

Katalog und Textkorpus zu Diskettenmagazinen der 1980er und 1990er (Re-)Digitalisierung frühen digitalen Kulturerbes

Roeder, Torsten

dh@torstenroeder.de
Universität Würzburg, Deutschland
ORCID: 0000-0001-7043-7820

Yannik, Herbst

yannik.herbst@uni-wuerzburg.de
Universität Würzburg, Deutschland
ORCID: 0000-0002-6547-9599

Johannes, Leitgeb

johannes.leitgeb@stud-mail.uni-wuerzburg.de
Universität Würzburg, Deutschland
ORCID: 0009-0006-2058-9133

Madlin, Marenc

madlin@mrmuseum.de
Universität Würzburg, Deutschland
ORCID: 0009-0003-7434-827X

Tomash, Shtohryn

tomash.shtohryn@stud-mail.uni-wuerzburg.de
Universität Würzburg, Deutschland
ORCID: 0009-0000-4597-603X

Überblick

Dieser Vortrag präsentiert die Ergebnisse eines Drittmittelprojekts, das sich die Erschließung von Diskettenmagazinen der 1980er und 1990er Jahre zum Ziel gesetzt hat. bei den sogenannten „Diskmags“ handelt es sich um digitale multimediale Zeitschriften, die auf elektromagnetischen Floppy-Disks verbreitet wurden und auf klassischen Heimcomputersystemen wie z.B. Apple II, C64, Sinclair ZX Spectrum, Schneider CPC 464 und vielen anderen lesbar waren. Anfangs vor allem als kommerzielle Magazine für die frühen, hochpreisigen Systeme produziert (Beispiel dazu in Abbildung 1), entstanden mit der Markteinführung günstigerer Geräte aber auch bald Magazine aus der rapide wachsenden Heimcomputer-Community (Beispiel dazu in Abbildung 2). Diskmags stellten in der Zeit vor dem Breitbandanschluss und dem Web 2.0 ein relevantes „born digi-

tal“-Medium dar, in dem sich unterschiedlichste frühe digitale Kulturszenen hauptsächlich durch Text, aber auch mittels Bild, Animation und Ton untereinander austauschten. Inhaltlich waren Diskmags aber nicht auf die Heimcomputerszene beschränkt, sondern es existierten beispielsweise auch Fanzines und Literaturmagazine (Beispiel dazu in Abbildung 3).



Abbildung 1: Das kommerzielle Magazin »Softdisk« (1981-1985, System: Apple II) war das erste Diskmag, das auf 5.25"-Diskette publiziert wurde; davor gab es bereits ähnliche Magazine auf Datensette.



Abbildung 2: Das Szene-Magazin »Bad Mag« (1992-1993, System: Amstrad/Schneider CPC 464) wurde auf der ungewöhnlichen 3-Zoll-Diskette verbreitet und versammelte Beiträge zur der Demoszene in Großbritannien, Irland und Westdeutschland.



Abbildung 3: Das Magazin »Warp« (1995-1996, System: Atari ST) richtete sich ausschließlich an Star Trek Fans und funktionierte über eine gamifizierte Menüsteuerung.

Die Überlieferungslage der Diskmags ist prekär, da sie nicht durch Bibliotheken oder Archive gesammelt wurden (vgl. Roeder 2020). In dem hier vorgestellten Projekt entstand ein Katalog, durch den jetzt mehrere tausend Diskettenmagazine und weit über zehntausend Ausgaben erstmals systemübergreifend und nach wissenschaftlichen Kriterien recherchierbar sind (Zwischenergebnisse und illustrierte Beispiele siehe Roeder et al. 2023). Mithilfe von Textmining-Methoden wurde aus den Binärdateien deutschsprachiger Magazine ein Textkorpus erstellt, das nicht nur für die Volltextsuche, sondern auch für die Erforschung der Sprache und der Thematiken der frühen digitalen Kultur nachgenutzt werden kann. Dieser Beitrag wertet die bisherigen Projektergebnisse aus und stellt die angewendeten Methoden zur Diskussion.

Gefördert wurde das einjährige Projekt in erster Linie als Kooperationsprojekt des NFDI-Konsortiums Text + (siehe <https://www.text-plus.org/forschungsdaten/kooperationsprojekte/>), in dessen Datenservices sowohl die Katalog- als auch Textdaten einfließen. Außerdem erhielt das Projekt Zuschüsse durch den Unibund der Universität Würzburg und durch die Vogel Stiftung Dr. Eckernkamp. Gehostet wird das Projekt am Zentrum für Philologie und Digitalität der Universität Würzburg unter der Domain diskmags.de.

