

Andreas Marquet, Annabel Walz

Social-Media-Archivierung aus institutioneller Perspektive: Überlieferungsbildung, technische Zugänge und rechtliche Rahmenbedingungen

Abstract: The article explores the challenges and solutions faced by memory institutions such as libraries and archives when it comes to social media content. It first briefly outlines the evolution of institutional social media archiving and then discusses two major hindrances for a more wide-spread, systematic archiving: the lacking legal framework and technical challenges. The legal areas that are affected in the field of social media archiving are mainly copyright, data protection and privacy protection. In both areas, a balance must be struck between private or business interests on the one hand and society's interest in preserving cultural heritage and making it accessible on the other. Although web content is recognised as potential cultural heritage the legal framework for its preservation is often still lacking. As for the technical challenges the article provides an overview of a range of possibilities for archiving social media, e. g. API access, web crawling, web recording and export functions. It discusses some of their advantages and disadvantages when it comes to preserving the characteristics of social media and gives a glance at the problems for digital preservation in the longterm. Finally, the article deals with how different institutions address the question of how to select the social media content for preservation. Both libraries and archives have reason to consider social media as part of their collection or documentation profile and both provenance-based and pertinence-based approaches are used as the basis for the development of selection criteria. Considering the challenges in preservation of social media this combined approach will hopefully help to broaden the material still available in the future, as will a strengthened collaboration with researchers in the field. Furthermore, the article stresses the importance and vivacity of international exchange and collaboration in the field of web and social media archiving.

1 Einleitung

Als Plattformen, die alle Bereiche des Lebens prägen, gehören Social Media zu den für Gedächtnisinstitutionen wie Bibliotheken und Archive bewahrenswerten Quellengattungen. Ihre Neuartigkeit gegenüber traditionell in diesen Institutionen

gesammelten Objekten und ihre ständige Veränderlichkeit stellen die Institutionen aber vor große Herausforderungen: Es muss neu durchdacht werden, was gesammelt werden soll und wie das technisch umzusetzen ist. Beides erfolgt innerhalb eines rechtlichen Rahmens, der erst dabei ist, den veränderten Erfordernissen angepasst zu werden – und natürlich bedarf es Ressourcen zur Erfüllung dieser neuen Aufgaben, obwohl die alten Aufgaben weiterhin bestehen bleiben.

Der vorliegende Aufsatz umreißt die Problemstellungen und Lösungsansätze, vor denen Institutionen bei der Social-Media-Archivierung stehen. Dazu wird nach einem kurzen Überblick über die institutionelle Landschaft zunächst der rechtliche Rahmen mit Schwerpunkt auf der deutschen Situation geschildert. Anschließend werden die möglichen technischen Herangehensweisen dargestellt, die zur Verfügung stehen, um die im letzten Abschnitt erläuterten konzeptionellen Sammlungsentscheidungen umzusetzen. Der Aufsatz wird dabei die Social-Media-Archivierung insgesamt behandeln, da Instagram und TikTok bislang in deutschsprachigen Institutionen, soweit öffentlich bekannt, nicht systematisch archiviert werden.¹

2 Landschaft

Der Übergang von der Web- zur Social-Media-Archivierung ist ähnlich fließend wie der vom Web 1.0 zum Web 2.0. Institutionen, die schon früh mit der Webarchivierung begonnen haben, haben zumindest in Teilen Social Media und deren Vorläufer einfach mit eingesammelt – beginnend mit GeoCities, dann Plattformen wie Blogger, Friendster, MySpace und Flickr.² Mit dem Erstarken von Social Media erweiterten einige große Gedächtnisinstitutionen zunehmend ihre Sammlungsprofile, sodass diese Inhalte bewusst adressiert wurden. 2008 haben die National Archives im Vereinigten Königreich mit einem speziell auf Social Media ausgerichteten Programm begonnen³, das erste Twitter-Archiv in der Library of New Zealand datiert auf 2009. Die Library of Congress (LoC) in den USA sammelte von 2010 bis 2017 in Absprache mit Twitter vollständig alle auf der Plattform abgesetzten

¹ Es ist denkbar, dass Institutionen Social-Media-Archivierung betreiben, dies aufgrund der unsicheren Rechtslage aber nicht öffentlich machen. Für einen Überblick zu internationalen Archivierungsinitiativen, die auch Instagram und TikTok berücksichtigen, siehe den Beitrag von Beatrice Cannelli in diesem Band.

² Vgl. Justin Littman et al., API-based Social Media Collecting as a Form of Web Archiving, in: *International Journal on Digital Libraries* 19/1 (2018), S. 21–38, hier S. 23, doi.org/10.1007/s00799-016-0201-7.

³ Vgl. ebd., S. 23.

Tweets.⁴ Die Bibliothèque nationale de France (BnF) berücksichtigt seit Beginn des Aufbaus ihres Webarchivs 2006 auch Facebook.⁵

Mittlerweile gehören zumindest in Westeuropa, Nordamerika und Ozeanien Social Media in vielen nationalen Gedächtniseinrichtungen zum Sammlungsbereich.⁶ Neben den Institutionen auf nationaler Ebene findet Social-Media-Archivierung aber auch in mittleren und kleinen Einrichtungen statt.⁷ Gerade bei Communities, die keine festen Verwaltungsstrukturen haben, dienen soziale Medien oft als hauptsächliche Kommunikationsplattform und sind deshalb entscheidend für die Archivierung.⁸ Andererseits fehlt es in kleineren Institutionen oft an den Ressourcen, die notwendig sind, um die Expertise und Infrastruktur für die Umsetzung der Archivierung zu erwerben und zu erhalten. Auch der Aufwand zur Klärung der Rechtslage stellt für solche Organisationen eine besondere Hürde dar.

In Deutschland hat die Deutsche Nationalbibliothek in Reaktion auf die Ankündigung des eingeschränkten Forschungszugangs für Twitter (mittlerweile X) in einem Projekt von Februar bis April 2023 einen Teil der deutschsprachigen Twitter-Inhalte archiviert.⁹ Schon zehn Jahre zuvor begann das Archiv für Christlich-Soziale Politik der Hanns-Seidel-Stiftung mit der Übernahme von Facebook-Accounts

4 Vgl. Eveline Vlassenroot et al., *Web-Archiving and Social Media: an Exploratory Analysis: Call for Papers Digital Humanities and Web Archives – A Special Issue of International Journal of Digital Humanities*, in: *International Journal of Digital Humanities* 2/1–3 (2021), S. 107–128, hier S. 109, doi.org/10.1007/s42803-021-00036-1. Danach stellte die LoC auf selektive Sammlung um. Vgl. Gayle Osterberg, *Update on the Twitter Archive at the Library of Congress*, in: *Timeless: Stories from the Library of Congress* (blog), 26.12.2017, blogs.loc.gov/loc/2017/12/update-on-the-twitter-archive-at-the-library-of-congress-2. Zu diesem Zeitpunkt war das technische Problem der Zugänglichkeit der riesigen Datenmenge noch nicht gelöst. Seitdem scheint es keine Neuigkeiten zu diesem Projekt zu geben.

5 Vgl. Vlassenroot et al., *Web-Archiving and Social Media*, S. 109.

6 Im Projekt BESOCIAL (kbr.be/en/projects/besocial/) wurde für solche Institutionen der aktuelle Stand der Social-Media-Archivierung 2020 zusammengetragen. Eine Zusammenfassung der Ergebnisse findet sich in Vlassenroot et al., *Web-Archiving and Social Media*.

7 Im Projekt „Best practices for archiving social media in Flanders and Brussels“ wurden Stand und Best Practices für private, kleinere und mittlere Archive in der genannten Region erhoben. Vgl. meemoo.be/en/projects/best-practices-for-archiving-social-media-in-flanders-and-brussels.

