

Kai Naumann

## **Ungelöstes Problem oder ignorierte Aufgabe? Web-Archivierung aus der Sicht deutschsprachiger Archive**

Welchen Stellenwert hat die Archivierung von Webseiten für die Überlieferungsbildung in deutschsprachigen Archiven, welche Handlungsmöglichkeiten stehen den Archivarinnen und Archivaren offen? Die vorliegende Ausarbeitung behandelt unter diesem Leitsatz zunächst die Ergebnisse einer Befragung im Internet. Die im Rahmen der Umfrage genannten Hindernisse auf dem Weg zu einer Überlieferung und Nutzung von Webseiten werden aus der Sicht der Praxis bewertet. Anschließend werden aus Verfassersicht die wichtigsten Schritte bei der Überlieferungsbildung aus Webseiten benannt und der wichtigste Handlungsbedarf in den betroffenen Bereichen erläutert. Zuletzt wird der Frage nachgegangen, welche Vor- und Nachteile ein Engagement der Archive in der Web-Archivierung in den nächsten 100 Jahren haben könnte und in welchen Bereichen dieser komplexen Aufgabe die Einflussnahme sich besonders lohnen dürfte.

### I.

Die Befragung, die der Verfasser über den Dienst [surveymonkey.com](http://surveymonkey.com) vom 15.3. bis 8.4.2010 abgehalten hat, diente in erster Linie dem Ziel, eine Grundvorstellung davon zu bekommen, wie die Gedächtnisorganisationen im deutschsprachigen Raum zur Web-Archivierung stehen. Die Befragung bildet keine Grundgesamtheit ab und ist auch methodisch nicht perfektioniert, genügt aber der Zielsetzung. Der Link auf den Online-Fragebogen wurde über die Mailinglisten [archivliste](#), [demuseum](#), [inetbib](#), und das Weblog [Archivalia](#) verbreitet. An der Befragung nahmen 98 Personen teil, von denen sich 61 den Archiven zurechneten, 29 den Bibliotheken, 5 den Museen, 6 der Verwaltung und 8 der Wissenschaft.

Die Befragung sollte Aufschluss auf drei Fragen geben:

1. Betrachten die Gedächtnisorganisationen Web-Archivierung überwiegend als irrelevante Aufgabe oder als ungelöstes Problem?
2. Haben die Teilnehmer schon in einzelnen Bereichen angefangen?
3. Welche Gründe zum Nicht-Handeln sehen die Teilnehmer?

In dieser Untersuchung werden Aufgaben und Probleme methodisch unterschieden. Aufgaben unterscheiden sich von Problemen darin, dass eine Lösung bekannt ist. Viele

Probleme, die in der Alltagssprache so genannt werden, stellen sich eigentlich als schwer erfüllbare, aber erfüllbare Aufgaben heraus. Insgesamt verfolgte die Befragung das Ziel, die Gründe für zögerliches Handeln im Bereich der Web-Archivierung zu sammeln und zu bewerten, ob die Befragten diese Hindernisse zu Recht oder eher aus Unkenntnis als Problem betrachten.

Im Einzelnen wurde gefragt (in Klammern mögliche Antworten):

1. Würden Sie gern gemeinsam mit anderen Einrichtungen die Archivierung von Webseiten im Rahmen Ihres beruflichen Auftrags betreiben? (Ja / Nein)
2. Wann, denken Sie, werden Sie die Voraussetzungen zu dieser gemeinsamen Arbeit geschaffen haben? (2010 / 2015 / nach 2020 / nie)
3. Haben Sie für Ihre Überlieferungsbildung schon einmal ein Webangebot gespiegelt oder spiegeln lassen? (Ja / Nein)
4. Wann haben Sie vor, die Existenz dieser Spiegelungen (nicht die Spiegelungen selbst) öffentlich bekannt zu machen? (Erledigt / 2010 / 2015 / nach 2020 / nie)
5. Was ist Ihre Wunschvorstellung für die Überlieferungsbildung im Bereich von Webseiten? Welche Hindernisse gibt es auf dem Weg dorthin? (Freitext)
6. In welchem Bereich sind Sie tätig? (Art der Einrichtung, s.o.)

