

## Schritt für Schritt auf neuen digitalen Wegen

### Webseiten- und Social-Media-Kanal-Archivierung im Stadt- und Stiftsarchiv Aschaffenburg

Johannes Schuck – Stadt Aschaffenburg, Stadt- und Stiftsarchiv

1969 wurden erstmals zwei Computer via Telefonleitung miteinander verbunden. Ziel war es, Forschungsergebnisse zwischen zwei Universitäten auszutauschen. Zunächst konnten nur einzelne Buchstaben gegenseitig ausgetauscht werden, doch schon kurze Zeit später waren es ganze Texte. Nach und nach wurden weltweit immer mehr Computer miteinander verbunden. Das stetig wachsende Netz von verbundenen Computern erhielt den Namen Internet. Darüber konnten in Echtzeit – wie E-Mails heute – Nachrichten von einem Ort zu einem anderen geschickt werden.

Ein nächster bahnbrechender Schritt zum heutigen Internet gelang 1990 Tim Berners-Lee mit seiner Idee zum World Wide Web. Diese sah eine Art Speicher vor. Nun war es nicht nur möglich, Nachrichten zu versenden, sondern auch Informationen zu speichern. Und jeder konnte von einem „Internet-Computer“ aus diese Informationen abrufen. Von nun an entwickelte sich das World Wide Web immer schneller. Mittlerweile nutzt mehr als die Hälfte der Erdbevölkerung das Internet. Weder 1969 noch 1990 haben die „Erfinder“ wohl damit gerechnet, dass miteinander verknüpfte Computer die Welt derart verändern und beschleunigen würden. Zahlreiche Menschen surfen heutzutage täglich via Smartphone, Tablet oder am Computer im Internet. Sowohl privat als auch beruflich.<sup>1</sup>

#### Die digitale Transformation in der Berufswelt – Informationsverlust für Archive?

Seit der Jahrtausendwende setzt ein neuer – rein digitaler – Veränderungsprozess ein, der die Arbeits- und somit auch die Archivwelt grundlegend verändert hat bzw. zukünftig weiterhin maßgebend prägen wird. Die klassische analoge Ablage rückt fortschreitend in den Hintergrund. Informationen werden zunehmend rein elektronisch verarbeitet, verwaltet und gespeichert. Mehr und mehr müssen elektronische Datensätze immer und möglichst überall für Mitarbeiter\*innen einer Stadtverwaltung zur Verfügung stehen. Insbesondere dort, wo Arbeitsplätze verstärkt mobil werden, und der klassische Büroalltag mehr und mehr verschwindet.

Dies setzt zunehmend elektronische Verfahren und Systeme voraus, die das leisten können. Doch was passiert mit dieser schier endlosen Datenflut? Der größte Teil der Daten und Informationen kann in der Regel nach einer gewissen Zeit – beispielsweise

nach dem Ende der Aufbewahrungsfrist – gelöscht werden, da sie nicht mehr aufbewahrt werden müssen und als „Datenmüll“ auf Servern, Festplatten oder Clouds liegen. Doch was ist mit denjenigen digitalen Informationen, die sensibel und historisch wertvoll sind? Wichtig für nachfolgende Generationen, um das Handeln der Generation widerzuspiegeln. Diese müssen derart „archiviert“ werden, dass sie – analog zu Papier – noch in Jahren, Jahrzehnten und im Idealfall in Jahrhunderten ihre gespeicherten digitalen Informationen preisgeben können. Das ist für die gesamte Archivwelt eine ungeheure Herausforderung und setzt eine grundlegende Umstellung bisheriger Archivierungsprozesse voraus. Hier verschiebt sich nicht nur der Umgang mit der „Archivalie“, sondern auch die Anforderungen an die zuständige Archivarin oder den zuständigen Archivar sind ganz andere. Es rücken IT-technische Themen in den Fokus der Bestandserhaltung, die zuvor keinerlei Rolle gespielt haben. Selten hat ein neuer Datenträger die Archivwelt derart zu neuen Archivierungskonzepten, sowohl strategisch als auch technisch, gezwungen. Aus diesem Grund wurde 2019 im Stadt- und Stiftsarchiv Aschaffenburg die Stelle eines „Archivars für Informationsmanagement“ geschaffen, dessen Aufgabe es unter anderem ist, sich diesen neuen – rein digitalen – Quellen strategisch zu nähern und Archivierungslösungen zu entwickeln.

