

## Schritt für Schritt auf neuen digitalen Wegen

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

von Johannes Schuck

Spätestens 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. Informationen werden zunehmend rein elektronisch verarbeitet und gespeichert. Mehr und mehr müssen elektronische Datensätze immer und möglichst überall für Mitarbeiterinnen und Mitarbeiter der Stadtverwaltung zur Verfügung stehen. Dies setzt zunehmend elektronische Verfahren und Systeme voraus, die das leisten können. Doch was passiert mit dieser Datenflut? Der größte Teil der Daten und Informationen kann in der Regel nach einer gewissen Zeit gelöscht werden, da sie nicht mehr gebraucht werden 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 konserviert werden, dass sie – analog zu Papier – noch in Jahren, Jahrzehnten und im Idealfall in Jahrhunderten ihre gespeicherten digitalen Informationen zur Verfügung stellen können.

Das ist für die gesamte Archivwelt eine ungeheure Herausforderung. Im Stadt- und Stiftsarchiv Aschaffenburg laufen gerade mehrere Projekte parallel, die sich mit der fortschreitenden Digitalisierung des Büroalltags und der Gesellschaft auseinandersetzen und jeweils ganz individuelle Herausforderungen an das Team der Archivarinnen und Archivare stellen.

Mit kleinen digitalen Schritten zu einem neuen Serviceangebot – so könnte man die Entwicklungen der letzten Monate und Jahre im Stadt- und Stiftsarchiv beschreiben.

Im Jahr 2017 begann die intensive Auseinandersetzung mit den unterschiedlichsten „digitalen Themen“. Ziel war es, eine „Digitalstrategie“ für das Stadt- und Stiftsarchiv zu entwickeln, um digitalen Content zu speichern und zu veröffentlichen – losgelöst vom „klassischen Archiv“. Die Definition von Themenfeldern, Zielen und einer Veröffentlichungsstrategie standen hierbei im Fokus.

Ergebnis war die Aufteilung in drei Themenbereiche: Die Digitalisierung und der digitale Zugang der vorhandenen Alt-Bestände (Beispielsweise via bavarikon<sup>1</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>2</sup> bzw. einer gesonderten Webpräsenz<sup>3</sup>, um die Stadtgeschichte digital, barrierefrei und partizipativ zu präsentieren (Aschaffenburg 2.0<sup>4</sup>, Podcasts, Audioguide<sup>5</sup>).

1 Vgl.: <https://www.bavarikon.de/stadtarchiv-aschaffenburg>  
[Stand: 21.12.2021, gilt ebenfalls für alle nachfolgenden Hinweise auf Internetseiten].

2 Vgl.: <https://aschaffenburgzweinull.stadtarchiv-digital.de>.

3 Vgl.: <https://stadtarchiv-digital.de>.

4 Vgl.: <https://aschaffenburgzweinull.stadtarchiv-digital.de>.

5 Vgl.: <https://audioguide.stadtarchiv-digital.de>.

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

Einer dieser Bausteine zum digitalen Gesamtangebot des Stadt- und Stiftsarchivs ist die Webseitenarchivierung. Im Laufe der Jahre sind Webseiten und auch Social-Media-Kanäle zu „Schaukästen“ der Verwaltungsbehörden geworden, und somit ein neuer rein digitaler Kanal ihrer Öffentlichkeitsarbeit. Aufgrund ihrer 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 diese sich 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>6</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 strukturiertem Text (HTML), in den unterschiedlichste Medien (Videos, Bilder, Audios etc.) integriert sind, die über verschiedene Browser dargestellt werden.

Im Idealfall besteht eine Webseite aus einer sogenannten HTML-Datei, welche in einem Verzeichnis im Pfad einer speziellen individuellen Domain gespeichert ist. Häufig sind diese Dateien aber in verschachtelten Verzeichnisstrukturen abgelegt. Die permanente Weiterentwicklung der unterschiedlichen Skript- und Programmiersprachen lässt eine immer umfangreichere und komplexere Programmierung von Webseiten zu. Das ist für die Besucherinnen und Besucher der Webseite von Vorteil, jedoch für die Archivierung eine enorme Herausforderung. Hierfür werden Spezialprogramme benötigt, die in einem technisch komplexen Verfahren die zu archivierenden Webseiten und Social-Media-Kanäle „crawlen“<sup>7</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 oder 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. Beispielsweise ob neben Web-

seiten auch Social-Media-Kanäle archiviert werden sollen, und in welcher Art und Weise man die neuen digitalen Archivalien als Archiv seinen Benutzerinnen und Benutzern zur Verfügung stellen möchte? 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 vom Archivar für Informationsmanagement geleitet wurde. Hierfür wurde eine neue Stelle geschaffen.