8 Vgl. z. B. für die Entstehung neuer religiöser Communities: Katrien Weyns, *Towards Best Practices for the Archiving of Social Media by Private Archival Institutions in Flanders, Wanted: Social Media Data – Archiving Practices and Research Use*, Brüssel 2022, in: @KBR, „Social Media Archiving from an Institutional Point of View“, YouTube-Video, 7.10.2022, ab Min. 23:40, www.youtube.com/watch?v=skVeVxQQVxo

9 Vgl. Claus-Michael Schlesinger/Britta Woldering, *Tweets archivieren. Initiative, Umsetzung, Ausblick, Show & Tell – Social Media-Daten in der Forschungspraxis*, Berlin 2023, nfdi4culture.de/fileadmin/user_upload/documents/Show_and_Tell/2023-05-26_NFDI_show_tell_Schlesinger_Woldering.pdf

unter Nutzung der zu diesem Zeitpunkt angebotenen Open-Graph-Schnittstelle.¹⁰ Als früher Akteur in der Archivierung von Social-Media-Daten ist außerdem das Leibniz-Institut für Sozialwissenschaften, GESIS, zu erwähnen, wo Inhalte aus sozialen Medien als Forschungsdaten archiviert werden.¹¹ Weitere Projekte und Initiativen zur Übernahme von Inhalten aus sozialen Medien sind die Übernahme des Twitter-Accounts von Thorsten Schäfer-Gümbel im Hauptstaatsarchiv in Wiesbaden im Jahr 2020¹², ein Twitter-Projekt des Sächsischen Staatsarchivs¹³ und die Übernahme aufbereiteter Social-Media-Inhalte von Facebook, Twitter, Instagram und YouTube der Pressestelle des Landtags von Baden-Württemberg durch das Landesarchiv Baden-Württemberg.¹⁴ Auf kommunaler Ebene übernimmt das Stadtarchiv Aschaffenburg die städtischen Accounts von mehreren Plattformen, wobei dafür ein Dienstleister genutzt wird.¹⁵ Das Stadtarchiv Münster hat gemeinsam mit dem Landschaftsverband Westfalen-Lippe (LWL) eine Software zur Übernahme von Twitter-Daten entwickelt und damit kommunale Accounts archiviert.¹⁶ Unter Verwendung dieser Software wurden auch im Archiv der sozialen Demokratie (AdsD) der Friedrich-Ebert-Stiftung zwischen September 2022 und Februar 2023 37 Twitter-Accounts¹⁷ archiviert, wobei die Inhalte soweit möglich ab Einrichtung der jeweiligen Accounts gesammelt wurden.

¹⁰ Gesammelt wurden der CSU-Organisationsaccount sowie Accounts der MdBs, MdLs und MdEPs, sofern deren Zustimmung vorlag. Vgl. Katharina Köhn, Parteienstiftungen und Social Media. Facebook-Archivierung mit OWA und Laurentius, Webarchive und Social Media, Baden-Baden 2013, files.dnb.de/nestor/veranstaltungen/Webarchivierung2013-10-18/3_Hanns-Seidel-Stiftung.pdf

¹¹ Vgl. Lars Kaczmarek et al., Social Media Monitoring of the Campaigns for the 2013 German Bundestag Elections on Facebook and Twitter (GESIS-Working Papers 31), 2014, doi.org/10.48550/arXiv.1312.4476.

¹² Vgl. Florian Stabel/Sebastian Tripp, Hauptstaatsarchiv Wiesbaden übernimmt Twitter-Daten von Thorsten Schäfer-Gümbel, Hessisches Landesarchiv, Wiesbaden 2022, web.archive.org/web/20220419214334/https://landesarchiv.hessen.de/twitter-daten_schaefer-guembel

¹³ Vgl. Stephanie Kortyla, Archivierung von Daten aus Twitter: Twitter-eigene Exportschnittstelle (2018–2021), Archivierung von Unterlagen aus digitalen Systemen, Vortrag AUdS 2021 Köln, 2.3.2021, www.sg.ch/content/dam/sgch/kultur/staatsarchiv/auds-2021/auds-camp-i/Vortrag_Kortyla_Twitter_Archivierung.pdf

¹⁴ Vgl. Paulina Eckert, Tagungsbericht: Zwischen Link-Graphen und Paragraphen – Archivierung und Nutzung von Netz-Quellen, in: H-Soz-Kult, 3.2.2023, www.hsozkult.de/conferencereport/id/fdkn-133549

¹⁵ Vgl. Johannes Schuck, Schritt für Schritt auf neuen digitalen Wegen. Webseiten- und Social-Media-Kanal Archivierung im Stadt- und Stiftsarchiv Aschaffenburg, in: Archivpflege in Westfalen-Lippe 96 (2022), S. 17–20, nbn-resolving.de/urn:nbn:de:hbz:6:2-1646179

¹⁶ Vgl. Peter Worm, Neue Ansätze für die Archivierung eines Twitter-Accounts, in: Archivpflege in Westfalen-Lippe 96 (2022), S. 26–39, nbn-resolving.de/urn:nbn:de:hbz:6:2-1646179

¹⁷ Gemäß dem Sammlungsprofil des AdsD handelt es sich dabei um die Accounts des Parteivorstands und der Landesverbände der SPD, ausgewählter Parteifunktionär*innen und Spitzenkandidat*innen sowie die Accounts der Gewerkschaften, deren Bestände im AdsD aufbewahrt werden.

3 Rechtliche Rahmenbedingungen

Digitale Objekte sind als erhaltenswertes Kulturerbe längst international anerkannt, wie die entsprechende Deklaration der UNESCO von 2003 zeigt. Zu den verschiedenen hierunter zu subsumierenden Manifestationen und Erscheinungsformen werden auch Webseiten gezählt, die „in einer wachsenden Vielfalt von Formaten“¹⁸ auftreten. Ferner haben die Mitgliedsstaaten der UNESCO sich darauf verständigt, für die notwendige Sicherung und die angestrebte Zugänglichkeit digitaler Inhalte rechtliche Rahmenbedingungen zu schaffen, die ein „faires Gleichgewicht zwischen den legitimen Rechten der Urheber und anderer Rechteinhaber einerseits und dem öffentlichen Interesse am Zugang zu Materialien des digitalen Erbes andererseits (...) entsprechend internationaler Normen und Vereinbarungen“¹⁹ garantieren. Während diese Postulate den Charakter allgemeiner Erklärungen zum digitalen Kulturerbe besitzen, verweist der notwendige Maßnahmen benennende Abschnitt auf hiermit unmittelbar verbundenen Handlungsbedarf, der sich auch auf die rechtlichen Rahmenbedingungen bezieht. So werden landesweite gesetzliche Regelungen zur Ablieferung und Sicherung digitaler Inhalte wie auch der Zugang dazu „innerhalb zumutbarer Restriktionen“²⁰ eingefordert, ohne freilich durch die Bereitstellung die „übliche Verwertung“ zu beeinträchtigen. Bezogen auf die Zugänglichmachung von Social-Media-Inhalten soll demnach dem gewinnorientierten Geschäftsmodell der Plattformbetreiber*innen Rechnung getragen werden.

Diese Charta vom Beginn der 2000er Jahre zeigt das Spannungsfeld zwischen verschiedenen, zum Teil divergierenden Interessen, Rechtsrahmen und -gebieten sowie beteiligten Akteur*innen, die bei der Erhaltung und Bereitstellung von digitalen Kommunikationsplattformen nach wie vor die Arbeit der Gedächtniseinrichtungen im rechtlichen Sinne determinieren. Neben den Inhaber*innen der Urheberrechte gilt dies für weitere Rechteinhaber*innen, worunter auch Dritte zu verstehen sind, deren schutzwürdige Belange im Bereich des Persönlichkeitsrechts und des Datenschutzes zu fassen sein dürften. Demgegenüber wird ausdrücklich auf das öffentliche Interesse am Zugang zum digitalen kulturellen Erbe abgehoben, an dem sich Restriktionen bei Überlieferung und Zugang messen lassen müssen. Eine Einschätzung der rechtlichen Rahmenbedingungen für die Archivierung speziell der Plattformen Facebook, Twitter und Instagram wurde 2022 mit Bezug auf die Stellung des AdsD von diesem beauftragt und von der auf

18 UNESCO, Charta zur Bewahrung des digitalen Kulturerbes, Paris 2003, www.unesco.de/sites/default/files/2018-03/2003_Charta_zur_Bewahrung_des_digitalen_Kulturerbes.pdf

