439-493-X.

<http://www.mainz.de/Gutenberg/...> (Stand Ende April 2008).

Gutenberg in Zahlen:

- mehr als 4000 Bücher
- über 1 200 000 Buchseiten
- ca. 20 000 Gedichte
- 1800 Märchen
- 1200 Fabeln
- 3500 Sagen

Stand: Oktober 2007

Seitenblick II: Das europäische Projekt «Europeana»

<http://www.europeana.eu/>

Am 25. August 2006 verabschiedete die Europäische Kommission eine Empfehlung zur Digitalisierung und Langzeiterhaltung digitaler Objekte, die den EU-Mitgliedsstaaten empfahl, Digitalisierungszentren aufzubauen, um das Kulturerbe Europas online über die Europäische Digitale Bibliothek verfügbar zu machen. Im November 2006 wurde die Idee einer Europäischen Digitalen Bibliothek von den Kultusministern aller EU-Mitgliedsstaaten unterstützt und durch das Europäische Parlament in seiner Resolution vom 27. September 2007 bestätigt. Aus der Resolution heraus entstand schliesslich das Projekt «Europeana».

Das Projekt wird vom EDLnet (the European digital library network) entwickelt. Diesem thematischen Netzwerk gehören um die 90 Partner. **Eine vollständige Liste sämtlicher beteiligter Partner findet sich unter** <http://www.europeana.eu/partne...> **(Stand Ende April 2000)**, an, die das Projekt gemeinsam voranbringen.

Die Website «Europeana» soll den direkten Zugang zu mehreren Millionen digitaler Objekte (Bilder, Bücher, Filme, Zeitungen, Tondokumente) aus Europas Bibliotheken, Museen, Archiven und audiovisuellen Sammlungen über ein einziges Portal bieten. «Europeana» wird im November 2008 online gehen. Zurzeit ist eine Testversion aufgeschaltet, die allerdings Umfang und Grösse des Projekts bereits erahnen lässt. Angestrebt werden 2 Mio. digitale Objekte, die den Usern frei zur Verfügung gestellt werden sollen.

Die Inhalte werden aus Beständen ausgewählt werden, die in europäischen Museen, Bibliotheken, Archiven und Sammlungen von Tondokumenten bereits in digitalisierter Form vorliegen.

Im Prototyp, der im November 2008 aufgeschaltet wird, ist eine dreisprachige Benutzerführung vorgesehen (deutsch, französisch, englisch), weitere Sprachen sollen folgen.

Bis 2010 soll der Bestand, so die ehrgeizige Zielsetzung, auf 6 Mio. digitale Objekte aufgestockt werden.

Es wird spannend sein, zu beobachten, wie sich diese zweite grosse Bibliothek nach Alexandria entwickeln wird – und wie und ob beispielsweise Schweizer Bestände in den Korpus Eingang finden werden.

1 Anzusteuern über <http://gutenbergshop.abc.de/ca...> (Stand Ende April 2008).

2 Gemäss Paragraph 64 des Urhebergesetzes.

3 <http://www.gaga.net/pgdp/faq/P...> (Stand Ende April 2008).



Rolf Hubler

Journalist (Biel-Bienne)

Abstract

Français

Le projet allemand Gutenberg est un vaste projet qui veut digitaliser et rendre accessibles gratuitement sur internet les classiques de la littérature allemande. Inauguré en 1994, c'est le Spiegel qui, depuis 2002, en est le principal partenaire: il assume les frais de la mise en ligne et assure le soutien technique. Le contenu et les tâches rédactionnelles reviennent au team Gutenberg. Ce projet s'est bien développé, mais reste encore cantonné au domaine des livres. D'autres projets en Europe visent à étendre le champ de mise en ligne à d'autres sources, telles que bandes sonores, films, photos, etc. L'auteur propose un panorama de la situation actuelle.

Le projet Gutenberg édite des livres d'auteurs décédés depuis plus de 70 ans, en raison de la prescription des droits d'auteurs en Allemagne, même si la question des droits d'exploitation en ligne n'est pas encore vraiment réglée. Sur ce site, les textes constituent naturellement le centre d'intérêt. Ils sont présentés selon trois dimensions: un registre alphabétique des auteurs, les différents genres littéraires et une recherche de textes. Des lacunes sont présentes qui ne permettent pas d'utiliser le site pour des recherches scientifiques. L'impression de documents laisse aussi à désirer. Les informations sur le projet et sur les éditions du Spiegel peuvent être utiles, mais pas indispensables. Sous la rubrique «spécial», le site offre encore des «handy-livres», qui peuvent être chargés et lus sur un téléphone mobile. L'utilité d'une telle offre laisse songeur. Plus intéressante est l'offre «Gutenberg-DE Edition 10» qui contient dix DVD présentant l'archive complète du projet Gutenberg. Le projet va se poursuivre, aucun terme n'ayant été fixé. Ainsi, des textes de Thomas Mann pourront être mis en ligne dès 2026, ou ceux de Hermann Hesse dès 2033. En 2007, ce projet représente plus de 4000 livres et 1 200 000 pages.

Au fur et à mesure de l'agrandissement de la toile, un partenariat sera développé pour les travaux de relecture et de correction avant la mise en ligne. Pour devenir une véritable institution, le projet Gutenberg doit améliorer l'aspect technique; pour l'instant, il souffre encore de trop graves lacunes dans ce domaine.

Il existe en anglais une version de ce projet, qui a démarré en 1971, qui met simplement à disposition des ouvrages essentiellement de langue anglaise. En France, le projet Gallica (

<http://gallica.bnf.fr>

), débuté en 1997, est l'un des plus grands projets actuels. Collection d'ouvrages en latin et de textes français et de langues étrangères, libres de droits, Gallica a prévu d'introduire sur le site environ 100 000 nouveaux livres chaque année. Malgré une amélioration de la navigation, la qualité de Gallica laisse encore à désirer, certaines pages étant difficilement lisibles ou inutilisables.

Le projet européen Europeana(

<http://www.europeana.eu/>

) a été lancé en 2006, sur recommandation de la Commission européenne, pour digitaliser et conserver à long terme différents documents relatifs entre autres à l'héritage culturel de l'Europe. EDLnet, la bibliothèque digitale européenne, a mis au point le projet auquel participent 90 partenaires. La page web «Europeana» (

<http://www.europeana.eu/>

) doit offrir l'accès à des millions d'objets (images, livres, films, journaux, etc.) provenant des bibliothèques, musées, archives et collections audiovisuelles d'Europe. C'est en novembre 2008 qu'aura lieu la mise en ligne dans trois langues d'abord: allemand, français et anglais. D'autres langues devraient suivre. Jusqu'en 2010, cette bibliothèque, la plus grande après celle d'Alexandrie, devrait réunir plus de 6 millions d'objets.