Hintergrund

In den 1980er Jahren etablierten sich Heimcomputer nach und nach als Unterhaltungskonsolen und Arbeitsgeräte gleichermaßen. Die Basis, die diese Entwicklung vorantrieb, bestand aus einer Subkultur von Computer-Enthusiasten, in der sich Anwender, Computerspieler, Programmierer, Spieleentwickler, Democoder sowie Hacker und Cracker zusammenfanden. Das authentische Erlebnis jener frühen digitalen Kultur ist uns heute jedoch nicht mehr präsent. Definiert wurde es zum einen durch die Hardware, die einerseits durch ein typisches „look and feel“, andererseits durch eine spezifische Ausstattung (Geschwindigkeit, Speicher, Soundkanäle, Farben, Bildauflösung, Peripherieanschlüsse) oft sehr charakteristisch war. Zum anderen war die Bedienung von Programmen mehr hybrid als „born digital“: Computern lagen dicke Handbücher bei, Anwendungsprogramme wurden oft mit umfangreichen Bedienungsanleitungen ausgeliefert, und auch in Spielen war das Druckbeiwerk manchmal unerlässlich. Programme wurden oft nicht nur digital kopiert, sondern händisch aus Printmagazinen abgetippt. Software erforderte Einarbeitungszeit und eine hohe Frustrationsschwelle, bei vergleichsweise geringer Stabilität der Betriebssysteme.

Die intensive Selbstdokumentation, die von der heutigen digitalen Massenkultur durch zahlreiche Medienkanäle geleistet wird, steht für die damalige Zeit nicht in derselben

Dichte zur Verfügung. Eine besondere Rolle fällt deshalb den zahlreichen Periodika zu, die in jener Zeit produziert wurden und Einblicke in die Vielfalt der Heimcomputer-szene geben. Einen guten Teil decken hier die professionellen Printmagazine ab, die allerdings das Geschehen innerhalb der diversen Subszene nur selten fokussierten. Daher lohnt die nähere Betrachtung von Diskettenmagazinen: eine Art multimedialer Born-Digital-Journale, die mit dem Beginn der Heimcomputerkultur entstanden sind und nicht nur die charakteristischen multimedialen Ansätze jener Zeit spiegeln, sondern auch Dokumente der vielfältigen Kreativität und Lebendigkeit der einstigen Heimcomputer-Community darstellen.

Diskettenmagazine

Wie einleitend erläutert und veranschaulicht, handelte es sich bei Diskettenmagazinen um digitale Zeitschriften, die ausschließlich auf einem dafür geeigneten Computersystemen rezipiert werden konnten. Ein vereinheitlichtes System gab es nicht: Jedes Diskmag beinhaltete seine eigene Reader-Software.

Verbreitet wurden Diskettenmagazine stets auf ihrem namengebenden Medium, aber auf ganz unterschiedlichen Vertriebswegen. Einige konnte man regulär am Kiosk beziehen, andere wurden gegen Einsendung von Retourporto und Leerdiskette oder einen angemessenen Unkostenbeitrag per Post versendet, manche verbreiteten sich ausschließlich durch Privatkopien.

Primär wurden sie in Gegenden mit hoher Heimcomputerdichte hergestellt. Prinzipiell gilt: Wo es Heimcomputer gab, gab es Diskmags. Durch die Bindung an bestimmte Computersysteme war ihr Verbreitungsgrad außerdem weniger durch sprachliche Grenzen, sondern primär durch technische Hürden limitiert.

Inhaltlich boten Diskettenmagazine sowohl Informationen zu digitalen Technologien, Geräten, Spielen und Anwendungsprogrammen als auch Programmier- und Bastelanleitungen. Einige Diskmags widmeten sich intensiv einem Bereich aus Grafik, Musik, Games, Literatur oder Fandom. Fast immer wurde Software (Anwendungen, Spiele, Demos) mitgegeben. Diskettenmagazine konkurrierten kaum mit den textbasierten Bulletin Board Systems (BBS) der späten 1980er, jedoch übernahm das World Wide Web im Verlauf der 1990er, spätestens mit dem Aufkommen von Breitbandverbindungen, die Funktion der multimedialen Informationsverteilung, und viele der Diskettenmagazine wurden zusammengelegt oder beendet. Nur wenige wurden online weitergeführt, manche führten dabei die Bezeichnung „Diskmag“ im Titel weiter; in der Demoszene ist dies bis heute üblich.

Katalogisierung

Bibliotheken sammelten Diskettenmagazine nur, wenn ein wesentlicher Printanteil vorlag, was nur selten der Fall

war. Verlässliche Informationen und digitale Dokumente (vor allem Disk-Images als Binärdateien, aber auch Screenshots sowie Scans von zusätzlichem Printmaterial) finden sich am ehesten auf einschlägigen Fansites, die zudem häufig auf einzelne Systeme, Sprachen oder Themen begrenzt sind. Selbst die englischsprachige Wikipedia verzeichnet lediglich einen Bruchteil.