Die quantitative Auswertung von Frage 1 ergab, dass 83 % der Teilnehmer sich gern um die Archivierung bestimmter Webseiten kümmern würden. Schon bei Frage 2, die von 90 Teilnehmern beantwortet wurde, zeigte sich aber, dass die Aufgabe nicht für sehr viele Teilnehmer prioritär ist. Unter zehn Teilnehmern der Umfrage haben zwei schon angefangen, vier möchten mit dem Jahr 2015 in der Lage sein, Webseiten zu archivieren, und weitere vier glauben, erst nach 2020 oder gar nie handlungsfähig zu sein. Nimmt man darüber hinaus zur Kenntnis, dass nur ein winziger Teil der deutschen Archivare und sonstigen Gedächtnisfachleute an der Umfrage teilgenommen hat, wird man Web-Archivierung derzeit in der Fachwelt als nachrangige Aufgabe einstufen dürfen.

Mit den Antworten auf Frage 3 ergab sich der überraschende Befund, dass schon 40 Befragte ein Webangebot selbst oder durch andere gespiegelt hatten. In der deutschsprachigen Fachgemeinde sind demnach einige Institutionen bereits aktiv, haben erste Ergebnisse hergestellt. Im Gegensatz hierzu steht, wie die Antworten auf Frage 3 und 4 ergeben, die geringe Bereitschaft, über die Existenz erster Ergebnisse zu reden, nachdem man die ersten Schritte bereits vollzogen hat. Von vier Archivaren, die bereits praktisch Hand angelegt haben, sprechen drei nicht in der Öffentlichkeit nicht darüber. Warum übernehmen Archivare Webseiten gleichsam heimlich in ihre Bestände, ohne dies potenziellen

Nutzerkreisen bekannt zu machen? Kaum ein Presseorgan wird sich eine solche Meldung entgehen lassen, und das jeweilige Archiv kann sich mit einer öffentlichen Verlautbarung entgegen allen Klischees als moderne Einrichtung profilieren.

## II.

Möglicherweise ist die aus den Zahlen sprechende Vorsicht Ausdruck einer gewissen Schwellenangst. Viele möchten nicht zu den Pionieren gehören, die unbequeme Fehlschläge hinnehmen müssen. Andere sind vielleicht nicht hinreichend unterrichtet, welche Möglichkeiten ihnen offenstehen. Kommen wir daher zur qualitativen Analyse der Antworten auf Frage 5. Nach Ansicht der Teilnehmer stehen viele Hindernisse im Raum, die im folgenden – nach Themen sortiert – dargestellt werden. Im folgenden werden diese Hindernisse dargestellt und jeweils daraufhin bewertet, ob sie wirkliche Probleme oder nur schwierige Aufgaben darstellen.

Zunächst gab es diejenigen Stimmen, die Web-Archivierung als Aufgabe rundweg ablehnen oder gegenwärtig nicht als Aufgabe betrachten. Diese Einschätzung ist legitim. Ob sie opportun ist, wird noch einmal in Abschnitt IV beleuchtet. Zitate aus der Befragung:

- „Eines Archivs nicht würdig.“
- „Ich zweifle an dem historischen Wert.“
- „Eigene Arbeitsbelastung mit anderen Standard- und Kernaufgaben“
- „Mangelnde Mittel, mangelnde Zeit“
- „Es gibt so viele andere Dinge zu tun.“

Dann gaben viele Befragte an, die Ablieferungspflichtigen und sie selbst verstünden zu wenig von der Notwendigkeit und von den Möglichkeiten, Web-Publikationen auf Dauer zu erhalten:

- „Haupthindernis ist nach wie vor das Verständnis für die Problematik“
- „Hartnäckige Verständnislosigkeit des Archivträgers“
- „Das Hauptproblem besteht im Augenblick darin, dass sich die Produzenten ihrer Überlieferungs- und Anbietepflicht auch für Webseiten überhaupt nicht bewusst sind.“

Ebenso häufig war die Ansicht, dass geeignete Handlungsvorgaben noch fehlten.

