

# File not found.

Timo Rychert. Master-Thesis 2021.

## Einleitung: Broken Links Everywhere

In der täglichen Nutzung des Internets bewegen wir uns meist auf gewohnten Bahnen. Wir suchen in der Regel solche Websites auf, die wir für die Erledigung unserer Alltags- und Berufsaufgaben, zum Vergnügen oder zur Zerstreuung ansteuern: Nachrichtenportale, Social-Media-Plattformen, Webmail-Anbieter, Streaming-Dienste etc. Die entsprechenden digitalen Infrastrukturen sind wie selbstverständlich rund um die Uhr verfügbar, gut ausgebaut und geben uns zielgerichtet genau die Inhalte an die Hand, die wir beim Besuch dieser Websites voraussichtlich benötigen.

In Nachrichtenportalen finden wir so die aktuellsten Nachrichtenmeldungen, wir klicken uns durch neue Storys und Medieninhalte digitaler Kontakte und interagieren live mit anderen Nutzer\*innen.

In dem Moment aber, wo wir uns auf die Suche nach etwas Unvorhergesehenem, etwas nicht Tagesaktuellem, etwas Verborgenem machen, wandeln wir auf weniger ausgetretenen Online-Pfaden und finden uns – oft nach nur ein, zwei Abbiegungen – erstaunlich schnell in der digitalen Peripherie.

Hier wird dieses scheinbar gut funktionierende Netz, das doch angeblich so fluide, formbare und ewig anpassungsfähige Web, plötzlich brüchig und lückenhaft: Inhalte sind – wenn überhaupt – nur noch mit Mühe auffindbar, interaktive Anwendungen funktionieren nicht mehr und Links laufen ins Leere, da die verlinkten Websites schlicht nicht mehr existieren.

Schaut man sich ältere Online-Artikel an, so ist es fast der Regelfall, dass extern verlinkte Seiten heute nicht mehr aufrufbar sind. Aber auch bei solchen Inhalten, die eigentlich Anspruch auf Aktualität erheben – wie zum Beispiel Wikipedia-Artikel – sind Weblinks oder die Einzelnachweise oft nicht mehr abzurufen, da eben auch diese Lexikon-Einträge schon vor längerer Zeit angelegt wurden, doch das Web sich stetig weiter wandelt.

Diese Kurzlebigkeit hängt unmittelbar mit dem Medium des Digitalen zusammen. Anders als allgemein angenommen, ist es bei digitalen Medien eine große Herausforderung sie zu bewahren und somit über einen längeren Zeitraum nicht nur verfügbar, sondern auch nutzbar zu halten.

Betrachtet man die folgende Infografik genauer, stellt man fest, dass das Problem weit verbreitet ist und wir vermutlich alle schon mit dem Phänomen fehlender Inhalte und visueller Leerstellen konfrontiert waren ([→ Abb. 1](#)):



Abb. 1: Infografik (30.03.2021)

Das ist umso ärgerlicher, als dass ich nicht einfach nur den Inhalt nicht sehe, sondern ich bekomme auch genau mit, dass ich ihn nicht sehe: Die Leerstellen lassen uns wissen, dass hier einmal etwas war, was inzwischen nicht mehr vorhanden ist. Das obenstehende Bild führt in seiner jetzigen Form vermutlich zu einem gewissen Maß an Unbehagen, denn es vermittelt das Gefühl: Diese Website ist kaputt. Im vorliegenden Fall habe ich das „fehlende Bild“ absichtlich herbeigeführt, indem ich die Website auf eine nicht existierende Bilddatei verwiesen habe.

Dass uns aber das Ergebnis, die Leerstelle, nur allzu vertraut ist, erinnert uns daran, dass die Inhalte online sehr häufig tatsächlich fehlen. Die zugehörigen Websites hatten vielleicht an Bedeutung verloren und wurden deshalb nicht mehr richtig gepflegt und „verfallen“ jetzt langsam.

Große Teile unserer alltäglichen Netzkultur gehen somit unwiederbringlich verloren, wenn nicht aktiv Sorge getragen wird, dass diese Websites und Medieninhalte in irgendeiner Form archiviert und somit für die Nachwelt gesichert werden. Vertreter\*innen verschiedenster Fachbereiche sind sich dieser ungewollten Ephemerität zunehmend bewusst und versuchen dagegen anzugehen: In den heute noch weitestgehend analogen Archiven und Bibliotheken beginnt man sich Gedanken darüber zu machen, wie man digitales Kulturgut erhalten kann. Auch die Kunstwelt beschäftigt das Problem, wie man digitale Kunst oder gar Net Art konserviert, bzw. ob das überhaupt gelingen kann oder der richtige Ansatz ist. Denn im digitalen Raum, wo alles beliebig vervielfältigt werden kann, stellt sich im Kontext der Kunst natürlich mehr denn je die Frage: Was ist das Original und kann es so etwas wie ein digitales Original überhaupt geben?