19 Ebd., S. 2.

20 Ebd., S. 3.

Kulturerbe-Einrichtungen spezialisierten Kanzlei i.Rights.Law vorgenommen.²¹ Wird urheberrechtlich geschützter Content gesichert, erfolgt bereits beim Download technisch gesehen ein Kopierprozess, der in Deutschland nicht durch die relevanten Erlaubnisregelungen der Paragraphen 53 und 61 des Urheberrechts abgedeckt ist.²² Aufgrund der Vielzahl von Beiträgen in sozialen Medien erscheint eine Einzelfallprüfung jedes Postings aber praktisch nicht umsetzbar, wenngleich, wie unten ausgeführt, die Art des Beitrags (Text, Bild, Bewegtbild, Audio-Beitrag) mit Blick auf das Urheberrecht differenziert werden kann. Eine dezidierte rechtliche Erlaubnis für die Archivierung von Social-Media-Kanälen steht noch aus, mit Ausnahme des Gesetzes über die Deutsche Nationalbibliothek.²³ Jedoch kann mit der im öffentlichen Interesse geleisteten Arbeit von Gedächtnisinstitutionen argumentiert werden. Dies gilt unabhängig von der – privaten oder öffentlichen – Träger*innenschaft, setzt jedoch die freie Zugänglichkeit und ein gesellschaftliches Interesse an den Sammlungsaktivitäten der Einrichtung voraus. Zudem können Verträge mit den Betreiber*innen einzelner Social-Media-Kanäle geschlossen werden, um die Archivierung der Inhalte zu regeln. Allerdings umfasst dies nicht die Interaktion mit anderen als den Kontoinhaber*innen – und damit ein konstitutives qualitatives Merkmal sozialer Medien –, die Antworten, Retweets, Kommentare etc. Da der Kreis der Nutzer*innen potenziell unbeschränkt ist, erschiene eine vertragliche Regelung in keiner Weise praktikabel.²⁴

Das öffentliche Interesse begründet jedoch keineswegs die frei im Internet zugängliche Bereitstellung digitaler Objekte durch Gedächtnisinstitutionen. Vielmehr ist der Terminal-Zugang, also der Zugang an einer speziell hierfür und ausschließlich in den Räumlichkeiten der Gedächtniseinrichtung bereitgestellten technischen Umgebung, nicht nur die rechtlich gangbarste, sondern auch die praktisch wohl überwiegend umgesetzte Lösung. Zwar wäre denkbar, Posts aus einem Social-Media-Kanal online bereitzustellen, wenn keine urheberrechtlich relevante Schöpfungshöhe vorliegt. Dies wird auf einzelne Tweets auf Twitter/X zutreffen,

21 Vgl. Annabel Walz/Andreas Marquet (Hrsg.), *Sicher Sichern? Social Media-Archivierung aus rechtlicher Perspektive im Archiv der sozialen Demokratie*, Beiträge aus dem Archiv der sozialen Demokratie Bd. 17, Bonn 2022, library.fes.de/pdf-files/adsd/19590.pdf

22 Vgl. www.gesetze-im-internet.de/urhrg/

23 Hier wurde der Sammlungsauftrag auf „Medien in unkörperlicher Form“ übertragen, wovon „alle Darstellungen in öffentlichen Netzen“ gefasst sind. Gesetz über die Deutsche Nationalbibliothek (DNBG) Nichtamtliches Inhaltsverzeichnis, 2006, www.gesetze-im-internet.de/dnbg/BjNR133800006.html

24 Vgl. Paul Klimpel/Fabian Rack, *Einschätzung der rechtlichen Rahmenbedingungen für die Archivierung von Social-Media-Inhalten im Archiv der sozialen Demokratie*, in: Annabel Walz/Andreas Marquet (Hrsg.), *Sicher Sichern? Beiträge aus dem Archiv der Sozialen Demokratie* Bd. 17, Bonn 2022, S. 15–48, hier S. 20–24, library.fes.de/pdf-files/adsd/19590.pdf

doch schon bei zusammenhängenden Threads kritisch. Bild- und videolastige Medien wie TikTok und Instagram beinhalten, selbst wenn nur Schnappschüsse gepostet werden, immer urheberrechtlich geschützte Werke. Da selbst bei textlastigen Medien jeweils im Einzelfall der Urhaberschutz zu prüfen wäre, scheidet die Online-Bereitstellung in der Praxis zumeist aus.²⁵

Hier hebt sich der US-amerikanische Rechtsrahmen insofern merklich von den deutlich restriktiver gefassten europäischen Gesetzgebungen ab, als das Fair-Use-Prinzip weitreichende Ausnahmen für Archive und Bibliotheken gestattet.²⁶ Schließlich entwickelt sich ein weiteres Spannungsfeld entlang der Anforderungen nach einem freien Zugang zu Forschungsergebnissen im Sinne des Open Access. Während der Zugang zu Forschungsergebnissen durch frei verfügbare Publikationen zu deren schnellerer und besserer Verbreitung beitragen soll, unterliegt die Bereitstellung der zugrunde liegenden Quellen merklich restriktiveren Regelungen.²⁷ Die FAIR-Prinzipien sehen in Analogie hierfür einen nachhaltigen und freien Zugang zu Forschungsdaten vor. Das Akronym verweist auf die vier zentralen Aspekte: Findable (Auffindbar), Accessible (Zugänglich), Interoperable (Interoperabel) und Reusable (Wiederverwendbar), die durch Beachtung hiervon abgeleiteter Anforderungen umgesetzt werden können. Wenngleich die Zugänglichkeit nach diesen Prinzipien prinzipiell auch über Authentifizierungsverfahren für eingeschränkte Zugriffsmöglichkeiten gelöst werden kann, besteht kein Zweifel am grundsätzlichen Paradigma der Offenheit der FAIR-Prinzipien.²⁸

In der Offenheit und Interaktionsfähigkeit sozialer Medien ist schließlich die Bedeutung der datenschutz- und persönlichkeitsrechtlichen Bestimmungen begründet. Der im öffentlichen Interesse liegende Archivzweck ist jedoch ausdrücklich in der Datenschutz-Grundverordnung (DSGVO) anerkannt und erlaubt eine Ausnahme von der Zweckbindung der Datenverarbeitung. Gedächtniseinrichtungen leisten schließlich einen über ihren primären Sammlungsauftrag hinausgehenden Beitrag zur Wissenschafts- und Informationsfreiheit. Gleichwohl sind Rechte Betroffener abzuwägen sowie Auskunft- und Informationspflichten zu

25 Vgl. ebd., S. 17, 18. Für Archive von Webseiten, die ebenfalls neben dem selbst erstellten Content oftmals urheberrechtlich bewährte Materialien Dritter wie bspw. Agenturfotos beinhalten, gilt dies ebenfalls, wenngleich hier zumindest Ausnahmen konstatiert werden können. Vgl. Lise Jaillant, *How Can We Make Born-Digital and Digitised Archives More Accessible? Identifying Obstacles and Solutions*, in: *Archival Science International Journal on Recorded Information* 22 (2022), S. 417–436, hier S. 431, doi.org/10.1007/s10502-022-09390-7.

26 Vgl. Paul Klimpel, *Kulturelles Erbe digital. Eine kleine Rechtsfibel*, Berlin 2020, S. 37, digis-berlin.de/wp-content/uploads/2020/09/digiS_PKlimpel_Rechtsfibel.pdf; Jaillant, *Accessible Archives*, S. 431.

27 Vgl. Jaillant, *Accessible Archives*, S. 432.

28 Vgl. www.go-fair.org/

wahren.²⁹ Auch weitere organisatorische und technische Maßnahmen wie Pseudonymisierung, Anonymisierung oder die Vermeidung von Profilbildungen können für Archive und Bibliotheken angemessene Maßnahmen darstellen, um den datenschutzrechtlichen Anforderungen Genüge zu leisten.

Schließlich stellen die Betreiber*innen von Social-Media-Plattformen einen eigens zu beachtenden Aspekt bei der rechtlichen Bewertung dar. Bei diesen handelt es sich in der Regel um private Unternehmen, mit denen User*innen einen Vertrag eingehen, wenn sie die allgemeinen Geschäftsbedingungen bei Erstellung eines Nutzungsprofils annehmen. Die Geschäftsbedingungen der Plattformbetreiber*innen beziehen sich ganz überwiegend auf ihre Geschäftsmodelle. Zugänge zur Archivierung und Forschung sind selten geregelt.³⁰

Auch die Archive und Bibliotheken selbst unterliegen gemäß ihrer Verfasstheit und Trägerschaft unterschiedlichen rechtlichen Anforderungen. Während die Pflichtexemplargesetze für die Deutsche Nationalbibliothek sowie die Landesbibliotheken einen Rahmen und zugleich einen Handlungsauftrag konstituieren, sind die Archivgesetze des Bundes und der Länder dahingehend weniger eindeutig formuliert. Dennoch kann ein Mandat zur Archivierung von sozialen Medien abgeleitet werden, wenn Verwaltungshandeln, worunter auch die Kommunikation öffentlicher Einrichtungen in digitalen Medien zu fassen ist, abgebildet und zugleich die Forschung unterstützt werden soll.³¹ Demgegenüber steht privatrechtlich organisierten Einrichtungen dieser gesetzlich verbrieft Auftrag als Handlungslegitimation nicht zur Verfügung. Dennoch steht es ihnen natürlich frei, über vertragliche Regelungen entsprechende Aufträge zu kodifizieren.