- „klar definierte Standards“
- „Standards“
- „Es muss einfach sein (automatisierbar, komplexitätsreduzierend etc.).“
- „Wer ist für welche Webseiten zuständig.“

Zur Bewertung dieser Argumente ist festzuhalten, dass ein Mangel an Bewusstsein und Handlungswissen nicht als unlösbar angesehen werden kann. Verständnis zu wecken und Praxiserfahrung zu erwerben, ist eine Aufgabe, kein Problem. Zur Webarchivierung gibt es Lehrbücher<sup>1</sup>, Aufsätze, Bibliographien, laufende und abgeschlossene Forschungsprojekte sowie Publikationen und Foren im Netz<sup>2</sup>. Große Staaten wie Großbritannien und die USA, aber auch kleine Nationen wie Portugal und die Tschechische Republik unterhalten Web-Archivierungs-Programme, die gerade die Web-Überlieferung der staatlichen Behörden auf allen Ebenen zu ihren zentralen Aufgaben zählen. Ein niederländisches kommunales Rechenzentrum bietet in Zusammenarbeit mit einer britischen Firma einen Web-Archivierungsdienst für Städte und Gemeinden an<sup>3</sup>. Auch Weiterbildungsangebote existieren bereits<sup>4</sup>.

Natürlich wurden auch die Finanzen als weiteres Hindernis genannt. Geld und Personal sind Voraussetzungen, um eine Aufgabe umzusetzen. Das Einwerben neuer Mittel erscheint vielen unrealistisch.

- „Geld- u. Personalprobleme“
- „Hindernis: Kommunal финанzen“
- „ungeklärte Kostenfrage“
- „der notwendige Formatwandel und Erschließungsbedarf frißt zu viele Ressourcen.“

Eine Finanzierung zu ermöglichen, ist harte Arbeit – aber kein Problem im eigentlichen Sinne. Zunächst ist es allerdings erforderlich, neben bestehenden Aufgaben Zeit zur Planung und zur Gewinnung einer Finanzierung zu erübrigen. Das ist eine Aufgabe für Langstreckenläufer, die wissen, dass die Anstrengung sich eben erst nach einigen Jahren des Einsatzes lohnt.

Eine weitere Gruppe von Befragten brachte rechtliche Hindernisse vor.

- „Abklärung der urheberrechtlichen Problematik“

---

<sup>1</sup> Lehrbücher sind: A. Ball: Web Archiving (Version 1.1), (DCC State of the Art Report) Edinburgh 2010, online verfügbar unter <http://dcc.ac.uk>; A. Brown: *Archiving Websites: a practical guide for information management professionals*. London 2006; J. Masanès (Hg.): *Web Archiving*. Berlin 2006.

<sup>2</sup> Die beste bibliographische Basis ist PADI: <http://www.nla.gov.au/padi/topics/92.html>. Eine gute Übersicht liefern die Wikipedia-Artikel auf Englisch und Deutsch. Mailinglisten laufen unter <https://listes.cru.fr/sympa/info/web-archive> und <http://www.netpreserve.org/about/curator.php>.

<sup>3</sup> Archiefweb, ein Dienst von Gemeenteweb: <http://archieffweb.eu/>.

<sup>4</sup> Vgl. die Seminare des Vereins für Medieninformation und Mediendokumentation: <http://www.vfm-online.de>.