Aktuell laufen mehrere Projekte, die sich mit der fortschreitenden Digitalisierung des Büroalltags und der damit verbundenen neuartigen Archivierung auseinandersetzen. Basis dieser einzelnen Projekte ist eine zu Beginn erarbeitete „Digitalstrategie“ für das Stadt- und Stiftsarchiv Aschaffenburg, um sich strategisch der Speicherung, Verwaltung und Veröffentlichung von digitalem Content zu nähern – losgelöst von „klassischen Archivierungsansätzen“. Die Definition von Themenfeldern, Zielen und einer Veröffentlichungsstrategie standen hierbei im Vordergrund.

Resultat war die Aufteilung in drei thematisch unterschiedliche Säulen: Die Digitalisierung von und der digitale Zugang zu vorhandenen Altbeständen (beispielsweise via *bavarikon* oder *Deutsche Digitale Bibliothek*<sup>2</sup>), die Auseinandersetzung mit neuen, rein digitalen Informationsquellen und deren Archivierungsmethodiken (Fachverfahren, E-Akte, Dokumenten-Management-System, Digitale Langzeitarchivierung, Webseitenarchivierung) und die

Schaffung eines „Digitallabors“<sup>3</sup> bzw. einer gesonderten Webpräsenz<sup>4</sup>, um die Stadtgeschichte digital, barrierefrei und partizipativ zu präsentieren (Aschaffenburger 2.0<sup>5</sup>, Podcasts, Audioguide<sup>6</sup>).

### **Herausforderungen bei der Archivierung von Webseiten und Social-Media-Kanälen**

Einer dieser Bausteine zum digitalen Gesamtangebot des Stadt- und Stiftsarchivs ist die Archivierung von Webseiten und Social-Media-Kanälen. Im Laufe der Jahre sind diese Informationsplattformen zu „Schaukästen“ der Verwaltungsbehörden geworden, und somit ein neuer, rein digitaler Kanal ihrer Öffentlichkeitsarbeit. Aufgrund der Vielzahl an unterschiedlichsten Informationen, die über diese Medien an die Öffentlichkeit gegeben werden, rückt diese facetten- und informationsreiche Quelle mehr und mehr in den Fokus der Landes-, Kommunal- und Stadtarchive. Als mittlerweile etabliertes Kommunikations- und Publikationsmedium müssen sie das auch. Doch das birgt auch viele technische, rechtliche und strategische Herausforderungen für die jeweils zuständigen Archive.

Die Schnelllebigkeit durch die permanente Aktualität der Webseiten führt dazu, dass sich diese in kürzester Zeit nahezu vollständig verändern können, oder die veröffentlichten Informationen bereits hinfällig sind. Die durchschnittliche Lebensdauer einer Webseite beträgt laut „Internet Archive“ nur 77 Tage.<sup>7</sup> Demnach reicht eine einmalige Archivierung keinesfalls aus. Es bedarf eines nachhaltigen archivischen Konzeptes, um dieses digitale Medium – zumindest ansatzweise – mit all seinen gelieferten Informationen zu archivieren. Doch die Geschwindigkeit des Informationsflusses ist nur ein Teil der Herausforderung bei der Archivierung von Webseiten.

Die Bestandserhaltung von Webseiten und Social-Media-Kanälen wird insbesondere durch die immer komplexeren Strukturen und unterschiedlichen Datei-Formate erschwert. In der Regel besteht eine Webseite aus einem strukturierten Text (HTML), in den unterschiedlichste Medien (Videos, Bilder, Audios etc.) integriert sind, die über verschiedene Browser dargestellt werden.

Zumeist besteht eine Webseite aus einer sogenannten HTML-Datei, welche in einem Verzeichnis im Pfad einer speziellen individuellen Domain gespeichert ist. Dabei sind diese Dateien häufig in verschachtelten Verzeichnisstrukturen abgelegt. Die permanente Weiterentwicklung der unterschiedlichen Skript- und Programmiersprachen ermöglicht eine immer umfangreichere und komplexere Programmierung von Webseiten. Das ist für die Besucher\*innen der Webseite von Vorteil, da die Webseiten stets modern und neu gestaltet werden, für die Archivierung stellt dies jedoch eine enorme

Herausforderung dar. Hierfür werden Spezialprogramme eingesetzt, die in einem technisch komplexen Verfahren die zu archivierenden Webseiten und Social-Media-Kanäle „crawl“<sup>8</sup> und gesondert abspeichern.

