

THOMAS STÄCKER

Wollen Sie C:*.* wirklich löschen? – die Zukunft der Provenienzforschung im elektronischen Zeitalter

Wollen Sie C:*.* wirklich löschen? – Wer erinnert sich noch an die misstrauische Frage, die uns der von einem DOS genannten System angetriebene Rechner stellte, um das versehentliche Löschen eines ganzen Verzeichnisses zu verhindern? DOS, für Disk Operating System, ist angesichts der mühelos gewordenen grafischen Oberflächen schon fast ganz in Vergessenheit geraten. Auch heute warnt der Computer noch den unvorsichtigen Nutzer davor, den Inhalt eines ganzen Verzeichnisses in den Papierkorb zu befördern. Doch braucht man sich nicht mehr große Sorgen zu machen. Das zusammengeknüllte elektronische Manuskript läßt sich problemlos aus dem Abfalleimer recyceln. Erst explizites Leeren bewirkt, dass es endgültig im elektronischen Orkus verschwindet. Dennoch ist diese Leichtigkeit des Beseitigens ein Menetekel.

Das Wegwerfen, sei es absichtlich oder versehentlich, hat schon immer die Zunft der Archivare und Bibliothekare beunruhigt. Das Wegwerfen ist angesichts der unvordenklichen Interessen zukünftiger Generationen eine Bedrohung, es löst die Beunruhigung aus, jemand könne doch wider Erwarten Interesse an dem Weggeworfenen haben, es könne Wichtiges verschwinden. Diese elementare Unsicherheit war in der Vergangenheit nicht selten Motiv für das Sammeln des scheinbar Nebensächlichen, das Beiseitelegen und Aufheben des Unwichtigen. Das Abgelagerte der Tradition stand auf einem festen Substrat, dem Papier, das, auch wenn es der Abt Trithemius im 15. Jh. als flüchtigen Datenträger geißelte, sich zumindest in seiner säurefreien Form als Hadernpapier erstaunlich haltbar erwies. Ehemals Unwichtiges, das sorglos beiseite gelegte Kalenderblatt, die Flugschrift des letzten Theologenzwistes oder die neue Zeitung vom unlängst zu Neapel gesichteten Wundermensch mit Schuppenhaut finden sich so aufbewahrt zwischen den Koryphäen der literarischen Hochkultur und bieten der heute prosperierenden, auf soziologische und ökonomische Faktoren rekurrierenden wissen- und kulturgeschichtlichen Forschung reiche Nahrung. Ebenso sprichwörtlich geduldig ist Papier in der Aufnahme der anderen Texte, der, man könnte sie *mutatis mutandis* mit dem französischen Literaturtheoretiker Gérard Genette Paratexte¹ nennen, den Peritexten und Epitexten, den handschriftlichen Eintragungen, die auf den Vorbesitzer weisen, den Stempeln, die die Bibliothek kenntlich machen, aus der das Buch entwendet wurde, das Exlibris oder Supralibros, das Besitzerstolz und Selbstverständnis des Eigentümers zur Schau trägt, Immissionen, die es uns erlauben, Bücher in ein Verhältnis zu ihren ehemaligen Nutzern zu bringen, die Lesebiografien offenbaren und die in Gestalt hingebungsvoll ausführlicher Annotationen das Interesse des seinerzeit Forschenden lebendig werden lassen, eine Spurenlese in den geistigen

1 Genette, Gérard: Paratexte: das Buch vom Beiwerk des Buches, Frankfurt am Main 2001.

Netzwerken unserer Vorväter, auf deren Schultern wir, wie es Bernhard von Chartes erstmals im 12. Jh. so schön formulierte, als Zwerge stehen. Eine Geschichte, die auf Papier, weiter zurückgreifend, auf Pergament gebaut ist, und deren Substrat gegenwärtig mit dem ubiquitären Siegeszug elektronischer Medien in wachsendem Maße in Frage gestellt wird.