Forschenden sind Download und Auswertung zu Zwecken des Text- und Data-Mining ausdrücklich auch dann gestattet, wenn dies den Geschäftsbedingungen der Plattformbetreiber*innen zuwiderläuft. Zum Zweck des Text- und Data-Mining angefertigte Kopien müssen nach Abschluss der wissenschaftlichen Auswertung gelöscht und dürfen nicht dauerhaft gesichert werden.³² Aus dieser beschränkten urheberrechtlichen Erlaubnis lässt sich freilich der Bedarf nach vertrauenswürdigen Archiven für Web- und Social-Media-Inhalte insofern ablei-

²⁹ Vgl. Klimpel/Rack, Rechtliche Rahmenbedingungen, S. 29–31.

³⁰ Vgl. ebd., S. 36, 37.

³¹ Eine umfassende Gesamtschau der bundes- und landesrechtlichen gesetzlichen Regelungen steht noch aus. Beispielhaft sei hier auf Baden-Württemberg verwiesen, wo Bibliotheksservice-Zentrum, Landesbibliotheken und Landesarchiv einen kooperativen Ansatz begründet haben. Hier erfolgt auch eine Ableitung des Handlungsauftrags auf Grundlage des Landesarchivgesetzes. Vgl. Kai Naumann, Gemeinsam stark. Web-Archivierung in Baden-Württemberg, Deutschland und der Welt, in: *Der Archivar* 65/1 (2012), S. 33–41, hier S. 39.

³² Vgl. Klimpel/Rack, Rechtliche Rahmenbedingungen, S. 22, 23.

ten, als die Nachvollzieh- und Überprüfbarkeit wissenschaftlicher Erkenntnis den freien Zugang zu gesicherten Quellen voraussetzt.

4 Technische Umsetzung der Archivierung

Soziale Medien beruhen auf Webtechnologien und können deshalb mit ähnlichen Techniken archiviert werden.³³ Damit steht man bei der Archivierung dieser Quellen vor ähnlichen Schwierigkeiten wie bei „traditionelleren“ Formen von Webinhalten.³⁴ Der erste Schritt ist dabei, überhaupt Daten in ein Archiv zu übernehmen. Bei der Übernahme sollte der sich anschließende Schritt aber immer schon mitgedacht werden, nämlich wie die Daten auf Dauer so bewahrt werden können, dass sie auch für zukünftige Generationen noch nutzbar sind. Um das zu gewährleisten, benötigt man zukünftig eine Umgebung, in der die originalen Formate abgespielt werden können. Dazu muss man eine entsprechende Hard- und Softwareumgebung emulieren. Eine alternative Herangehensweise ist, die ursprünglichen Dateien in Formate zu migrieren, die auf längere Zeit stabil zu bleiben versprechen. Bei beiden Strategien muss gewährleistet bleiben, dass die Daten integer und authentisch bleiben, also keine Veränderungen am digitalen Objekt stattfinden, die nicht für den Archivierungsprozess notwendig und dokumentiert sind, und dass nachvollziehbar bleibt, dass es sich um das echte ursprüngliche Objekt handelt.³⁵

Bei Webinhalten allgemein und Social Media im Besonderen kann immer nur ein transformierter Ausschnitt eines flüchtigen „Originals“ gesichert werden. Das liegt nicht nur daran, dass sich der Inhalt selbst jederzeit verändern kann. Auch die Präsentation der Inhalte ist davon abhängig, welche Software-Version und welches

³³ Zur Frage, inwiefern bei Webinhalten der Begriff „Archiv“ angemessen ist, vgl. z. B. Andrea Ammendola, Webarchivierung in NRW aus Sicht der Universitäts- und Landesbibliothek Münster, Masterarbeit Technische Hochschule Köln, 2020, S. 4–7, nbn-resolving.org/urn:nbn:de:hbz:79p-bc-opus-16219

³⁴ Vgl. dazu grundlegend: Niels Brügger, *The Archived Web: Doing History in the Digital Age*, Cambridge/ Massachusetts 2018; Niels Brügger, *Understanding the Archived Web as a Historical Source*, in: Niels Brügger/Ian Milligan (Hrsg.), *The SAGE Handbook of Web History*, London 2019, S. 16–29.

³⁵ Mit diesen Fragen, die digitale Objekte generell betreffen, setzt sich die digitale Langzeitarchivierung auseinander. Einen guten ersten Einblick für Forscher*innen in Fragen der digitalen Langzeitarchivierung (mit Fokus auf Forschungsdaten) bietet Jörg Heseler et al., *Grundlagen der digitalen Langzeitarchivierung*. Eine Handreichung zur digitalen Langzeitarchivierung aus Perspektive der NFDI4Culture Community (1.0.2), 17.1.2024, doi.org/10.5281/zenodo.10511404. Einen allgemeineren ersten Einstieg bietet Digital Preservation Coalition, *Digital Preservation Handbook*, 2. Aufl., Glasgow 2015, www.dpconline.org/handbook. Eine neue Auflage ist aktuell in Vorbereitung und wird über die genannte URL verfügbar sein.

Gerät verwendet werden, um die Inhalte anzuzeigen – das Konzept des zu bewahrenden Originals stößt hier deshalb an seine Grenzen. Folglich muss bereits bei der Übernahme bedacht werden, welche Eigenschaften des „Originals“ möglichst erhalten werden sollen. Zudem ist die Nachvollziehbarkeit der Entstehung der Archivversion entscheidend für die Möglichkeit zur quellenkritischen Bewertung.³⁶

Social Media bringen aber auch Eigenheiten mit, die einerseits zusätzliche Archivierungsmöglichkeiten bieten, andererseits weitere technische Probleme nach sich ziehen. Durch die (meist) gegebene Plattformgebundenheit ihrer Inhalte bieten Social Media häufiger als andere Webseiten Programmierschnittstellen (Application Programming Interfaces, Abk. APIs), über die auf die der Anwendung zugrunde liegenden Daten in strukturierter Form zugegriffen werden kann, was sich für die Archivierung nutzen lässt.³⁷ Als Antwort auf die Abfrage über eine APIs werden in textueller Form z. B. für einen Post der *username* des*der Verfassenden, die zugehörige Identifikationsnummer, die URL zum geposteten Bild oder Video, der Begleittext, der Zeitpunkt, wann der Post abgesetzt wurde, die Anzahl und Art der Reaktionen auf den Post sowie die Kommentare samt der ID der Kommentator*innen zurückgegeben.³⁸ Welche Inhalte genau ausgegeben werden, hängt von der Abfrage ab und davon, welche Datenfelder von der API angeboten werden.

Damit können Plattformbetreiber*innen steuern, welche Daten in welcher Form über APIs zur Verfügung gestellt werden und wer darauf Zugriff bekommt.³⁹ Die Modalitäten für diesen Zugriff kann die Plattform jederzeit ändern. Beispiele dafür sind die Einschränkung der API-Nutzung durch Facebook 2018 infolge des

36 Vgl. dazu z. B. Emily Maemura et al., If These Crawls Could Talk: Studying and Documenting Web Archives Provenance, in: *Journal of the Association for Information Science and Technology* 69/10 (2018), S. 1223–1233, doi.org/10.1002/asi.24048; Gregory Wiedemann, Describing Web Archives: A Computer-Assisted Approach, in: *Journal of Contemporary Archival Studies* 6, Article 31 (2019), elischolar.library.yale.edu/jcas/vol6/iss1/31/

37 Vgl. Sara Day Thomson/Neil Beagrie, Preserving Social Media, DPC Technology Watch Reports, 2016, S. 7 und 9–11, dx.doi.org/10.7207/twr16-01.

38 Es handelt sich hier nur um ein generisches Beispiel. Die genaue Struktur der API und damit die Abfragemöglichkeiten können von den Plattformbetreiber*innen jederzeit geändert werden. Beispiele für solche Abfragen mittels der Facebook Graph API, die der Instagram Graph API zugrunde liegt, finden sich z. B. hier: developers.facebook.com/docs/graph-api/overview. Beispiele für Abfragen der TikTok Research API finden sich hier: developers.tiktok.com/doc/research-api-specs-query-user-pinned-videos

39 Vgl. Helen Hockx-Yu, Archiving Social Media in the Context of Non-print Legal Deposit, IFLA WLIC, Lyon 2014, S. 3, library.ifla.org/id/eprint/999. Vgl. auch beispielhaft die Notwendigkeit, sich bei TikTok für einen Forschungszugang zu bewerben: developers.tiktok.com/products/research-api/. Die Nutzung der API für eine Archivierung von Accounts durch Gedächtnisinstitutionen ist dabei als Anwendungsfall nicht vorgesehen.