All dies macht das Spiegeln und anschließende Speichern von Webseiten sehr komplex und für Archive nahezu unmöglich, da sich kaum ein Archiv eigene IT-Fachangestellte noch die Programmierung einer entsprechenden Software leisten kann. Demnach ist man auf Unterstützung angewiesen. Hilfreich ist es aber, wenn man sich im Vorfeld als Archiv grundlegende Gedanken über diese neue historische Quelle macht. Sollen neben Webseiten auch Social-Media-Kanäle archiviert werden und in welcher Art und Weise möchte man die neuen digitalen Archivalien als Archiv seinen Benutzerinnen und Benutzern zur Verfügung stellen? Derartige Grundsatzfragen nehmen erheblichen Einfluss auf das Verfahren der Archivierung und somit auch auf deren technische Umsetzung.

### **Eine Zeitreise in der digitalen Welt – Sonderprojekt „Webseitenarchivierung“**

Auch das Stadt- und Stiftsarchiv Aschaffenburg hat sich dieser neuen historischen Quelle angenommen und 2019 eine in das digitale Sammlungskonzept passende Archivierungslösung erarbeitet. Die Umsetzung erfolgte in einem internen Sonderprojekt „Webseitenarchivierung“, welches von der neu geschaffenen Stelle des Archivars für Informationsmanagement geleitet wurde.

Unsere Idee war es 2019, die archivierten Webseiten und Social-Media-Kanäle – losgelöst vom „klassischen Lesesaal“ – über eine gesonderte Webpräsenz allen Interessierten frei zugänglich und kostenfrei zur Verfügung zu stellen. Bei all diesen Überlegungen war uns sehr schnell klar, dass wir das städtische Webseitenangebot niemals vollständig archivieren können. Zunächst sind einige Seiten relevanter als andere und die bereits oben beschriebene Schnelllebigkeit lässt dies schlichtweg nicht zu. Es geht eher um gezielte und detailreiche Momentaufnahmen, die im Idealfall eins zu eins darstellen, wie dieser Moment aussah. Bei den Social-Media-Kanälen verhält sich dies wieder etwas anders. Hier ist es relativ einfach möglich, sämtliche Posts zu archivieren.

Um dies Realität werden zu lassen, wurde ein interner Kriterienkatalog erarbeitet, der ausschlaggebend für die weiteren Schritte war, um die Ziele einer aktuell rechtssicheren, vollständigen und voll funktionsfähigen Archivierung von Webseiten zu gewährleisten. Hierbei wurden archivintern die folgenden Standards gewählt:

- Vollständige Archivierung (Bild + Video)
- Schneller Zugriff auf gesonderter Plattform (für die Bereitstellung<sup>9</sup> für Nutzer\*innen)

- Rechtssichere Digitale Signatur (nach SHA1)
- Vertrauenswürdiger Zeitstempel (ANSI x9.95 Standard)
- Zugriffsrechteverwaltung für die archivierten Seiten
- Speicherung in einem langzeitarchivierungsfähigen Format (ISO-konform WARC-Format)

Erster Ansprechpartner für eine mögliche Umsetzung war das Amt für IT und Digitalstrategie der Stadtverwaltung Aschaffenburg. In den dafür angesetzten Besprechungen wurde aber deutlich, dass diese vom Archiv gestellten technischen Herausforderungen nicht in der Stadtverwaltung intern umgesetzt werden konnten. Daher war man auf externe Unterstützung angewiesen.

Doch wie und wo findet man den richtigen Partner? Es folgten Kontaktaufnahmen zu verschiedenen Archiven, auf Landes- und Kommunalebene, aber auch zu Wirtschaftsarchiven. Zudem wurden auch eigene Recherchen für Anbieter derartiger Spezialsoftware angestellt. So kristallisierten sich drei bis vier Anbieter mit unterschiedlichen Archivierungsansätzen heraus, zu denen ein intensiver Kontaktausgang aufgebaut wurde. Entscheidungsgrundlage hierbei war stets der vorab erarbeitete Kriterienkatalog mit den Anforderungen an die Softwareanwendung. Schließlich entschied sich das Stadt- und Stiftsarchiv Aschaffenburg für eine Zusammenarbeit mit dem damaligen Start-up „MirrorWeb“, welches in Deutschland durch die Firma Walter Nagel vertreten wird. Hier überzeugte insbesondere die technische Umsetzung der Archivierung, die Erfüllung aller seitens des Stadt- und Stiftsarchiv getroffenen Anforderungen zur Archivierung und der Gedanke, gemeinsam mit einem jungen Start-up etwas Neues aufzubauen. So konnten auch unsere individuellen Wünsche gut eingepflegt werden.