In der ersten Phase des Erschließungsprojektes stand deshalb die umfassende Katalogisierung im Vordergrund. Dazu wurden mehrere bestehende Verzeichnisse ausgewertet und zusammengeführt. Aktuell sind dies fünf Datenquellen: Demozoo (<https://demozoo.org>), Pouet (<https://www.pouet.net>), C64 Scene Database (<https://csdb.dk>), Internet Archive (<https://archive.org>, mehrere Collections) und ZXpress (<https://zxpress.ru>); weitere könnten zukünftig folgen. Parallel dazu wurde Sekundärliteratur ausgewertet (Volko 2012).

Diese Datenquellen wurden mithilfe von Scraping-Techniken geharvestet und in einem GitHub-Repository abgelegt (<https://github.com/zpd-digital-editions/Diskmags>). Für den Abgleich der Daten kam die bewährte Software OpenRefine (<https://openrefine.org>) zum Einsatz. Die Zuordnung der Einzelausgaben zu den jeweiligen Titeln war aufgrund vieler ähnlich lautender Titel und vielen Titelvarianten keine triviale Aufgabe; widersprüchliche Angaben wurden entsprechend nachrecherchiert. Insbesondere die Angaben zu Sprachen und Ursprungsländern stellten sich als problembehaftet heraus. Die vereinheitlichte Datensammlung wurde anschließend in ein Semantic MediaWiki exportiert und steht unter diskmags.de für die Recherche und Nachbearbeitung zur Verfügung. Mithilfe des Wikis können die existierenden, teils sehr aktiven Communities bei der Datenkuratierung direkt einbezogen werden. Es wird angestrebt, die erhobenen Katalogdaten über die Infrastruktur von Text+ in Normdatenkataloge und ins Linked Open Data Network einzuspielen.

Ergab deshalb die erste Schätzung vor Projektstart noch, dass vermutlich mit etwa 200 bis 300 Diskettenmagazinen zu rechnen sei, musste diese Zahl selbst nach Bereinigung zahlreicher Doppeleinträge erheblich nach oben korrigiert werden. Aus den genannten Datenbanken wurden Nachweise zu ca. 2.500 unikalen Titeln mit weit über 10.000 Einzelausgaben extrahiert. Die rein quantitative Relevanz dieses ungewöhnlichen Mediums wurde somit um den Faktor 10 unterschätzt.

Die Auswertung der Datensammlung ergab mehrere Erkenntnisse. Nach dem aktuellen Stand der Datensammlung ist der Höhepunkt der Diskmag-Kultur zwischen den späten 1980ern und dem Ende der 1990er anzusetzen (vgl. Abbildung 4). Davor existierten Diskmags vor allem für eher exklusive und hochpreisige Geräte wie Apple II. Die vor allem durch Commodore eingeleitete Preissegmentierung des Heimcomputermarktes beförderte die Herausbildung einer breiteren Heimcomputer-Community und einer entsprechenden Diskmags-Kultur. Mit dem Erscheinen leistungsstärkerer Computer setzte sich dies in mehreren Wellen z.B. auf Systemen wie Commodore Amiga und Atari fort, überraschenderweise aber auch auf dem eigentlich

deutlich älteren Spectrum ZX, der in den postsowjetischen Ländern durch Nachbauten eine immense Rezeption erfuhr. In der späten Zeit dominierten dann MS-DOS-basierte Diskmags, bevor durch CD-ROMs und die Möglichkeiten des Web 2.0 das Diskmag als Medienformat weitgehend obsolet wurde. Es existiert allerdings weiterhin als Community-Produkt in einigen bis heute aktiven Retrocomputing-Szenen. Dies ist ein Glücksfall für die Erschließung, da so das teils obskure Wissen um die Funktionsweise der Hard- und Software durch Zeitzeugen zugänglich ist.

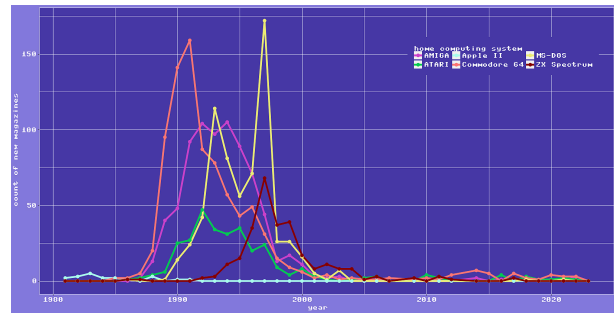


Abbildung 4: Anzahl neuer(!) Diskettenmagazine nach Jahr und System (Auswahl). Grafik: Johannes Leitgeb. Stand: Juli 2023.