Die mangelnde Langlebigkeit oder eben Flüchtigkeit der elektronischen Medien ist ein Dauerbrenner in der bibliothekarischen und archivischen Fachdiskussion. Allerdings spielte sie bislang im Rahmen der Provenienzforschung, die noch weitgehend dem Papier vertraut, kaum eine Rolle. Und doch gewinnt diese bisher als weit entfernte Abstraktion verhandelte Sache auch im Bereich der Provenienzforschung zunehmend an Bedeutung, mehr als man es auf den ersten Blick vielleicht wahrhaben will und manchem lieb ist. Es sind vor allem die Archive, die an vorderster Front stehen und von Problemen wie unarchivierbaren nachgelassenen Disketten oder CDs umgetrieben werden; aber auch die Bibliotheken stehen vor wachsenden Problemen, die mit der Medienrevolution, dem Wechsel oder, um genauer zu sein, dem Verschwinden des Trägermaterials der Überlieferung einhergehen. Probleme, die auch um die Zukunft der Provenienzforschung fürchten lassen.

In diesem Beitrag soll eine Transposition gewagt werden, eine Übersetzung der klassischen Provenienzforschung und deren Methodik in die neue Medien-Welt. Dabei stehen zwei Problemfelder zur Diskussion. Zum einen die Frage der Provenienzermittlung bei elektronischen Dokumenten, zum anderen das Problem der elektronischen Langzeitarchivierung. Der Beitrag schließt mit einem Ausblick in Gestalt einer Art Science Fiction der Provenienzforschung und einem kurzen Resümee.

a) Provenienzermittlung bei elektronischen Dokumenten

Wir sind es, aus der Papierzeit kommend, gewohnt, bestimmte Merkmale als Provenienzträger zugrunde zulegen. Das sind bei personenbezogener Provenienzermittlung – dem außerhalb des Mittelalter häufigsten Falls – insbesondere handschriftliche Eintragungen in der Handschrift oder im Druck, meist auf dem Spiegel, dem Vorsatz oder Titelblatt.

Ein typisches Beispiel liefert ein Homerdruck (Abb. 1), der die Illias und die Odyssee enthält². Hier findet sich ein Eintrag, dass ein gewisser E. M. den Band 1543 für 1 Thaler und 13 Denarius (= 12 solidus = Pfennig oder Groschen) gekauft hat; daneben sind aber noch weitere Provenienzsichten zu erkennen. So hat sich der Sohn des berühmten Humanisten Johannes Caselius (1533–1613), Johann Ulrich Caselius, hier verewigt. Die andere Hand stammt wohl vom Vater selbst, der erst in

2 Signatur HAB: P146 Helmst 2°.

ΠΟΙΗΣΕΙΣ ΟΜΗ-
 ΡΟΥ ΑΜΦΩ ΗΤΕ ΙΛΙΑΣ ΚΑΙ Η
 ΟΔΥΣΣΕΙΑ, ΥΠΟ ΤΕ ΙΑΚΩΒΟΥ ΤΟΥ ΜΙΚΥΛΛΟΥ

*καὶ Ἰωαχίμου Καμεραρίου, ἀποδοτίας πρὸς τὴν νυνὶ
 ἐκδοσὴν ἡρασκυλάδου-
 σα.*

OPVS VTRVMQVE HOMERI ILIA-
 DOS ET ODYSSEAE, DILIGENTI OPERA
 Iacobi MICYLLI & Ioachimi CAMERARII recognitum.

*Μὴν 100 στίχοι
 ποιεῖ ἑκάστη Τροία
 καὶ ἡ Ἰουλίαν
 σύγχρονος ἡαία*

Ioannes Vovius Casolus



*Πορφύριος φιλοσοφῶν ἑπισημὰ ἐπισημασθεῖς.
 Τῶ αὐτῷ Πορφύριῳ περὶ τῆς ἢ Οδυσ-
 σεῖας τῶ νυμφαίου αὐτῶ.*

*Porphyrj philosophi Homeriarum quaestionum liber.
 Eiusdem de Nympharum antro in Odyssea opusculum.*

Cum privilegio Imperatorio & Regio.