Die eigentliche Archivierung findet mithilfe der von MirrorWeb zur Verfügung gestellten Spezialsoftware nach einem festgelegten Zeitplan statt, losgelöst von den Mitarbeiter\*innen des Archivs ganz automatisiert. Dieser sieht vor, die Webseiten zweimal im Jahr ganzheitlich zu übernehmen. Hierbei werden die Webseiten gecrawlt und in ihrer Gesamtheit archiviert. Slideshows, Filme und Bilder werden ebenfalls übernommen. Dadurch ist im Nachgang die archivierte Webseite in der Form ihrer Veröffentlichung nutzbar und durchsuchbar. Somit verschwimmt der Eindruck, auf einer „historischen“ Seite zu surfen. Dies wird auf einer extra dafür eingerichteten Webpräsenz realisiert.<sup>10</sup> Die ausgewählten Social-Media-Kanäle werden seit dem ersten Beitrag auf dieser Plattform bis heute archiviert.

Über das sogenannte „Dashboard“<sup>11</sup> hat das Team des Stadt- und Stiftsarchivs die Möglichkeit, die Intervalle der Archivierung zu steuern, die Vollständig-

keit der archivierten Webseiten zu kontrollieren, die Zugriffsrechte zu verwalten (beispielsweise Ausblendung von Webseiten wie dem städtischen Intranet für die Öffentlichkeit), die Metadaten abzufragen, die Datensätze zu exportieren und vieles mehr. Es ist das zentrale Werkzeug, mit dem unsere Webseiten- und Social-Media-Kanal-Archivierung gesteuert und geplant wird. Auch Testcrawls perspektivischer Seiten, die in das Portfolio aufgenommen werden sollen, werden dort im Hintergrund verwahrt und von den Archivar\*innen überprüft. Dies ermöglicht es, vorab zu kontrollieren, welche Größe eine potenziell neue Webseite bei der Archivierung hat oder ob beispielsweise Besonderheiten beim Crawlen auftreten. Neuerdings ist es sogar möglich, hier Screenshots von den archivierten Webseiten zu erstellen, um diese via dem Dateiformat PDF oder PNG zu exportieren. Dies ist sicherlich für eine historische Gegenüberstellung in Fachaufträgen oder beispielsweise auf Social-Media-Kanälen eine gute visuelle Veranschaulichung. Auch der zukünftige Export der Archivpakete in ein digitales Langzeitarchiv lässt sich über das Dashboard planen, vorbereiten und anstoßen.

Aktuell werden elf Webseiten und vier Social-Media-Kanäle archiviert. Diese setzen sich aus zeitlich begrenzten Veranstaltungsseiten, allgemeinen Webseiten der Stadtverwaltung, dem Twitter-Kanal sowie dem YouTube-Kanal der Stadt Aschaffenburg zusammen. Die Kanäle Instagram und Facebook mussten dabei leider zunächst ausgeklammert werden, da zwar zu Beginn des Projektes 2019 eine technische Archivierung möglich gewesen wäre, jedoch rechtliche Hürden der Betreiber der Social-Media-Kanäle eine Archivierung verhinderten.

Diese Herausforderungen konnten behoben werden und so werden seit dem Frühjahr 2022 ausgewählte Facebook- und Instagram-Kanäle der Fachämter der Stadt Aschaffenburg archiviert. Ebenso wird die Palette der zu archivierenden Webseiten regelmäßig ergänzt, um das Angebot einer digitalen Zeitreise zu erweitern.

Auch wenn es technisch umsetzbar gewesen wäre, haben wir uns aktuell ganz bewusst gegen eine Archivierung von Kommentaren auf den Social-Media-Kanälen entschieden, da im Falle einer Archivierung Rechte Dritter betroffen wären. Für das Team des Stadt- und Stiftsarchivs Aschaffenburg war es wichtig, die Informationen und die Informationskanäle unseres Trägers zu archivieren und der Öffentlichkeit niederschwellig wieder zur Verfügung zu stellen. Dies verfolgen wir mit unserer aktuell umgesetzten Strategie.

Noch im Jahr 2022 werden wir insbesondere bei den Facebook- und Instagram-Kanälen die Kommentare mit einsammeln, diese verbleiben jedoch im Dashboard und werden nicht über unser

öffentliches Zugangsportale bereitgestellt. Trotz der Nichtveröffentlichung der Kommentare sichern wir dennoch die historische Diskussion zu den Posts der Stadt Aschaffenburg, halten diese aber so lange bewusst zurück, bis eine Veröffentlichung sowohl datenschutzrechtlich als auch archivrechtlich grundlegend geklärt ist.