Cambridge-Analytica-Skandals⁴⁰ oder die Einführung von Gebühren auch für Forscher*innen durch Twitter im Frühjahr 2023.⁴¹

Die API-Daten sind außerdem nicht geeignet, die Oberfläche von Social-Media-Anwendungen zu dokumentieren. Für eine schwerpunktmäßige Archivierung der Oberfläche sind Screenshots oder Screencasts eine mögliche Ergänzung.⁴² Auch beim Web-Crawling wird über die Oberfläche, wie sie sich im Browser darstellt, auf die Social-Media-Inhalte zugegriffen. Das Web-Crawling stammt als Technik aus dem Bereich der „traditionellen“ Webarchivierung: Ausgehend von sog. *seeds* verfolgen Crawler dabei nach vorgegebenen Regeln die auf den Seiten vorgefundenen Links und sammeln die dort hinterlegten Inhalte ein. Um mit den dynamisch gebauten Social-Media-Webseiten umgehen zu können, müssen solche Crawler aufwendig konfiguriert werden.⁴³ Teilweise versagen sie auch ganz, wobei neuere Generationen von Web-Crawlern⁴⁴ besser mit dynamischen Inhalten umgehen können.

Mitunter kann es bei der Archivierung helfen, Web-Recorder zu verwenden, die nicht automatisiert Links folgen, sondern bei denen ein Mensch mit einer Seite interagiert und alle Inhalte, die so aufgerufen werden, vom Crawler heruntergeladen werden.⁴⁵ Eine weitere Herangehensweise ist Web-Scraping, wobei nur ausgewählte Inhalte einer Webseite in transformierter Form als strukturierte Daten gespeichert werden.⁴⁶ Generalisierte Aussagen dazu, welche Crawl- und Scra-

40 Vgl. Anat Ben-David, Gegenarchivierung von Facebook, in: Christian Schwarzenegger et al. (Hrsg.), Digitale Kommunikation und Kommunikationsgeschichte. Perspektiven, Potentiale, Problemfelder, Berlin 2022, S. 133–155, hier S. 133, 134, doi.org/10.48541/dcrv10.5.

41 Vgl. Eva-Maria Weiß, Twitter-API: 1500 Tweets kostenlos, 100 US-Dollar für das günstigste Abo, in: heise online, 30.3.2023, heise.de/news/Twitter-bestaetigt-1500-Tweets-kostenlos-100-US-Dollar-fuer-das-guenstigste-Abo-8239039.html

42 Vgl. Thomson/Beagrie, Preserving Social Media, S. 26. Mit Wissen um das ursprüngliche Aussehen der Seite lässt sich damit aus den API-Daten die ursprüngliche Erscheinung ggf. rekonstruieren.

43 Vgl. dazu z.B. die Hinweise zur Crawler-Konfiguration für TikTok: support.archive-it.org/hc/en-us/articles/4410475150868-Archiving-TikTok und die Ausführungen in: Arndt Klindt Myrvoll, Social Media Archiving at the Royal Danish Library – Reflections and Status Quo, Wanted: Social Media Data – Archiving Practices and Research Use, Brüssel 2022, in: @KBR-BEL, „Social Media Archiving from an Institutional Point of View“, YouTube-Video, 7.10.2022, ab Min. 4:56, www.youtube.com/watch?v=skVeVxQQVxo

44 So z.B. der browserbasierte Crawler Browsertrix; vgl. github.com/webrecorder/browsertrix-crawler

45 Ein Beispiel dafür ist die Chrome-Extension archive.page; vgl. webrecorder.net/tools#archive-webpage

46 Ein Beispiel für einen auf die Twitter-Archivierung spezialisierten Scraper ist twint; vgl. github.com/twintproject/twint. Dieser Scraper stellte auch die Basis für das oben erwähnte von Stadtar-

ping-Software bei welcher Plattform eingesetzt werden kann, lassen sich kaum sinnvoll machen, weil sich die Oberflächen häufig ändern und deshalb auch die Crawler bzw. ihre Konfiguration angepasst werden müssen.⁴⁷ Diese Änderungen wie auch die zuweilen aufwendige Installation und Pflege der Archivierungssoftware führen dazu, dass die praktische Umsetzung gerade in kleineren und mittleren Archiven eine große Herausforderung darstellt. Eine Reaktion darauf ist der Rückgriff auf Dienstleister*innen.⁴⁸

Eine weitere Herangehensweise von Institutionen, um Social-Media-Inhalte zu sichern, ist der Aufruf zu Datenspenden. Dabei werden Nutzer*innen einer breiten Öffentlichkeit oder Hinterleger*innen, deren Bestände sich ohnehin bereits in einem Archiv befinden, gebeten, Daten aus ihren Social-Media-Accounts über die Exportfunktion herunterzuladen und an die Institution zur Bewahrung für die Zukunft zu übergeben.⁴⁹ Dieser Zugang bietet die Möglichkeit, auch Daten aus dem privaten Bereich zu archivieren, dafür können aber in der Regel nur die Posts und Inhalte gesichert werden, die von den Account-Inhaber*innen selbst stammen.

Diese verschiedenen Formen der Übernahme ziehen unterschiedliche Formate nach sich, in denen die Daten gespeichert werden. Bei Übernahmen über eine API und Web-Scraping liegen die Daten in der Regel als strukturierte Textdateien (z. B. json, csv) vor. Screenshots und -casts werden als Bilder oder Videos gesichert. Beim Export zur Selbstarchivierung liegen die Daten in einer Mischung aus Text- (z. B. json oder html) und Mediendateien vor. Meist handelt es sich in diesen Fällen also um verbreitete Text-, Bild- und Videodateien, die als Formate für die Nachnutzung und dauerhafte Archivierung keine außergewöhnlichen Anforderungen nach sich ziehen.

chiv Münster und LWL entwickelte Skript zur Twitter-Archivierung dar. Vgl. zur Software: Michael Jerusalem et al., Ein neuer Ansatz für die Langzeitarchivierung von Twitter-Accounts, in: Hypotheses (blog), 9.3.2021, archive20.hypotheses.org/10031.

47 Für einen Überblick zu verschiedenen Tools vgl. z. B. die Zusammenstellung des BESOCIAL-Projekts: kbr.be/wp-content/uploads/2020/07/202009_BeSocial_social_media_archiving_tools_CC_BY_4.0.xlsx und eine Zusammenstellung von Anleitungen für verschiedene Social-Media-Archivierungs-Tools (leider nur auf Niederländisch): www.projectceest.be/wiki/Publicatie:Workflows_voor_de_captatie_van_individuele_sociale_media_accounts. Ein Überblick über Webarchivierungs-Tools, die teilweise auch für Social Media eingesetzt werden können, findet sich hier: coptr.digipres.org/index.php/Web.

48 In den USA verwendet die große Mehrheit der Institutionen den Anbieter Archive-it. Vgl. National Digital Stewardship Alliance, 2022 Web Archiving Survey Results (2023), S. 24, doi.org/10.17605/OSF.IO/N5MYR. In Deutschland arbeitet z. B. das Stadtarchiv Aschaffenburg mit dem Anbieter MirrorWeb zusammen. Vgl. Schuck, Schritt für Schritt auf neuen digitalen Wegen.

49 Vgl. dazu z. B. das Facebook-Archivierungsprojekt der National Library in Neuseeland: natlib.govt.nz/collections/donations/donating-digital-items/facebook-archive-project

Beim Web-Crawling und -Recording dagegen werden die Dateien meist im sog. WARC-Format gespeichert. Es handelt sich hier um ein Containerformat, in das alle gefundenen Inhalte in der Reihenfolge, wie sie beim Crawl gesammelt werden, zusammen mit den beim Aufruf zwischen Client und Server ausgetauschten Informationen geschrieben werden.⁵⁰ Um diese Daten nutzen zu können, benötigt man Software, die sie für Menschen handhabbar machen. Das kann z. B. durch eine Replay-Software sein, mittels der die Inhalte so „original“ wie möglich wiedergegeben werden sollen.⁵¹ Da es sich beim WARC-Format aber um strukturierte Daten handelt, können diese auch mit computationellen Methoden analysiert werden.⁵² Das WARC-Format ist dadurch einerseits sehr mächtig, andererseits ist die Frage der Langzeitarchivierung dieses Formats noch nicht gelöst.⁵³

5 Sammlungsprofil

Der Einfluss von Social Media auf gesellschaftliche Entwicklungen wie auch auf die politische Kultur und Mobilisierung ist längst belegt. Nicht wegen der fehlenden Einsicht in die Relevanz und Überlieferungswürdigkeit, sondern vielmehr wegen technischer und rechtlicher Probleme wie auch mangelnder Ressourcen betreiben staatliche und kommunale Gedächtniseinrichtungen bislang nicht flächendeckend Social-Media-Archivierung.⁵⁴ Auch für die Archive von Unternehmen, Verbänden oder Nicht-Regierungsorganisationen, Kirchen und Religionsgemeinschaften mit ihren Sammlungsaufträgen ist diese Form der Außendarstellung und Kommunikation – und deren Archivierung – bedeutsam.