Die vorliegende Thesis will einen Blick auf dieses digitale Archivieren und die damit verbundenen Probleme werfen. Da diese Frage ein ex-

trem weites Feld umfasst, werde ich mich dabei auf einen Teilbereich konzentrieren: das Archivieren von Websites und Webinhalten. Meine Wahl fällt deshalb auf Websites, weil hier das Grundproblem digitaler Vergänglichkeit besonders gut sichtbar ist: Jede\*r von uns ist vermutlich schon einmal auf den Statuscode 404 gestoßen und hat sich über das Verschwinden der Inhalte geärgert. Die Problematik tritt somit meist dann erst auf, wenn es bereits zu spät ist: Die Website ist nun nicht mehr auffindbar. Wenn sie nicht in irgendeiner Form vorab archiviert worden ist, lebt sie von nun an nur in der Erinnerung ihrer Betrachter\*innen fort. Es gibt inzwischen eine ganze Bandbreite von Ansätzen und Tools, die versuchen diesem Phänomen entgegenzuwirken, und deren genauere Betrachtung sich lohnt. Keines dieser Tools hat sich bislang als die perfekte Lösung durchsetzen können, noch sieht es so aus, als könne dies irgendwann einmal geschehen. Die Websites selbst, vor allem aber die Anforderungen an ihre Archivierung, sind dafür zu unterschiedlich.

Mich führt auch ein persönlicher Zugang zum Thema: Als Gestalter sammle ich stetig visuelle und konzeptionelle Referenzen. Dabei war ich schon viele Male mit der Frage konfrontiert, auf welche Weise ich denn am besten eine Website „abspeichern“ kann, um sie später bei Bedarf – zum Beispiel bei der Recherche für ein neues Projekt – schnell und verlässlich wieder zur Hand zu haben. Und letztlich muss ich mir als Produzent digitaler Inhalte und Websites Gedanken darüber machen, wie ich diese Websites dokumentieren will und ob und wie ich diese Arbeiten in die Zukunft hinüberretten kann.

Wer sich mit dem Vorgang des Archivierens einzelner digitaler Archivstücke beschäftigt, muss aber auch die resultierenden Archive in den Blick nehmen. Funktionieren diese Archive? Wenn ja, auf welche Weise und vor allem in wessen Sinne? Archive geben immer eine bestimmte Wirklichkeit vor: Sie machen zeitlich, räumlich und inhaltlich Unzusammenhängendes gleichzeitig und betrachten es aus einem bestimmten Blickwinkel. Welche Sorge können und müssen wir als Nutzer\*innen oder Ersteller\*innen solcher Archive tragen, um dafür zu sorgen, dass sie offen sind, sowohl im Sinne des Zugangs und der Nutzbarkeit, als auch im Sinne einer möglichen Neu-Interpretation, der Eröffnung anderer Lesarten oder gar einer kompletten Umnutzung? Wie können wir sicherstellen, dass wir die Einschränkungen, denen wir beim Archivierungsvorgang unterliegen, auch klar kenntlich machen, um nicht den Eindruck einer vermeintlichen Authentizität archivierter „Originale“ zu erwecken? Beispielaufgabe werde ich dazu einen Blick auf bestehende Webarchiv-Plattformen werfen und anhand dieser einige Aspekte dieser Probleme beleuchten.

Die vielzitierte Formel „Das Internet vergisst nie“ hat sicherlich ihre Begründung als datenschutzrelevante Ermahnung zur sparsamen Preisgabe persönlicher Informationen. Allerdings stellt sich, nun da das Internet kein Neuland mehr ist, zunehmend heraus, dass – nimmt man

das Große und Ganze in den Blick – vor allem auch das Gegenteil richtig ist: Das Internet vergisst *immer*. Es befindet sich in einem permanenten Verfallsprozess und nur solche Inhalte überdauern die Zeit, die weiterhin relevant sind, in halbwegen regelmäßigen Abständen von Nutzer\*innen aufgesucht werden und daher von den Anbietern dieser Inhalte entsprechend gepflegt und vorgehalten werden. Aber selbst wenn diese Inhalte vorerst erhalten bleiben, kommen mit jedem Augenblick riesige Mengen neuer Daten hinzu, die die alten überlagern und somit unsichtbar machen.

Die vorliegende Thesis soll diese Vorgänge einordnen und Wege aufzeigen, diesem Vergessen etwas entgegenzusetzen.

## 1. Websites speichern

Wie kann man eine Website archivieren? Dieser Frage möchte ich nachgehen, indem ich an einem konkreten Fallbeispiel untersuche, wie ich vor ein paar Jahren selbst eine Website archiviert habe. Anhand dieses Archivierungsvorgangs und am jetzt vorliegenden Ergebnis dieses Prozesses lassen sich bereits viele Problemstellungen des Archivierens von Webinhalten aufzeigen. Dabei sei vorangestellt, dass es hier nicht um eine Musterlösung gehen soll; da die Anforderungen eines jeden Konservationsvorgangs jeweils unterschiedlich sind, kann es eine solche Musterlösung ohnehin nicht geben. Mir ist zudem wichtig, ein echtes „Archivstück“ zu verwenden und nicht etwa eines, was ich extra für diese Untersuchung angefertigt habe. So lässt sich am besten feststellen, welche Informationen sich noch heute aus dem archivierten Inhalt gewinnen lassen und welche hingegen verloren sind.