– „Es muss Compliance-Anforderungen erfüllen.“ (d.h. Beweiswert vor Gericht im Handels- und Steuerrecht, KN)

– „Urheberrecht als Haupthindernis“

In der Tat sind Rechtsfragen beim Archivieren von aktuellen Netzpublikationen relevanter als beim Umgang mit jahrzehntealten Materialien, deren rechtlicher Status zwar oft unklar ist, mit denen man als Archivar aber alltäglich ohne allzu große Sorgen umgeht, weil man ihnen aus Altersgründen in der Regel keinen besonders hohen Streitwert mehr beimisst.

Allerdings sind zur Überwindung der rechtlichen Hindernisse Lösungen bekannt. Zum einen hält das Urheberrecht speziell für Archive und andere Gedächtnisinstitutionen sogenannte Schranken bereit, die wesentliche Schritte der Web-Archivierung möglich machen, sofern man sich an gewisse Einschränkungen der Verfügbarkeit hält<sup>5</sup>. Zum anderen stellt sich oft heraus, dass man bei kleineren Unklarheiten in gesetzlichen Regelungen gut damit beraten sein kann, es auf einen Versuch ankommen zu lassen und nicht auf eine Regelung durch Gesetz- und Verordnungsgeber zu warten.

Als Beispiel kann man Bilanz über Rechtsfragen ziehen, die beim Landesarchiv Baden-Württemberg im Rahmen des Baden-Württembergischen Online-Archivs BOA<sup>6</sup> von Dritten aufgeworfen wurden. In BOA werden derzeit 64 Web-Präsenzen von Behörden und Einrichtungen des Landes gemäß Landesarchivgesetz archiviert, die Spiegelungen sind vollständig öffentlich zugänglich und werden von Google indexiert. Ursprünglich wurde das Urheberrecht in diesem Zusammenhang als schwierigster Aspekt betrachtet. Um sich gegen urheberrechtliche Forderungen abzusichern, ließ das Landesarchiv über die für Web-Präsenzen des Landes zuständige Stelle beim Innenministerium einen Leitfaden zum Urheberrecht veröffentlichen, um ein Problembewusstsein zu schaffen<sup>7</sup>. Zudem wurden alle betroffenen Stellen schriftlich von der Archivierung unterrichtet.

Nach drei Jahren und acht Monaten Betrieb von BOA stellt sich aber heraus, dass das Urheberrecht bisher nicht Gegenstand eines Antrags war. Von drei in dieser Zeit gestellten Anträgen auf Sperrung oder Veränderung der Spiegelungen bezogen sich zwei auf das Persönlichkeitsrecht, also auf Personennamen und Fotos von Mitarbeitern der betroffenen Institutionen, die nicht mehr im Internet sichtbar sein wollten. Nur einem Antrag musste stattgegeben werden. Die dritte Anfrage betraf eine irrtümliche Verwechslung: Ein veraltetes Formular war vom Webserver einer Behörde gelöscht worden, aber in BOA weiterhin

---

<sup>5</sup> Vgl. den instruktiven Bericht von Christiane von Nessen: Presseveröffentlichungen als Onlineangebot eines Stadtarchivs – Chancen und Grenzen, in: M. Stumpf/K. Tiemann (Hg.): Kommunalarchive und Internet, Münster 2009, S. 27-35, hier ab S. 31.

<sup>6</sup> <http://www.boa-bw.de>.

<sup>7</sup> [http://www.verwaltungsreform-bw.de/PUBLIKATIONEN/Documents/Bilder\\_und\\_Urheberrechte.pdf](http://www.verwaltungsreform-bw.de/PUBLIKATIONEN/Documents/Bilder_und_Urheberrechte.pdf).

verfügbar. Aus unerfindlichen Gründen warf eine Google-Suche damals das veraltete Formular noch vor dem neuen Formular aus. Der Finder druckte es und füllte es aus und stiftete Verwirrung in der Behörde, die sich dann an das Landesarchiv wandte. Die Lösung war hier eine Sperre auf Zeit, und Veränderungen am Server von BOA, die dazu führten, dass Google Suchergebnisse aus BOA unter „ferner liefern“ einstuft. Somit war die Verwechslungsgefahr gebannt.<sup>8</sup>