E. M.

*Πλάταρχος. ὅτι δευκτέον ἐστὶ τὰ ποιήματα,
 ὅτι τοῖς φιλοσοφῶσιν μέλλουσι ἀλλ' ἢ ποιή-
 μασι πρὸς φιλοσοφῆσθαι.*

Αναο 1.5.43. Emi 112 & 130

Abb. 1

Rostock, dann ab 1576 an der Universität Helmstedt wirkte. Der Sohn starb schon in frühen Jahren, so dass der Vater ihn beerben musste. Man weiß also nicht sicher, ob die Eintragungen des Vaters vor oder nach dem Tod des Sohnes gemacht wurden, zumindest sind beide Provenienzen vertreten. An der HAB finden sich noch ca. 50 Bände mit Caseliusprovenienz, oft mit reichen Kommentaren des hervorragenden Gräzisten versehen.

Hier gelingt die Entzifferung entweder durch primäre Lektüre oder aber es bedarf eines gewissen Ermittlungsaufwandes oder probalistischer Schlüsse aus dem biographischen Kontext der fraglichen Personen (Vater beerbt Sohn) oder des Sammlungskontextes durch Sichtung umfangreichen Materials; auch um Abkürzungen, wie E. M. aufzulösen. Leider ist diese ein Einzelfall und daher eine Sackgasse.

Wie aber stellt sich die Situation beim elektronischen Dokument dar? Man kann unterstellen, dass auch im elektronischen Medium dieselben psychologischen Mechanismen wirken, die in der Vergangenheit zu dieser Art Eintragungen geführt haben.

Wenn man ein elektronisches Dokument unter Provenienzgesichtspunkten betrachtet, ist klar, dass man nicht so etwas wie eine handschriftliche Eintragung und sonstige Manipulation – einen Paratext – am und im Text erwarten darf. Die Suche wird sich auf andere Merkmale erstrecken müssen, Merkmale, die nicht in den sichtbaren Text eingetragen werden, sondern aus sekundären Quellen geschöpft sind.

Unterstellen wir einmal, jemand hätte sich ein Werk herunter geladen und es sich in WORD angesehen bzw. bearbeitet. Stellen wir uns vor, es sei ein provenienzrelevantes Dokument. Wenn wir nach der Herkunft fragen, bietet die Oberfläche des Textes zunächst keinen Anhaltspunkt. Gleichwohl gibt es Möglichkeiten, den Autor oder den Besitzer des Dokuments anhand anderer Merkmale festzustellen. Um etwas über dieses Dokument zu erfahren, könnte man zunächst im Eigenschafteneditor unter Statistik nachsehen (Abb. 2).

Dort erfahre ich, dass dieses Dokument im November 2000 zuletzt geändert und kurz vorher ausgedruckt wurde (im Ausdruck wurde wohl noch ein Fehler gefunden). Man kann also mit ziemlicher Sicherheit davon ausgehen, dass es zu dieser Zeit in seiner Endfassung abgespeichert wurde. Wir haben jetzt also schon einen relativ guten Anhaltspunkt bzw. einen *terminus ante quem*. Die Veranstaltung, die in dem fraglichen Dokument besprochen wird, fand vor dem November 2000 statt, in unserer Fiktion, das Werk, das hier betrachtet wurde, wurde im November 2000 bearbeitet. Auch die Lokalisierung des Dokumentes – wo entstand es, wer ist dafür verantwortlich – ist auf diese Art und Weise möglich. Allerdings wird es uns hier nicht so leicht gemacht. Autor ist ein „Projekt 11“ und das Gruppierungsmerkmal Firma lautet FEA e. V. Letzteres scheint auf einen eingetragenen Verein hinzuweisen. Nach einem Blick in Google ergeben sich Treffer mit der Provenienz „Potsdam“, die relativ eindeutig sind. Nun müsste nur noch das ominöse „Projekt 11“ gefunden werden. Dies gelingt allerdings über Google nicht mehr.

Natürlich kann man Texte auch noch weiteren Analysen unterziehen. Das hängt jeweils vom Format ab. Ein Blick mit dem Hexadezimaleditor auf den Dateicode mag