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. In 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 rechts-sichern, vollständigen und vollfunktionsfähigen Archivierung von Webseiten zu gewährleisten. Hierbei wurden die Standards bewusst hoch gewählt:

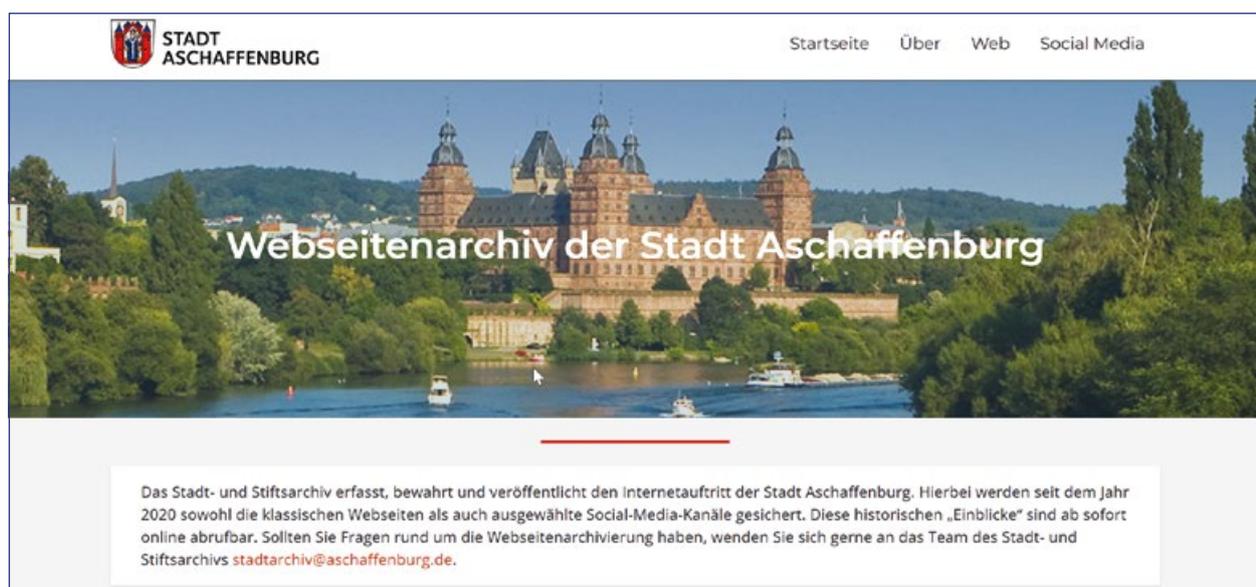
- vollständige Archivierung (Bild + Video),
- schneller Zugriff auf gesonderter Plattform (für die Bereitstellung<sup>8</sup> für Nutzerinnen und Nutzer),
- rechtsichere Digitale Signatur (nach SHA1),
- vertrauenswürdiger Zeitstempel (ANSI x9.95 Standard),
- Zugriffrechteverwaltung für die archivierten Seiten,
- Speicherung in einem langzeitarchivierungsfähigen Format (ISO-konform WARC-Format).

Erster Ansprechpartner für eine Umsetzung war die interne IT-Abteilung der Stadtverwaltung. Hier wurde aber sehr schnell klar, dass diese technische Herausforderung derart komplex war, sowohl in Bezug auf die notwendige Spe-

<sup>6</sup> Vgl.: <https://archive.org/>.

<sup>7</sup> Als Crawler oder Webcrawler werden Softwareprogramme bezeichnet, die das Internet bzw. einzelne Webseiten durchsuchen. Hierbei analysieren und indizieren sie Inhalte von Websites wie Texte, Audios, Bilder oder auch Videos. Erst diese Sammlung und Bündelung aller Informationen einer Webseite durch einen Crawler macht eine Archivierung möglich.

<sup>8</sup> Vgl.: <https://webseitenarchiv.stadtarchiv-digital.de>.



Webseitenarchivierung der Stadt Aschaffenburg (Screenshot)

zialsoftware als auch von Umsetzung innerhalb des städtischen Netzwerkes. Man war auf externe Unterstützung angewiesen.