Auch Anforderungen der Rechtssicherheit nach Handels- und Steuerrecht, der sogenannten „Compliance“, können von Web-Archivierungs-Systemen grundsätzlich eingehalten werden. Spiegelungen können, sofern sie als definierte Bitfolge abgelegt sind, ohne weiteres digital signiert werden. Es ist nicht aber einmal in jedem Rechtsfall nötig, die strengen Regeln der Behörden einzuhalten, denn jedes digitale Objekt kann der freien richterlichen Beweiswürdigung unterliegen.

Zusammenfassend ist zum Rechtsbereich festzustellen, dass tatsächlich einige Rechtsfragen den Status von Problemen erreichen, da die Lösung unter Juristen teilweise umstritten, also nicht klar bekannt ist. In vielen, ja den meisten Bereichen jedoch stellt unsere Rechtsordnung, wie einige bereits bestehende Projekte in Deutschland und Europa zeigen, zur Einrichtung eines Web-Archivs bereits wesentliche Lösungen bereit.

Viele Befragte waren schließlich auch der Ansicht, die Technik sei noch nicht hinreichend ausgereift.

- „Probleme machen Datenformate und Speichermengen“
- „technische Umsetzung unklar; Software noch nicht vorhanden“
- „noch kein Stand Alone Viewer“
- „die Frage der dauerhaften Sicherung und Zugänglichkeit ist leider noch nicht gelöst“

Hierzu ist erstens festzustellen, dass in diesem sehr kleinen Geschäftsfeld zweifelhaft ist, oft die zu erwartenden Gewinne viele Unternehmer dazu motivieren können, erheblich in die Fortentwicklung bestehender Anwendungen zu investieren. Web-Archivierung gibt es in großem Maßstab seit 1996<sup>9</sup>. Gut marktgängige Technologien sind nach 15 Jahren voll ausgereift, was auf die Web-Archivierung nicht zutrifft. Der Verfasser hat im Gespräch mit

---

<sup>8</sup> Nachtrag kurz vor Drucklegung des Beitrags: Inzwischen ist auf <http://www.boa-bw.de> für jeden Nutzer eine Anleitung verfügbar, wie man selbstständig Suchergebnisse über die vom Landesarchiv verantworteten BOA-Inhalte, die in Google indexiert wurden, aus der Google-Suche löschen kann. Diese Maßregel ist ein zwar nicht archivgesetzlich normierter, aber praxisnaher Ausgleich zwischen den Interessen der Allgemeinheit an der Erhaltung staatlicher Unterlagen und den Interessen des Einzelnen am Schutz seiner Privatsphäre.

<sup>9</sup> Im Jahr 1996 nahm das Internet-Archiv seine Tätigkeit auf: <http://www.archive.org>.

Web-Experten den Eindruck gewonnen, dass die Technologie der Web-Archivierung kaum weiter ausreifen wird, weil sie stets dem rapiden Wandel in der Web-Technologie hinterherlaufen muss<sup>10</sup>. Zweitens ist die Meinung, es gebe noch keine Lösungen für Sicherung und Zugänglichkeit, nicht wahr: es existieren zahlreiche Antworten und zahlreiche Lösungen, die man sich nur zu eigen machen muss<sup>11</sup>. Drittens ist bei den Speicherkosten auch in Zukunft mit stetig fallenden Preisen zu rechnen.

Wem die etwas künstliche Unterscheidung von Problem und Aufgabe inzwischen auf die Nerven geht, der kann aufatmen, da sie ab hier schon wieder irrelevant wird. In der Tat waren sich viele Befragte bewusst, dass ihre Alltagsprobleme eigentlich nur unbewältigte Aufgaben sind. Die Bewältigung setzt aber eine gemeinsame Willensbildung voraus, sie fällt in die Domäne der Politik. Innerhalb der Fachgemeinde wäre demnach zu entscheiden, was wann im Bereich Web-Archivierung für wieviel Geld unternommen wird, und die Aufgabe gemeinsam im Feld der Kultur- und Wissenschaftspolitik voranzutreiben.