⁵⁰ Vgl. z. B. Astrid Schoger/Konstanze Weimer, Das Dateiformat WARC für die Webarchivierung, nbn-resolving.de/urn:nbn:de:0008-2021042614. Dieses Format zieht wiederum Probleme in der digitalen Bestandserhaltung nach sich. Vgl. Fußnote 53.

⁵¹ Ein Beispiel dafür ist die Darstellung der in WARC-Format gespeicherten Inhalte des Internet Archive in der Wayback-Machine, z. B. web.archive.org/web/20230620000456/http://www.tiktok.com/explore

⁵² Hinweise zu Software, die dafür genutzt werden kann, findet sich z. B. in folgender Liste, insbes. in den Rubriken Search & Discovery und Analysis: github.com/iipc/awesome-web-archiving.

⁵³ Vgl. z. B. Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen, Webarchivierung. Eine Studie der KOST, 31.5.2016, kost-ceco.ch/cms/index.php?web_archiving_study_de

⁵⁴ Vgl. Christoph Neuberger, Digitale Öffentlichkeit und liberale Demokratie, in: *Aus Politik und Zeitgeschichte* 72/10–11 (2022), S. 18–25. Außerdem aktuell für das Beispiel Irland die Ergebnisse von Sharon Healy/Helena Byrne, *Scholarly Use of Web Archives Across Ireland: The Past, Present & Future(s)*, WARCnet Special Report, Aarhus 2023, S. 92, cc.au.dk/fileadmin/dac/Projekter/WARCnet/Healy_Byrne_Scholarly_Use_01.pdf

Der Umgang mit Webobjekten wie den Inhalten von sozialen Medien weist neben technischen und rechtlichen Herausforderungen auch den Bedarf nach konzeptioneller Einordnung auf. Auch wenn bei definierten Gruppen und Follower*innenkreisen eine zumindest von den Absender*innen unterstellte Einschränkung der Leser*innen nicht auszuschließen ist, dürfte die Unkontrollierbarkeit eines einmal abgesetzten Posts grundsätzlich den meisten User*innen bewusst sein. Demnach haben soziale Medien den Charakter von Publikationen, weshalb es naheliegt, dass primär Bibliotheken Aktivitäten zu deren Sicherung entfaltet haben.⁵⁵ Zugleich konstituieren sich Webobjekte aber auch dynamisch, abhängig von verschiedenen Variablen wie dem Zeitpunkt des Aufrufs, der geografischen Zugriffsregion, der technischen Umgebung des Clients und natürlich den algorithmischen Vorgaben der einzelnen Dienste und Anwendungen. Insofern weisen Webobjekte einen unikalen Charakter auf, der gemeinhin für Archivgut reklamiert wird.⁵⁶ Daher kann es zu Überschneidungen der Sammlungsaufträge von Archiven und Bibliotheken kommen, die im besten Fall – angesichts der dynamischen Entwicklung des Web – zu einer Ergänzung der jeweiligen Sammlungen führen. Gerade angesichts rasch aktualisierter Inhalte reduzieren sich mit einer derart kollektiv wie zugleich dezentral und unkoordiniert entstehenden Sammlung die Überlieferungslücken. Wenn also die Ressourcen einzelner Gedächtniseinrichtungen nicht ausreichen, um eine dichte Überlieferung von Webinhalten sicherzustellen, kann dies unter Umständen bis zu einem gewissen Grad durch die Sammlungsaktivitäten anderer Einrichtungen aufgefangen werden.

In Dokumentations- und Sammlungsprofilen spezifizieren Gedächtniseinrichtungen ihren Sammlungs- resp. Archivierungsauftrag. Derartige Dokumente haben grundsätzliche Geltung für einen unbestimmten Zeitraum, was freilich Anpassungen nicht ausschließt. Nichtsdestoweniger weisen sie ein Abstraktionsniveau auf, das durch nachgelagerte Richtlinien zum Bestandsaufbau konkretisiert wird. Für Archive sind hierbei insbesondere Bewertungsmodelle wichtige Ergänzungen, da sie archivwürdiges von archivunwürdigem Schrift- und Sammlungsgut entlang transparenter Kriterien nachvollziehbar unterscheiden. Solche Modelle beziehen sich zumeist auf einzelne Bestände oder Sammlungen. Die Sicherung von Social-Media-Kanälen ist also im Kontext dieser aufeinander aufbauenden Ziel- und Umsetzungsabsichten konzeptionell gefasst.

⁵⁵ Vgl. Kieran Hegarty, *The Invention of the Archived Web: Tracing the Influence of Library Frameworks on Web Archiving Infrastructure*, in: *Internet Histories* 6/4 (2022), S. 432–451, hier S. 436, doi.org/10.1080/24701475.2022.2103988.

⁵⁶ Vgl. Manfred Reimann/Marcus Stumpf, *Grundfragen und Organisation des Archivwesens*, in: Marcus Stumpf (Hrsg.), *Praktische Archivkunde. Ein Leitfaden für Fachangestellte für Medien- und Informationsdienste – Fachrichtung Archiv*, 4. Aufl., Münster 2018, S. 25–56.

Aufgrund der Schnellebigkeit von sozialen Medien und auch des zum Teil ephemeren Charakters einzelner Inhalte wird die Schwierigkeit deutlich, die Granularität der Dokumentation angemessen zu fassen. Bewusst weit gefasste Formulierungen, wie im Gesetz über die Deutsche Nationalbibliothek benannte *unkörperliche Werke*⁵⁷, müssen praxistauglich übersetzt und dabei die organisatorischen, technischen und rechtlichen Determinanten beachtet werden.

Kriterien für die Sammlungstätigkeit können sich zunächst aus der jeweiligen Plattform selbst vor dem Hintergrund der Bedeutung für den Sammlungsauftrag ergeben. Die unterschiedlichen Designs und Zielgruppen von Kanälen, die Verbreitung in einzelnen Domänen und ihre darüber hinausgehende Rezeption sind demnach Teil einer Überlieferungsstrategie.

Neben der Auswahl von Plattformen ist schließlich auch zu definieren, welche Inhalte hieraus übernommen werden sollen. Archive verfolgen einen provenienzbezogenen Ansatz, d. h., sie richten ihre Sammlungsaktivitäten entlang der die Unterlagen produzierenden Person oder Organisation aus. Auf diese Weise entsteht insbesondere bei größeren Organisationen entlang ihrer Strukturen der Prozess der Überlieferungsbildung. Funktionale Accounts können bspw. durch private ergänzt werden, wenn die dahinterstehenden Personen in exponierter Rolle Social-Media-Aktivitäten betreiben. Wird ein derartiges Modell allzu starr angewandt, besteht freilich die Gefahr, dass einzelne Accounts mit besonders hoher Sichtbarkeit nicht eingeschlossen sind.

Die Auswahl setzt ggf. unterschiedliche technische Zugangspunkte voraus. Dies bezieht sich etwa auf Selbstexport-Funktionen, die nur die eigenen Postings, aber nicht die erhaltenen Reaktionen und Kommentare umfassen. Darüber hinaus können Konjunkturen oder thematische Zuschnitte von Interesse sein – also ein pertinenzbezogener Ansatz. Neben politischen Aktivitäten und einzelnen Ereignissen wie etwa Kriegen oder Wahlen bildet bspw. der Hashtag #MeToo auf Twitter/X große Teile der zugehörigen Debatte um sexuelle Übergriffe in der Medienbranche und darüber hinaus ab, die durch die Taten des Hollywood-Produzenten Harvey Weinstein ausgelöst wurde. Die Breite dieser Debatte kann mit einem thematisch-inhaltlichen Ansatz besser abgebildet werden als durch die Sicherung ein-

57 Vgl. Gesetz über die Deutsche Nationalbibliothek. Die Archive der politischen Stiftungen haben sich in einem gemeinsamen Sammlungs- und Dokumentationsprofil allgemein gefasst und „Webseiten und Accounts in sozialen Netzwerken“ als „Herausforderungen“ identifiziert. Konrad-Adenauer-Stiftung et al. (Hrsg.), Dokumentations- und Sammlungsprofil der Archive der politischen Stiftungen in Deutschland, München 2022, S. 3, www.fes.de/index.php?eID=dumpFile&t=f&f=73521&token=edb37f3d89238b25859e60493c64034e52265c47

zelner Kanäle. In diesem Sinn kann auch das Plädoyer verstanden werden, einen „more targeted approach“⁵⁸ zu verfolgen.