Doch wie und wo findet man den richtigen Partner? Es folgten Kontaktaufnahmen zu den Archiven, wie Landes- und Kommunalarchiven, aber auch zu Wirtschaftsarchiven. Zudem wurden auch eigene europaweite Recherchen für Anbieter derartiger Spezialsoftware angestellt. Schlussendlich kristallisierten sich drei/vier Anbieter mit unterschiedlichen Archivierungsansätzen heraus, mit denen ein intensiver Kontaktaustausch stattfand. Grundlage hierbei war stets der vorab erarbeitete Kriterienkatalog. Schließlich entschied sich das Stadt- und Stiftsarchiv Aschaffenburg für eine Zusammenarbeit mit dem damaligen Startup „MirrorWeb“, welches in Deutschland durch das Unternehmen Walter Nagel vertreten wird. Hier überzeugte insbesondere die technische Umsetzung der Archivierung, die Erfüllung aller seitens des Stadt- und Stiftsarchiv getroffenen Anforderungen zur Archivierung gemäß der individuellen archivischen Grundsätze und der Gedanke, gemeinsam mit einem jungen Startup etwas Neues aufzubauen.

Die eigentliche Archivierung findet mit Hilfe der von MirrorWeb zur Verfügung gestellten Spezial-Software nach einem festgelegten Zeitplan statt, losgelöst von den Mitarbeiterinnen und Mitarbeitern des Archives 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 verschwindet der Eindruck, auf einer „historischen“ Seite zu surfen. Dies wird auf einer extra dafür eingerichteten Webpräsenz realisiert.<sup>9</sup> Die ausgewählten Social-Media-Kanäle werden von dem ersten Beitrag auf dieser Plattform bis heute archiviert.

Über das sogenannte „Dashboard“<sup>10</sup> hat das Team des Stadt- und Stiftsarchiv die Möglichkeit, die Intervalle der Archivierung zu steuern, die Vollständigkeit der archivierten Webseiten zu kontrollieren, die Zugriffsrechte zu verwalten (beispielsweise Ausblendung von Webseiten wie dem städtischen Intranet für die Öffentlichkeit), die Metadaten abzufragen und vieles mehr. Es ist der zentrale Ort, wo 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 Archivarinnen und Archivaren überprüft. Neuerdings ist es sogar möglich, hier Screenshots von den archivierten Webseiten zu machen, um diese via PDF oder PNG zu exportieren. 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, 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 aber mittlerweile behoben werden und so wird das Stadt- und Stiftsarchiv ab 2022 einige der Facebook- und Instagram-Kanäle der verschiedenen Fachämter der Stadt Aschaffenburg in das Portfolio der zu archivierenden Accounts aufnehmen. Ebenso wird die Palette der zu archivierenden Webseiten umfang-

<sup>9</sup> Vgl.: <https://webseitenarchiv.stadtarchiv-digital.de>.

<sup>10</sup> Eine graphische Benutzeroberfläche für die Visualisierung von Daten, für ein gezieltes Informationsmanagement.

reich 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 Stiftsarchiv Aschaffenburg war es wichtig, die Informationen und die Informationskanäle unseres Trägers zu archivieren und wieder bereitzustellen, und dies verfolgen wir mit unserer aktuell umgesetzten Strategie. Nichtsdestotrotz ist mittelfristig auch die Archivierung von Kommentaren vorgesehen, jedoch muss sich hierfür erst die rechtliche Grauzone klären.

Unsere Kriterien für die Auswahl der zu archivierenden Webseiten waren bzw. sind unterschiedlichster Natur. Neben der Informationsdichte und der Relevanz der veröffentlichten Behörde spielt auch die Gestaltung und eine naheliegende Offline-Schaltung bzw. Modernisierung von Webseiten eine entscheidende Rolle.

Auch das städtische Intranet war bzw. ist im Fokus des Stadt- und Stiftsarchivs. Hier findet im Frühjahr 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 Archivierungs-lösung zu erarbeiten. Dabei ist zu berücksichtigen, dass diese historische Quelle der Öffentlichkeit nicht online zugänglich gemacht wird.

Eine zentrale Frage bei der Umsetzung des Projektes „Archivierung städtischer Webseiten“ war, wie stellen wir die rein digitalen historischen Informationen unseren Nut-

zerinnen und Nutzern zur Verfügung? In unserem Lesesaal? Digitaler Zugriff von zu Hause? Wo bleibt die klassische Kontrolle der Archivarinnen und Archivare?

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

Das Ergebnis ist, dass wir – mit Hilfe einer extra geschaffenen Webpräsenz<sup>11</sup> – themenbasiert, unentgeltlich und barrierefrei unseren Nutzerinnen und Nutzern 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. ■



**Johannes Schuck**  
**Stadtarchiv Aschaffenburg**  
[Johannes.Schuck@aschaffenburg.de](mailto:Johannes.Schuck@aschaffenburg.de)

<sup>11</sup> Vgl.: <https://stadttarchiv-digital.de>