- „Organisierte fachliche Kooperation“
- „Erledigung ... durch einen im öffentlichen Bereich angesiedelten Dienstleister“
- „Entwicklung eines Marktes für einschlägige Archivierungssysteme“
- „Es sollte eine zentrale Lösung für das Archivieren geben, die (gegen Kostenbeteiligung) buchbar ist.“
- „Kooperative Überlieferungsbildung im Verbund mit anderen Gedächtnisorganisationen, gemeinsame technische Lösungen und Konzepte zur Langzeitverfügbarkeit.“

### III.

Im folgenden Abschnitt werden zu Themen der Web-Archivierung einige aus der subjektiven Sicht des Verfassers wichtige Fragen formuliert. Zunächst das Thema der Bewertung, das wohl aus Sicht der Archivwelt zur Zeit das wichtigste ist. Internetpräsenzen sind für den heutigen Erstnutzer behördlicher Leistungen die erste Anlaufstelle. Insofern

---

<sup>10</sup> So auch R. Schmitz: Selektive Webarchivierung – Auswahl und Bewertung bei der Archivierung von Webpräsenzen, in: M. Stumpf/K. Tiemann (Hg.): Kommunalarchive und Internet, Münster 2009, S. 81-87, hier S. 84.

<sup>11</sup> Beispiele für existierende Software sind heritrix (<http://crawler.archive.org/>), Hanzo Engine (<http://www.hanzoarchives.com/>), HTTrack (<http://www.httrack.com>), NutchWAX (<http://archive-access.sourceforge.net/projects/nutch>), Offline Explorer (<http://www.metaproducts.com>), OWA (<http://www.oia-duesseldorf.de/>), wayback (<http://archive-access.sourceforge.net/projects/wayback/>), WARC tools (<http://code.google.com/p/warc-tools/>).

bieten sie sich als wesentliche Quellenbasis an, wenn die Tätigkeit einer Einrichtung nach vielen Jahrzehnten nachzuvollziehen ist. Sie verändern sich dabei in besonders hohem Maße. In einer Zeit, die sehr häufige Veränderungen der Behördenstruktur kennt, sind die Webseiten dieser Behörden einem noch häufigeren Wandel unterworfen. Es gibt URLs nicht nur für Behörden und landeseigene Betriebe, sondern auch für interministerielle Arbeitsgruppen, für Planungsverfahren mit Beteiligung des Landes und der Gemeinden, für Runde Tische mit Bürgerbeteiligung. Hier liegt für die Archive eine potenzierte Herausforderung: seit vielen Jahren bekannt ist uns die Schwierigkeit, Provenienzen mit verändertem Zuschnitt, zwischen den Ressorts wechselnde Zuständigkeiten in der Beständeübersicht nachvollziehbar abzubilden. Nun kommen die vielen beweglichen Internetpräsenzen und beschleunigen noch die Fluktuation.

Insofern scheint die Zeit gekommen, über kollaborative Bewertungsmodelle nachzudenken. Technisch sind die bisherigen Bewertungsmodelle Textdokumente mit einem oder mehreren Verfassern, die eventuell von Zeit zu Zeit aktualisiert werden. Sie sind fast immer nur in einem Bundesland gültig. Im Zeitalter des Internet wird sich das Verfassen von Bewertungsmodellen aber zweckmäßigerweise an den kollaborativen Möglichkeiten des Web 2.0 orientieren. Ein Bewertungsmodell gleiche einem Wiki mit einem Autorenkreis aus den Archivsparten der beteiligten Verwaltungsebenen. Eine solches Wiki sollte – möglichst bundesweit, vielleicht in Länderpartnerschaften – Tätigkeitsbereiche bewerten, aber auf die Möglichkeit bieten, bestimmte Institutionenbezeichnungen im Wandel der Zeit aufzuzeichnen, bestimmte URLs vorzugeben und deren Übernahmehäufigkeit zu bestimmen. Solch eine Anwendung fehlt. Wie anders aber wird man dem beschleunigten Wandel unserer Behördenwelt im 21. Jahrhundert beikommen können?