Social-Media-Archivierung erfordert neue Wege – in technischer, rechtlicher und organisatorischer Hinsicht. Darüber hinaus fordert die Schnelllebigkeit sozialer Medien Gedächtniseinrichtungen regelrecht heraus und kontrastiert auffällig mit langfristigen Sammlungsstrategien. Das konstitutive Merkmal der interagierenden Netzwerke kann jedoch auch zu dessen Vorteil gewendet werden, wenn Kooperationen gesucht und gefestigt werden. Insbesondere auf Social-Media-Korpora basierende Forschungsprojekte umfassen oftmals zu diesem Zweck eigens gesicherte Quellen, die den Charakter von Forschungsdaten haben. Diese Schnittmengen gilt es perspektivisch stärker in den Blick zu nehmen und technisch-organisatorische Lösungen für einen möglichst ungehinderten Zugang zu schaffen.

6 Zusammenarbeit und Vernetzung

Da die Verbreitung sozialer Medien – zumindest potenziell – grenzenlos ist, sind der Zusammenarbeit von Kulturerbe-Einrichtungen in diesem Bereich ebenfalls keine Grenzen gesetzt. Zwar können nationale gesetzliche Regulierungen voneinander abweichen, doch insbesondere technische und fachliche Herausforderungen werden von einer globalen Community diskutiert. Das International Internet Preservation Consortium (IIPC) ist ein internationaler Zusammenschluss zur Entwicklung von fachlichen Lösungen, vielfältigen Informationsressourcen und einschlägigen Konferenzen im Bereich Webarchivierung, dem die Sicherung von Social-Media-Kanälen ebenfalls zugerechnet wird.⁵⁹ Zudem werden auf internationalen und nationalen Konferenzen, die ausschließlich oder mindestens schwerpunktmäßig die digitale Langzeitarchivierung zum Gegenstand haben, regelmäßig Diskussionen zur Sicherung von sozialen Medien geführt. Die jährlich stattfindende Konferenz iPres ist die weltweit bedeutsamste in diesem Bereich.⁶⁰ Die Digital Preservation Coalition (DPC) organisiert Veranstaltungen und Fortbildungen, verfasst eigene Publikationen und Reports zur digitalen Langzeitarchivierung.⁶¹ Hervorgegangen aus einem Forschungsprojekt, hat die Open Preservation Foundation (OPF) neben Netzwerkaktivitäten und Veröffentlichungen auch verschiedene Tools unter

⁵⁸ Vlassenroot et al., *Web-Archiving and Social Media*, S. 115.

⁵⁹ Vgl. netpreserve.org/

⁶⁰ Vgl. ipres-conference.org/

⁶¹ Vgl. www.dpconline.org/

Open-Source-Lizenzen zur Verfügung gestellt.⁶² In Deutschland hat das Kompetenznetzwerk zur digitalen Langzeitarchivierung nestor wiederholt Veranstaltungen zur Webarchivierung durchgeführt.⁶³ Daneben entfalten einzelne Akteur*innen Initiativen, die ebenfalls vom Willen der Verfestigung und des Ausbaus der institutionenübergreifenden Kooperation geprägt sind.

Literaturverzeichnis

- Ammendola, Andrea: Webarchivierung in NRW aus Sicht der Universitäts- und Landesbibliothek Münster, Masterarbeit Technische Hochschule Köln, 2020, nbn-resolving.org/urn:nbn:de:hbz:79pbc-opus-16219 (26.3.2024).
- Ben-David, Anat: Gegenarchivierung von Facebook, in: Schwarzenegger, Christian/Koenen, Erik/Pentzold, Christian/Birkner, Thomas/Katzenbach, Christian (Hrsg.): Digitale Kommunikation und Kommunikationsgeschichte. Perspektiven, Potentiale, Problemfelder, Berlin 2022, S. 133–155, doi.org/10.48541/dcr.v10.5 (26.3.2024).
- Bundesministerium der Justiz/Bundesamt für Justiz: Gesetz über die Deutsche Nationalbibliothek (DNBG) Nichtamtliches Inhaltsverzeichnis, 2006, www.gesetze-im-internet.de/dnbg/BJNR133800006.html (11.10.2024).
- Brügger, Niels: The Archived Web: Doing History in the Digital Age, Cambridge/Massachusetts 2018.
- Brügger, Niels: Understanding the Archived Web as a Historical Source, in: Brügger, Niels/Milligan, Ian (Hrsg.): The SAGE Handbook of Web History, London 2019, S. 16–29.
- Day Thomson, Sara/Beagrie, Neil: Preserving Social Media, DPC Technology Watch Reports, 2016, doi.org/10.7207/twr16-01 (26.3.2024).
- Digital Preservation Coalition, www.dpconline.org/ (11.10.2024).
- Digital Preservation Coalition, Digital Preservation Handbook, 2. Aufl., Glasgow 2015, www.dpconline.org/handbook (26.3.2024).
- Eckert, Paulina: Tagungsbericht: Zwischen Link-Graphen und Paragraphen – Archivierung und Nutzung von Netz-Quellen, in: H-Soz-Kult, 3.2.2023, www.hsozkult.de/conferencereport/id/fdkn-133549 (26.3.2024).
- Go FAIR, www.go-fair.org/ (11.10.2024).
- Healy, Sharon/Byrne, Helena: Scholarly Use of Web Archives Across Ireland: The Past, Present & Future(s), WARCnet Special Report, Aarhus 2023, cc.au.dk/fileadmin/dac/Projekter/WARCnet/Healy_Byrne_Scholarly_Use_01.pdf (26.3.2024).
- Hegarty, Kieran: The invention of the archived web: tracing the influence of library frameworks on web archiving infrastructure, in: Internet Histories 6/4 (2022), S. 432–451, doi.org/10.1080/24701475.2022.2103988 (26.3.2024).
- Heseler, Jörg/Büttner, Alexandra/Arnold, Matthias: Grundlagen der digitalen Langzeitarchivierung. Eine Handreichung zur digitalen Langzeitarchivierung aus Perspektive der NFDI4Culture Community (1.0.2), 17.1.2024, doi.org/10.5281/zenodo.10511404 (26.3.2024).

⁶² Vgl. openpreservation.org/.

⁶³ Vgl. www.langzeitarchivierung.de/Web/nestor/DE/Home/home_node.html.