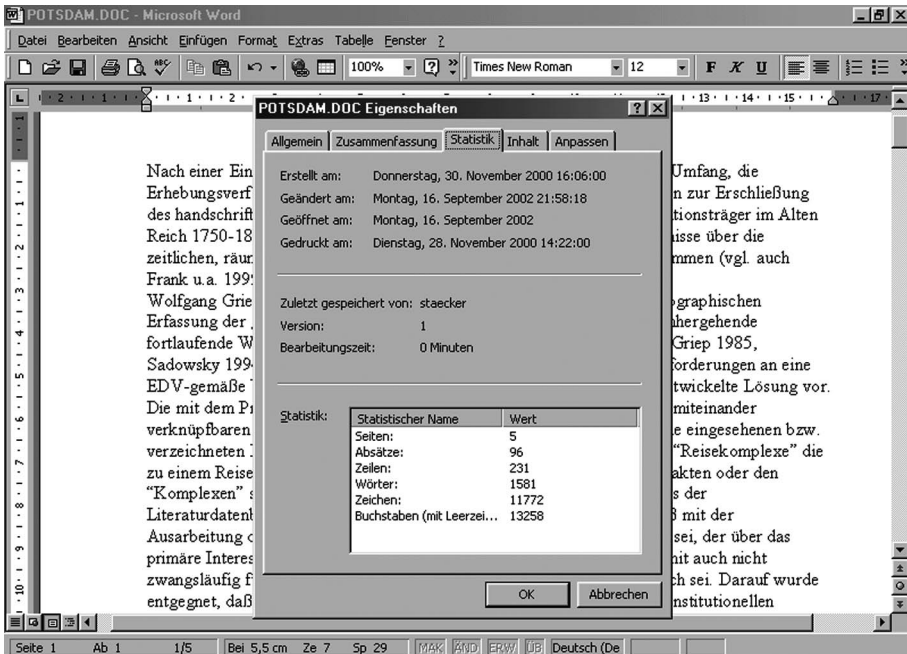


Abb. 2

in vertrackten Fällen manchmal weiterhelfen. So ist hier das Format des PDF-Files kodiert (1.2), aktuell ist 1.7. Man kann daraus schließen, dass entweder der Erzeuger ein altes Programm verwendet oder aber das Dokument schon ca. 6 Jahre alt ist. Ein möglicherweise für die Datierung wichtiges Indiz.

Wir können heute noch nicht abschätzen, welchen Gebrauch die Wissenschaft vom elektronischen Medium machen wird. Ob sie die elektronische Textversion als Gebrauchs- und Lesetext akzeptiert oder ob sich hybride Formen durchsetzen werden, die die Vorteile beider Medien verbinden, etwa dadurch, dass die Lektüre auf Papierbasis durch die Recherche per EDV ergänzt wird. Als sicher kann jedoch gelten, dass die Nutzung elektronischer Produkte auch im geistes- und kulturwissenschaftlichen Sektor zunimmt. Provenienzerschließung erhält vor diesem Hintergrund einen völlig neuen Charakter. Wie lässt sich an Hand einer elektronisch vervielfältigten Quelle feststellen, welchen Gebrauch ein Nutzer davon gemacht hat? Nach heutigem Ermessen spielt die individuelle Arbeitsumgebung dabei eine wichtige Rolle. Nicht nur, dass sich – dies wird heute eher der Regelfall sein – Ausdrucke von Artikeln oder gar ganzen Drucken im Besitz des Literaten oder Forschers befinden, man wird auch auf elektronische Kopien stoßen, die auf eine solche Nutzung hindeuten. Das Kopieren als zentrales Moment in der Nutzung elektronischer Texte exteritoriali-