Die Politik bestärkt nicht gerade die Bestrebungen, eine kooperative Bewertung des Internet zu leisten. Das Gesetz über die deutsche Nationalbibliothek und die Pflichtexemplargesetze der Länder, die bereits eine Ablieferungspflicht für Netzpublikationen vorsehen, betrauen meist nur wenige große Bibliotheken in staatlicher Trägerschaft ausdrücklich mit der Sicherung und Erhaltung von Webseiten. Sicherlich werden sich auch die meisten Archivgesetze ihrem Auftrag zur Überlieferungsbildung nach auf Webseiten anwenden lassen. Dennoch besteht die Gefahr, dass die Aufgabe einseitig den großen, vorwiegend bibliothekarisch ausgerichtete Einrichtungen zuerkannt wird, deren

Möglichkeiten zur kollaborativen und kooperativen Bewertung unterentwickelt sein werden<sup>12</sup>. Da aber der kommunale Bedienstete vor Ort wohl zweifellos die besten Aussichten hat, die für seine Gemeinde wesentlichen, historisch wertvollen URLs zu definieren, könnte somit ein Überlieferungsverlust drohen, wenn die Aufgabe einseitig von den „großen Spielern“ übernommen wird.

Kommen wir zur Übernahme, Qualitätssicherung und Erhaltung. Webseiten sind binär codierte Informationen, die von Server und Browser als eine von Menschen bedienbare Anwendung dargestellt werden. Webseiten sind keine starren Informationsträger, sondern interaktive Konstruktionen. Sie haben – salopp gesagt – ein Eigenleben. Für die Archivierung müssen sie in eine dauerhaft erhaltungsfähige Form gebracht werden. Das vorherrschende Verfahren der Spiegelung überführt das Trägersystem „Webserver“ in das Trägersystem „Dateisystem“. Hierbei geht zwangsläufig ein definiertes Maß an Eigenleben verloren. Da dieser Verlust von den verwendeten Serversystemen abhängt, kann er im einen Fall winzig, im anderen erheblich sein. Wie definieren wir Archivare, wieviel verloren gehen darf, in anderen Worten: wieviel Verlust können wir uns leisten? Wie gewinnen wir einen Überblick über die Möglichkeiten der Qualitätssicherung? Wie definieren wir Erhaltungspakete (Archival Information Packages) für Spiegelungen einer URL?<sup>13</sup> Wie lagern wir die Spiegelungen so, dass wesentliche Eigenschaften von Webseiten wie Suche, Navigation und Multimedialität erhalten bleiben?<sup>14</sup>

Die Nutzung erfolgt unter den Voraussetzungen der vorgenannten Schritte. Wie klären wir unsere Nutzer darüber auf, dass es bestimmte einkalkulierte Einschränkungen gibt, sogenannte Quirks<sup>15</sup>, die toleriert werden müssen? Wird eine Suchfunktion über alle Inhalte benötigt? Welche Lizenzkosten für Emulatoren sind langfristig bezahlbar? Wie schirmen wir urheberrechtlich relevante Inhalte vor unerlaubter Nutzung ab?

#### IV.

Zum Abschluss dieser Ausführungen ein Gedankenspiel. Nutzerin Meier recherchiert heute über Hans Schulz im Jahr 1910 in Bibliotheken und Archiven. Dort nimmt sie Bücher, Graue

---

<sup>12</sup> Erste Entwicklungen dieser Art fanden bereits im Saarland statt, wie berichtet in: I. C. Becker: Archivierung kommunaler Websites – Bewertungsgrundlagen, in: M. Stumpf/K. Tiemann (Hg.): Kommunalarchive und Internet, Münster 2009, S. 93-99.