Unsere Kriterien für die Auswahl der zu archivierenden Webseiten waren bzw. sind unterschiedlichster Natur. Neben der Informationsdichte und der Relevanz der veröffentlichenden Behörde spielen auch die Gestaltung und eine naheliegende Offline-Schaltung bzw. Modernisierung von Webseiten eine entscheidende Rolle. Ebenso wie die Frequentierung durch die Nutzer\*innen.

Auch das städtische Intranet war bzw. ist im Fokus des Stadt- und Stiftsarchivs. Hier findet im Frühjahr 2022 eine Ablösung des alten Intranets durch ein neues Social-Intranet statt. Daher bemühen wir uns hier, das alte Intranet zu sichern und für das neue Social-Intranet eine Archivierungslösung zu erarbeiten. Dabei ist zu berücksichtigen, dass diese historische Quelle nicht der Öffentlichkeit zugänglich gemacht wird, da dort Interna der Stadtverwaltung Aschaffenburg gespeichert werden.

Eine zentrale Frage bei der Umsetzung des Projektes „Archivierung städtischer Webseiten“ war, wie man die rein digitalen historischen Informationen unseren Nutzer\*innen zur Verfügung stellen kann: In unserem Lesesaal? Digitaler Zugriff von zu Hause? Wo bleibt die klassische Kontrolle der Archivar\*innen?

Fragen, die wir uns auch in anderen digitalen Projekten immer wieder stellen. Daher war es unser Ziel, eine gesamtheitliche und einheitliche Lösung zu finden.

Dies konnten wir realisieren, indem wir – mithilfe einer extra geschaffenen Webpräsenz<sup>12</sup> – themenbasiert, unentgeltlich und barrierefrei unseren Nutzer\*innen die neuen, rein digitalen historischen Überlieferungen zur Verfügung stellen. Hier hat auch die Webseitenarchivierung ihren eigenen Platz.

Somit steht einer Zeitreise durch die Historie der „Schaukästen“ nichts mehr im Wege!

Hingewiesen werden muss an dieser Stelle darauf, dass rein städtische Kanäle und Webseiten archiviert werden und daher die Stadt Aschaffenburg die Hoheit über die veröffentlichten Informationen hat und somit auch die Inhaberin sämtlicher Rechte ist. Dies vereinfacht sicherlich ein derartiges Veröffentlichungskonzept, wie es im Stadt- und Stiftsarchiv Aschaffenburg umgesetzt wird.

1. Vgl.: <https://www.zdf.de/kinder/logo/die-geschichte-des-internets-100.html> [letzter Zugriff: 12.01.2022].
2. Vgl.: <https://www.bavarikon.de/stadtarchiv-aschaffenburg> [letzter Zugriff: 21.12.2021].
3. Vgl.: <https://aschaffenburgzweimull.stadtarchiv-digital.de> [letzter Zugriff: 17.01.2022].
4. Vgl.: <https://stadtarchiv-digital.de> [letzter Zugriff: 22.01.2022].
5. Vgl.: <https://aschaffenburgzweimull.stadtarchiv-digital.de> [letzter Zugriff: 17.01.2022].
6. Vgl.: <https://audioguide.stadtarchiv-digital.de> [letzter Zugriff: 22.01.2022].
7. Vgl.: <https://archive.org/> [letzter Zugriff: 18.01.2022].
8. Als Crawler oder Webcrawler werden Softwareprogramme bezeichnet, die das Internet bzw. einzelne Webseiten durchsuchen. Hierbei analysieren und indizieren sie Inhalte von Webseiten, wie Texte, Audios, Bilder oder auch Videos. Erst diese Sammlung und Bündelung aller Informationen einer Webseite durch einen Crawler machen eine Archivierung möglich.
9. Vgl.: <https://webseitenarchiv.stadtarchiv-digital.de/> [letzter Zugriff: 23.01.2022].
10. Vgl.: <https://webseitenarchiv.stadtarchiv-digital.de/> [letzter Zugriff: 23.01.2022].
11. Eine grafische Benutzeroberfläche für die Visualisierung von Daten, für ein gezieltes Informationsmanagement.
12. Vgl.: <https://stadtarchiv-digital.de> [letzter Zugriff: 23.01.2022].