siert jedoch die Nutzungsspur. Das materiale Objekt, die Diskette, die CD, der USB Stick, die Festplatte sind nicht mehr wie in klassischen Papierzeiten auch Träger der Benutzerspuren oder Eigentumsvermerke, sondern fungieren als gegen ihren Inhalt gleichgültige Container, wenigstens solange ihr Gebrauch anhält, Texte umgeschrieben und überschrieben werden, mit Ausnahme der soeben geschilderten Möglichkeiten, ein so gut wie keine sichtbaren Spuren hinterlassender Vorgang. Die Situation ändert sich indessen, wenn das Verarbeiten der Texte endet. Dann werden diese Medien Elemente eines Benutzungskontextes, liefern wertvolle Informationen zu Provenienzen. Allein, es gibt wegen dessen Charakters als Kopie eines wie auch immer gearteten Archetyps natürlich keine Mehrfachprovenienzen bzw. Benutzungsschichten mehr. Die beschriftete Diskette, die Festplatte oder auch der Account bei der Telekom stehen im Provenienzkontext einer Person. Nimmt man dies ernst, so ergeben sich für die Provenienzermittlung und auch -sicherung erhebliche Konsequenzen. Dabei ist abzusehen, dass die noch vorhandenen individualisierbaren Elemente wie die Beschriftung auf der Diskette oder CD-ROM zugunsten anderer Provenienzkennzeichen zurücktreten werden. Die Provenienzforschung – jetzt sind wir schon beim Science Fiction angelangt – wird eine Archäologie elektronischer Quellen und Verweise betreiben müssen, um die Nutzung elektronischer Texte zu erschließen. Nutzungsspuren finden sich weniger am Text oder genauer: hinter dem Text, als vielmehr im Verhältnis zum Text, in Form von PDF oder HTML – Downloads und persönlichen WORD-Dateien auf dem selbst genutzten Rechner, Texte, die möglicherweise als Kompilate modifiziert in mittelalterlich anmutender Sorglosigkeit um Autorschaft und Urheberrecht großzügig ins Netz zurückgeflossen sind. Oder man denke an die heute schon unübersehbar gewordene Flut wissenschaftlicher Texte, von Professoren oder Studenten ins Netz gestellt, die oft nichts anderes sind und auch nichts anderes sein wollen als Versatzstücke, Gebrauchs- bzw. Verbrauchstexte. Die Kopien von Texten, die in solchen Kompilationen oder Sammlungen auftauchen, lassen Lektüerverhalten transparent werden und bieten neue Szenarien der Provenienzermittlung. Besonders ergiebig aber für die elektronische Archäologie, so die Prognose, werden die Email- oder Blogarchive sein, seien sie öffentlich oder nur institutionell zugänglich, vorausgesetzt, sie bleiben erhalten und werden langfristig archiviert.

Man kann bilanzieren, dass die Provenienzforschung von morgen in sehr viel stärkerem Masse als die heutige Buch bzw. Papier gestützte Forschung auf die Nutzung nicht objekt-immanenter Daten angewiesen sein wird. Den Archiven nicht weniger als den Bibliotheken wächst hier eine besondere Aufgabe zu. Doch im Einzelnen stehen diesen Einrichtungen handfeste rechtliche Hindernisse im Weg. Zwar könnte man wissenschaftliche Bibliotheken mit der Archivierung benutzerbezogener Daten betrauen. Es werden ja heute schon die meisten Benutzungsvorgänge elektronisch erfasst. Doch verbietet der Datenschutz die Speicherung personenbezogener Daten über einen längeren Zeitraum hinaus. Diese rechtlichen Regelungen gelten natürlich auch für die Verlage, die elektronische Dokumente vertreiben. Leider werden diese Daten meist gelöscht und stehen damit der zukünftigen Provenienzarchäologie

nicht mehr zur Verfügung. Angesichts des hohen Quellenwertes, den uns etwa alte Benutzerbücher heute bieten, eine bedauerliche Entwicklung. Auch das legale Ausdrucken hilft nur bedingt, weil die schiere Masse der Daten eine elektronisch Form der Archivierung erzwingt. Bei allem Recht, den der Datenschutz in der Gesellschaft genießen muss, sind wir angesichts dessen auf dem besten Wege zu einem für die kommenden Generationen dunklen Jahrhundert zu werden. Hier besteht ein hoher Regelungsbedarf. Es darf nicht sein, dass die Tradierung forschungsrelevanter Daten durch sicher berechnete Datenschutzansprüche verhindert wird, vielmehr sollten Mittel und Wege gefunden werden – z. B. über die öffentlichen Archive, die Erfahrung mit persönlichkeitschutzrelevanten Informationen haben – solche Daten langfristig zu sichern.