<sup>13</sup> Derzeit sind die vom Internet Archive und der International Internet Preservation Consortium IIPC entwickelten Formate ARC und der Nachfolger WARC hierfür im Gespräch.

<sup>14</sup> Hier steht vor allem auch die Alternative zwischen Lagerung in einem Dateisystem und auf einem Datenbankserver im Raum.

<sup>15</sup> Quirk (engl.) bedeutet Eigenart, Marotte.

Literatur und Nachlässe in Augenschein, hier sieht sie Verwaltungsakten und Nachlässe ein. Sie kommt auf 70 Nutzertage in Bibliotheken, 30 Tage in Archiven. Springen wir nun ins Jahr 2110 und nehmen an, deutschsprachige Archive haben sich 100 Jahre lang nicht an Web-Archivierung beteiligt. Nutzer Müller recherchiert nun im 2110 über Heike Schultz im Jahr 2010. In Bibliotheken nutzt er digitalisierte Bücher, digitale Nachlässe (Persönliche Homepages, Blogs, Tweets) und Webseiten von Institutionen. In den Archiven begegnen ihm noch einige papierne Verwaltungsakten. Ein papierner Nachlass ist nicht vorhanden, da Heike Schultz ihre Texte nicht mehr ausdrückte. Müller kommt auf nur noch zwei Nutzertage in Archiven, der Rest findet in Bibliotheken statt. Im Gedankenspiel haben die Archive also 93 Prozent Marktanteil verloren. Ein Einstieg in die Archivierung von Webseiten wäre damit ein offenkundiger Vorteil. Uns winkt die Möglichkeit, in Zeiten von eGovernment und Netzkultur das Gedächtnis der Verwaltung und der Gesellschaft zu bleiben.

Es gibt aber auch weitere Vorteile. Web-Archivierung ist nicht nur Archivierung des Internet, sondern auch Archivierung mit Internet-Technologien. CSS, JAVA, HTTP, PDF/A sind Web-Technologien, die vom „Technikmanager“ Archivar auch für die Archivierung von DMS-Systemen, GIS oder Fachanwendungen ansatzweise verstanden werden müssen. Und nicht zuletzt ist ein wichtiger Posten im Budget bei Webseiten geringer als bei Papierunterlagen: der Verzeichnungsaufwand ist wesentlich geringer, denn eine Klassifikation und Titelaufnahmen ergeben sich aus der Struktur der Seiten.

Die Nachteile seien nicht verschwiegen. Noch stärker als bei anderen digitalen Unterlagen wirkt sich die Notwendigkeit aus, mit Dienstleistern und kooperativ zu arbeiten. Damit endet das uralte Prinzip der Sachherrschaft des Archivs über seine Unterlagen. Unter dem Schlagwort „Cloud Computing“ steht unsere Zunft vor ganz neuen Herausforderungen bei der Sicherung von Integrität, Authentizität und dem Schutz vor unberechtigter Nutzung<sup>16</sup>. Ebenso nachteilig ist das Fehlen von Kennzahlen für die Bestandserhaltung. Aber: hätten wir ab 1860 jedes säure- und ligninhaltige Papier an der Schwelle des Archivs ablehnen dürfen? In diesem Sinne hofft der Verfasser bei den Lesern neues Wissen hinterlassen und vielleicht eine neue Einstellung zum Thema geweckt zu haben.

---

**16** Die britische Society of Archivists beauftragte im März 2010 die Universität von Aberystwyth mit einer Umfrage zum Thema Cloud Computing und bereitet eine Handreichung für ihre Mitglieder zu diesem Thema vor: <http://www.dil.aber.ac.uk/>.