Textkorpus

Das Textkorpus umfasst Plaintext aus mehreren vollständigen Jahrgängen deutschsprachiger Diskettenmagazine. Die Textextraktion aus mittlerweile obsoleten Datenträgern bzw. Datenformaten birgt eine Reihe von antizipierbaren Problemstellungen: So ist bereits die Zeichenkodierung stark herstellerabhängig und konnte zudem softwareseitig manipuliert werden, so dass die angezeigten Zeichen nicht immer dem jeweiligen Kodierungsstandard entsprechen. Ein Unicode-Mapping ist deshalb nicht immer eindeutig möglich, insbesondere wenn selbsterstellte Symbolzeichen verwendet wurden, deren Semantik manchmal über Blockgrafik-Elemente hinausging. Zudem wurde gerade Text aufgrund der relativ geringen Speicherkapazität von Floppys gerne komprimiert, wie auch eine entropische Analyse nahelegte. Zum Teil konnte hier jedoch auf Community-Initiativen zurückgegriffen werden, durch welche bereits teilweise eine Re-Digitalisierung vorliegt. Als entscheidend für die Textextraktion erwies sich jedoch die Erkenntnis, dass Texte häufig als Screencodes gespeichert wurden, die von den üblichen Zeichenkodierungen deutlich abweichen konnten. So z.B. lautet der ASCII- und Commodore-PETSCII-Code für den Buchstaben »A« einheitlich 65, der Screencode des Commodore 64 jedoch 1.

Im Unterschied zum Katalog sind bei der Erstellung des Textkorpus aber vor allem urheber- und personenschutzrechtliche Fragen zu berücksichtigen, wofür aufgrund der Komplexität der Rechtslage (z.B. nicht ermittelbare Urheber, nicht eindeutige Rechtfreigaben, ggf. sensible Daten)

zum Projektabschluss eine entsprechende Handreichung vorgelegt werden wird.

Ziel ist die Publikation eines Textkorpus in einem flach hierarchischen Format wie DTABf (siehe <https://www.deutschestextarchiv.de/doku/basisformat/>), möglichst mit Named Entity Recognition und Artikelgrenzen, abhängig von der Rechtslage zugangsbeschränkt, ggf. zumindest mit der Möglichkeit eines Stichwortindexes und einer Wortstatistik. Für die Auswertung des Textkorpus ist beabsichtigt, sowohl das spezielle Vokabular als auch die besondere Stilistik der Heimcomputer-Szene mit Auswertungsverfahren der DH zu untersuchen. Mittels Named Entity Recognition lassen sich ggf. Netzwerke aus Personen und Gruppen erschließen, Softwaretitel auffinden und hinsichtlich ihrer Rezeption untersuchen oder Diskussionen um Computersysteme, deren Leistungsspektren und deren spezifischer Auslotung verfolgen. Ab einer größeren Textmenge erscheint auch die Anwendung von Topic-Modeling-Verfahren sinnvoll.

Ausblick

In den Digital Humanities steht die Auseinandersetzung mit digitalem Kulturerbe als Erhaltungsgegenstand noch relativ am Anfang. Im Umgang mit älteren Computersystemen existiert in den Digital Humanities vor allem im europäischen Raum, von einzelnen Projekten abgesehen, noch keine allgemeine Methodik. Standardformate wie TEI berücksichtigen digital überliefertes Material und Medienformate bislang nicht oder nur unzureichend. Die bereits hohe technologische Distanz erfordert jedoch geradezu eine verstärkte Auseinandersetzung mit digitalen Überlieferungsformen, sowohl bezogen auf einzelne Objekte als auch auf die historischen digitalen Ökosysteme als Ganzes. Insbesondere zu den 1980er und 1990er Jahren besteht eine erhebliche Lücke hinsichtlich der wissenschaftlichen Erschließung. Das Erschließungsprojekt zu Diskettenmagazinen stellt in dieser Hinsicht eine kleine Pionierleistung dar und bietet zahlreiche Anknüpfungspunkte für zukünftige Forschungsprojekte, die sich mit älterer Software, Hardware, Datenträgern und Datenformaten befassen.

Bibliographie

Roeder, Torsten: Rescuing Diskmags: Towards Scholarly [Re-]Digitisation of an Early Born-Digital Heritage, in: *Magazén* 1,3, 2022, S. 139–58.

Roeder, Torsten; Herbst, Yannik; Leitgeb, Johannes; Marenc, Madlin; Shtohryn, Tomash: Preserving the Early Born-Digital Heritage of Floppy Disk Magazines. Zenodo, 2023.

Ruan, Jianhai; P. McDonough, Jerome: , in: IEEE International Symposium on IT in Medicine & Education, 2009, S. 745–48.

Volko, Claus-Dieter: Enzyklopädie der Diskmags. Norderstedt: BoD, 2012.