Dies führt zum zweiten Komplex dieses Beitrages, der Langlebigkeit elektronischer Dokumente. Der Titel deutete bereits auf die Problemlage hin. Elektronische Dokumente sind wegen ihrer flüchtigen Natur, ihrer transitorischen Trägermaterialien bzw. nahezu verschwindenden Materialität prädestiniert, durch ein kleines Malheur, eine tragische Unachtsamkeit auf nimmer Wiedersehen zu verschwinden. Das reicht von versehentlich gelöschten Dateien bis zum nicht durch eine USV (unterbrechungsfreie Stromversorgung) gesicherten Server mit entsprechenden Programmabstürzen, bis zum Headcrash einer ungesicherten Festplatte. Von verklebten Magnetbändern, nicht mehr lesbaren USB-Sticks oder zerkratzten CD-ROMs ganz zu schweigen. Die Möglichkeiten, Daten zu verlieren, sind vielfältig und allgegenwärtig. Um diese Gefahr des Datenverlustes im Bereich elektronischer Publikationen zu vermindern, haben Bibliotheken in den letzten Jahren intensive Anstrengungen unternommen. Besonderes Problem der Langzeitarchivierung sind die schnellen Programm- und Formatwechsel. So sieht z. B. meine derzeit älteste noch verfügbare Datei aus. Eine Seminararbeit über Caesars Stil, im Jahre 1989 abgespeichert (Abb. 3).

Die Formatierung ist noch grob erkennbar, die Schrift stimmt nicht mehr. Die Sonderzeichen sind verloren gegangen. Immerhin sieht man, dass die ASCII-7 Kodierung relativ stabil ist. Der Text ist noch einigermaßen lesbar. Das Programm, mit dem diese Datei erstellt wurde (vermutlich DOS-Word 4.0) besitze ich noch, nur befindet es sich auf 5 1/4 Zoll Disketten, die unmittelbar nicht mehr viel nutzen.

Wir sprechen hier von einem Zeitraum von knapp 20 Jahren und man kann sich leicht ausmalen, welche Entwicklungen uns noch bevorstehen.

Leider ist im Bereich der Langzeitarchivierung elektronischer Dokumente noch kein endgültiger Durchbruch erzielt worden. Doch existieren grundsätzliche Strategieüberlegungen, die in Projekten wie Nestor³ jetzt auf ihre Praktikabilität hin untersucht werden, doch kann von einer breiten Anwendung dieser Strategien noch keine Rede sein.

Problematisch ist, dass die Aufmerksamkeit innerhalb der Geistes- und Kulturwissenschaften gegenwärtig überwiegend den literarischen Erzeugnissen gilt, während die

3 <http://www.langzeitarchivierung.de>.

Dokument wird untersucht. Im Laufe der Zeit hat es einige Migrationsschritte erlebt. Die letzte Konversion in ein modernes Format hat vor 20 Jahren mit der generellen Umstellung des Archivs vom magnetischen auf das holografisch-optische Speichersystem stattgefunden und liegt im Format *verbum 6* vor. Welche Migrationsschritte im Einzelnen stattgefunden haben, ist nicht mehr nachvollziehbar. Die Dokumentation ist beim letzten großen, durch einen Blitzschlag verursachten Servercrash vor 50 Jahren, der auch die Sicherungsbänder betraf, zugrunde gegangen. Heute wird alles redundant an zahlreichen Orten archiviert und laufend automatisch aktualisiert. An jedem Monatsende entsteht zudem ein elektrisch unabhängiger Gesamtabzug der Datenbanken in chemisch inerten PVC-Speichern.

Im Zuge der Nachforschungen, wer das elektronische Buch genutzt haben könnte, werden neben einigen Kopien im Supranet, einem Nachfolger des alten Internet, im dem alle eingestellten Dateiinhalte automatisch indexiert und semantisch suchbar gemacht werden, im Modul Annotationsmaterialien der HAB Dokumente zu diesem Titel gefunden, die möglicherweise interessante Hinweise beinhalten und auf einen zeitgenössischen Nutzungskontext schließen lassen. Die gefundene Datei lässt sich jedoch nicht unmittelbar in eines der bekannten Textverarbeitungsprogramme laden, automatisch gestartete, vorhandene Emulatoren versagen ihren Dienst. Der bibliothekarische Sachverständige erläutert, dass der Text aus der Frühzeit der Archivierung elektronischer Dokumente stammt, in der es noch keine systematische Verzeichnung der technischen Metadaten gab. Eine nähere medienarchologische Untersuchung auf der Byteebene der fraglichen Datei ergibt schließlich, dass es sich um ein Format namens WORD 6.0 handelt. Ein schneller Blick in die Datenbank belehrt: es handelt sich um ein um die Jahrhundertwende geläufiges Format, das durch ein heute unbekanntes Programm namens MS-Word verarbeitet wurde, das auf einem heute ebenso unbekanntem Betriebssystem Windows lief.