- Hockx-Yu, Helen: Archiving Social Media in the Context of Non-print Legal Deposit, IFLA WLIC, Lyon 2014, library.ifla.org/id/eprint/999 (26.3.2024).
- IIPC – International Internet Preservation Consortium, netpreserve.org/ (11.10.2024).
- iPres – International Conference on Digital Preservation, ipres-conference.org/ (11.10.2024).
- Jaillant, Lise: How Can We Make Born-Digital and Digitised Archives More Accessible? Identifying Obstacles and Solutions, in: *Archival Science International Journal on Recorded Information* 22 (2022), S. 417–436, doi.org/10.1007/s10502-022-09390-7 (26.3.2024).
- Jerusalem, Michael/Malwitz, Marc/Worm, Peter: Ein neuer Ansatz für die Langzeitarchivierung von Twitter-Accounts, in: *Hypotheses* (blog), 9.3.2021, archive20.hypotheses.org/10031 (26.3.2024).
- Kaczmarek, Lars/Mayr, Philipp/Vatrapu, Ravi/Bleier, Arnim/ Blumenberg, Manuela/Gummer, Tobias/Hussain, Abid/Kinder-Kurlanda, Katharina/Manshaei, Kaveh/Thamm, Mark/Weller, Katrin/Wenz, Alexander/Wolf, Christof: Social Media Monitoring of the Campaigns for the 2013 German Bundestag Elections on Facebook and Twitter (GESIS-Working Papers 31), 2014, doi.org/10.48550/arXiv.1312.4476 (26.3.2024).
- Klimpel, Paul: Kulturelles Erbe digital. Eine kleine Rechtsfibel, Berlin 2020, digis-berlin.de/wp-content/uploads/2020/09/digiS_PKlimpel_Rechtsfibel.pdf (26.3.2024).
- Klimpel, Paul/Rack, Fabian: Einschätzung der rechtlichen Rahmenbedingungen für die Archivierung von Social-Media-Inhalten im Archiv der sozialen Demokratie, in: Walz, Annabel/Marquet, Andreas (Hrsg.): *Sicher Sichern? Social Media-Archivierung aus rechtlicher Perspektive im Archiv der sozialen Demokratie*, Beiträge aus dem Archiv der sozialen Demokratie Bd. 17, Bonn 2022, S. 15–48, library.fes.de/pdf-files/adsd/19590.pdf (26.3.2024).
- Klindt Myrvoll, Arndt: Social Media Archiving at the Royal Danish Library – Reflections and Status Quo, Wanted: Social Media Data – Archiving Practices and Research Use, Brüssel 2022, in: @KBR-BEL: „Social Media Archiving from an Institutional Point of View“, YouTube-Video, 7.10.2022, www.youtube.com/watch?v=skVeVxQQVxo (26.3.2024).
- Köhn, Katharina: Parteienstiftungen und Social Media. Facebook-Archivierung mit OWA und Laurentius, Webarchive und Social Media, Baden-Baden 2013, files.dnb.de/nestor/veranstaltungen/Webarchivierung2013-10-18/3_Hanns-Seidel-Stiftung.pdf (26.3.2024).
- Konrad-Adenauer-Stiftung/Friedrich-Ebert-Stiftung/Friedrich-Naumann-Stiftung/Hanns-Seidel-Stiftung/Heinrich-Böll-Stiftung/Rosa-Luxemburg-Stiftung (Hrsg.): *Dokumentations- und Sammlungsprofil der Archive der politischen Stiftungen in Deutschland*, München 2022, www.fes.de/index.php?eID=dumpFile&t=f&f=73521&token=edb37f3d89238b25859e60493c64034e52265c47 (26.3.2024).
- Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen: Webarchivierung. Eine Studie der KOST, 2016, kost-ceco.ch/cms/index.php?web_archiving_study_de (26.3.2024).
- Kortyla, Stephanie: Archivierung von Daten aus Twitter. Twitter-eigene Exportschnittstelle (2018–2021), Archivierung von Unterlagen aus digitalen Systemen, Vortrag AuDs 2021 Köln, 2.3.2021, www.sg.ch/content/dam/sgch/kultur/staatsarchiv/auds-2021/auds-camp-i/Vortrag_Kortyla_Twitter_Archivierung.pdf (26.3.2024).
- Littman, Justin/Chudnov, Daniel/Kerchner, Daniel/Peterson, Christie/Tan, Yecheng/Trent, Rachel/Vij, Rajat/Wrubel, Laura: API-based Social Media Collecting as a Form of Web Archiving, in: *International Journal on Digital Libraries* 19/1 (2018), S. 21–38, doi.org/10.1007/s00799-016-0201-7 (26.3.2024).
- Maemura, Emily/Worby, Nicholas/Milligan, Ian/Becker, Christoph: If These Crawls Could Talk: Studying and Documenting Web Archives Provenance, in: *Journal of the Association for Information Science and Technology* 69/10 (2018), S. 1223–1233, doi.org/10.1002/asi.24048 (26.3.2024).

- National Digital Stewardship Alliance: 2022 Web Archiving Survey Results (2023), doi.org/10.17605/OSF.IO/N5MYR (26.3.2024).
- Naumann, Kai: Gemeinsam stark. Web-Archivierung in Baden-Württemberg, Deutschland und der Welt, in: *Der Archivar* 65/1 (2012), S. 33–41.
- Nestor – Kompetenznetzwerk: Langzeitarchivierung und Langzeitverfügbarkeit digitaler Ressourcen in Deutschland, www.langzeitarchivierung.de/Webs/nestor/DE/Home/home_node.html (26.3.2024).
- Neuberger, Christoph: Digitale Öffentlichkeit und liberale Demokratie, in: *Aus Politik und Zeitgeschichte* 72/10–11 (2022), S. 18–25.
- Open Preservation Foundation (OPF), openpreservation.org/.
- Osterberg, Gayle: Update on the Twitter Archive at the Library of Congress, in: *Timeless: Stories from the Library of Congress* (blog), 26.12.2017, blogs.loc.gov/loc/2017/12/update-on-the-twitter-archive-at-the-library-of-congress-2 (26.3.2024).
- Reimann, Manfred/Stumpf, Marcus: Grundfragen und Organisation des Archivwesens, in: Stumpf, Marcus (Hrsg.): *Praktische Archivkunde. Ein Leitfaden für Fachangestellte für Medien- und Informationsdienste – Fachrichtung Archiv*, 4. Aufl., Münster 2018, S. 25–56.
- Schlesinger, Claus-Michael/Woldering, Britta: Tweets archivieren. Initiative, Umsetzung, Ausblick, *Show & Tell – Social Media-Daten in der Forschungspraxis*, 2023, nfdi4culture.de/fileadmin/user_upload/documents/Show_and_Tell/2023-05-26_NFDI_show_tell_Schlesinger_Woldering.pdf (26.3.2024).
- Schoger, Astrid/Weimer, Konstanze: Das Dateiformat WARC für die Webarchivierung, nbn-resolving.de/urn:nbn:de:0008-2021042614 (26.3.2024).
- Schuck, Johannes: Schritt für Schritt auf neuen digitalen Wegen. Webseiten- und Social-Media-Kanal Archivierung im Stadt- und Stiftsarchiv Aschaffenburg, in: *Archivpflege in Westfalen-Lippe* 96 (2022), S. 17–20, nbn-resolving.de/urn:nbn:de:hbz:6:2-1646179 (26.3.2024).
- Stabel, Florian/Tripp, Sebastian: Hauptstaatsarchiv Wiesbaden übernimmt Twitter-Daten von Thorsten Schäfer-Gümbel, Hessisches Landesarchiv, Wiesbaden 2022, web.archive.org/web/20220419214334/https://landesarchiv.hessen.de/twitter-daten_schaefer-guembel (archiviert am 19.4.2022, Archivversion abgerufen am 26.3.2024).
- UNESCO, Charta zur Bewahrung des digitalen Kulturerbes, Paris 2003, www.unesco.de/sites/default/files/2018-03/2003_Charta_zur_Bewahrung_des_digitalen_Kulturerbes.pdf (26.3.2024).
- Vlassenroot, Eveline/Chambers, Sally/Lieber, Sven/Michel, Alejandra/Geeraert, Friedel/Pranger, Jessica/Birkholz, Julie/Mechant, Peter: Web-Archiving and Social Media: an Exploratory Analysis, *Call for Papers Digital Humanities and Web Archives – A Special Issue of International Journal of Digital Humanities*, in: *International Journal of Digital Humanities* 2/1–3 (2021), S. 107–128, doi.org/10.1007/s42803-021-00036-1 (26.3.2024).
- Walz, Annabel/Marquet, Andreas (Hrsg.): Sicher Sichern? Social Media-Archivierung aus rechtlicher Perspektive im Archiv der sozialen Demokratie, *Beiträge aus dem Archiv der sozialen Demokratie* Bd. 17, Bonn 2022, library.fes.de/pdf-files/adsd/19590.pdf (26.3.2024).
- Weiß, Eva-Maria: Twitter-API: 1500 Tweets kostenlos, 100 US-Dollar für das günstigste Abo, in: *heise online*, 30.3.2023, www.heise.de/news/Twitter-bestaetigt-1500-Tweets-kostenlos-100-US-Dollar-fuer-das-guenstigste-Abo-8239039.html (26.3.2024).
- Weyns, Katrien: Towards Best Practices for the Archiving of Social Media by Private Archival Institutions in Flanders, *Wanted: Social Media Data – Archiving Practices and Research Use*, Brüssel 2022, in: @KBR, „Social Media Archiving from an Institutional Point of View“, YouTube-Video, 7.10.2022, www.youtube.com/watch?v=skVeVxQQVxo (26.3.2024).

Wiedemann, Gregory: Describing Web Archives: A Computer-Assisted Approach, in: *Journal of Contemporary Archival Studies* 6, Article 31 (2019), elischolar.library.yale.edu/jcas/vol6/iss1/31/ (26.3.2024).

Worm, Peter: Neue Ansätze für die Archivierung eines Twitter-Accounts, in: *Archivpflege in Westfalen-Lippe* 96 (2022), S. 26–39, nbn-resolving.de/urn:nbn:de:hbz:6:2-1646179 (26.3.2024).