Nun gelingt schon einmal eine Grobdatierung, der Text ist zeitgenössisch. Mit Hilfe dieser Information kann unter Hinzuziehung der Zentralstelle für Programmulation in Frankfurt ein System aufgesetzt werden, das die ursprüngliche Betriebssystemumgebung rekonstruiert und erlaubt die Datei anzuzeigen. Der Verfasser ist anhand des Textes schnell ermittelt. Es handelt sich um einen berühmten Gelehrten, der in seiner Jugend eine Zeitlang in Wolfenbüttel geforscht und anscheinend diesen Text auf dem Server hinterlassen hat, der wiederum später in das Supranet übertragen wurde. Aus dem Text ergeben sich interessante Hinweise auf das elektronische Dokument in BB. Nachdem die Zeichencodierung dieses vorzeitlichen Formats bekannt ist, führt eine globale Suche über den Archiverserver zu weiteren Belegen für die Nutzung des Dokumentes durch den Gelehrten. Weitere Hinweise auf sog. Emaillkontakte, ein früher geläufiges, heute nicht mehr gebräuchliches Kommunikationsverfahren, werden nach Überprüfung der persönlichkeitschutzrelevanten Faktoren im elektronischen Zentralarchiv für elektronische Autographen in Berlin systematisch verfolgt und lassen schließlich eindeutige Zuordnungen von Text und Leser zu.

c) Resümee

Wenn es sich so oder so ähnlich zutragen wird, dann bleibt Provenienzforschung auch im elektronischen Zeitalter lebendig. Voraussetzung ist, dass bei den Provenienz sichernden Institutionen ein Bewusstsein dafür entsteht, welche Schritte zu unternehmen und welche Art elektronische Dokumente zu sichern sind. Dabei scheint eine Mischung von Migrations- und Emulationsverfahren ein sinnvoller Weg zu sein. Dokumente, die dauernd in Benutzung oder die problemlos ohne Datenverlust konvertierbar sind, wie z. B. Imagedaten, sollten migriert werden. Eine Vielzahl von Emulatoren zu bevorzugen, die dann jeweils einzusetzen bzw. zu programmieren wären, wäre vermutlich wenig praktikabel. Dabei muss von Fall zu Fall überprüft werden, ob das migrierte Endprodukt noch alle Informationen enthält wie die migrierte Quelle. Das gilt insbesondere auch für die nicht sichtbaren Dateiinformationen. So mag es interessant sein, den Zeitpunkt und den Ort der Entstehung eines Images zu kennen. Ist dieses Image jedoch in ein neues konvertiert worden, lässt sich dies an der Datei möglicherweise nicht mehr ablesen, wenn die Metadaten im Header der Datei nicht erhalten wurden. Es bedarf dann zusätzlicher Metadaten.

Gewissermaßen beiseite gelegte Dokumente sollten emuliert werden. Dabei erscheint es nicht sinnvoll, wenn jede Einrichtung sich separat um die Emulation kümmert, es genügt, wenn zunächst einmal die Daten mit entsprechenden Metadaten versehen als Bitstream archiviert werden. Vorhandene untypische Formate oder Programme sollten an eine Zentralstelle gemeldet werden, die dann für die passenden Emulatoren sorgt. Von dort fließen diese frei verfügbar an die Nutzer zurück. Es obläge den jeweiligen Einrichtungen, ihren Nutzern auf passenden Geräten diese Emulationstechnik bereitzustellen.

Aus diesen nur skizzenhaften Überlegungen erhellt, dass Provenienzforschung zukünftig in steigendem Masse von äußerlichen, dokumentenexternen Dingen abhängen wird. Die Entmaterialisierung der Kopie verhindert weitgehend eine an das Substrat gebundene Provenienzermittlung, nicht mehr die Textura des Papiers, sondern der elektronischen Netze wird Provenienzzusammenhänge sichtbar werden lassen. So wandelt sich die Provenienzermittlung zu einer Art digitaler Archäologie, die hinter der Präsentationsoberfläche der elektronischen Texte selbst, mehr aber noch in den Kontexten dieser Texte, in den dislozierten Paratexten forscht.