Hier die archivierte Website.

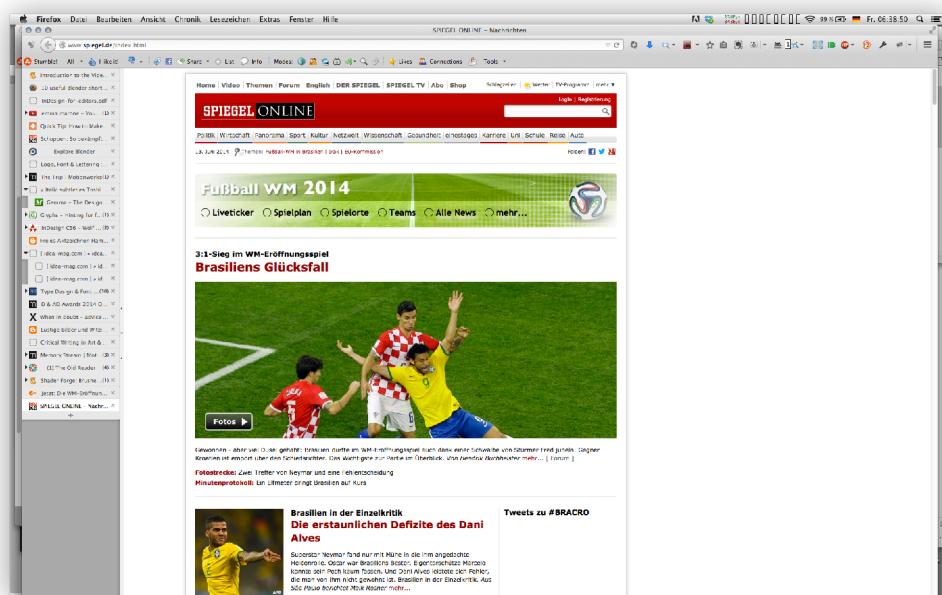


Abb. 2: Archivierte Website, Quelle: Siehe Text.

Ich lade dazu ein, zunächst ohne weiteren Kontext die Website zu betrachten, um herauszufinden, welche Informationen über das Archivstück und über den Archivierungsprozess man sich selbst erschließen kann (ein Klick auf die Seite erlaubt das Betrachten in voller Größe).

Eine Aufschlüsselung:

Zu sehen ist ein Screenshot des Onlinenachrichtenportals **Spiegel Online**. Die Domain<sup>1</sup> **spiegel.de** gehört zu den am häufigsten aufgerufenen Domains in Deutschland (Platz 31 im Alexa Ranking, Stand Januar 2021) ([Alexa, 2021](#)), weshalb viele Betrachter\*innen den Screenshot visuell schnell einordnen können sollten. Der Titel des Portals ist ebenfalls im Bild vorhanden, so dass diese Zuordnung auch ohne Vorwissen erfolgen kann.

Der Screenshot zeigt nicht nur den Inhalt der Website, sondern den ganzen Bildschirminhalt zum Zeitpunkt der Aufnahme. So lassen sich weitere Details des Archivierungsmoments erschließen: Die Seite wurde im Firefox-Browser unter Mac OS X geöffnet. Am oberen Rand des Bildes lässt sich ablesen, dass die Aufnahme um 06:38:50 Uhr gemacht wurde, die Website verrät das Jahr, bei genauerem Hinsehen auch das genaue Datum: 13. Juni 2014. Zeitgleich waren noch zahlreiche andere Websites geöffnet, deren Titel man am linken Rand des Browserfensters ablesen kann und die somit ein Stück weit Aufschluss darüber geben könnten, womit sich der\*die Archivar\*in zum damaligen Zeitpunkt sonst beschäftigt hat.

Der Aufbau des Bilds lässt vermuten, dass die Aufnahme nicht speziell inszeniert war: Wäre es darum gegangen, den Inhalt möglichst sauber zu präsentieren, so hätte man die Seite in den Vollbild-Modus schalten können, so dass Tabs, Browsersrahmen und Menüleiste des Betriebssystems gar nicht erst sichtbar wären. Das Browserfenster ist nicht einmal zur vollen Bildschirmgröße aufgezogen, man kann rechts und links Anwendungsfenster anderer Programme erahnen. Die Aufnahme zeigt also mutmaßlich einen relativ „authentischen“ Bildschirmmoment, so wie er in der alltäglichen Internetnutzung hätte stattfinden können.

Bei genauer Betrachtung gibt es allerdings einen Part in diesem Aufbau, der doch inszeniert scheint: Die **Spiegel Online**-Website ist genau an ihrem Einstiegspunkt abgebildet, also so, wie man sie unmittelbar nach dem Aufrufen der Seite sehen würde. Es wurde noch nicht heruntergescrollt oder auf eine Artikel-Unterseite gewechselt. Das könnte natürlich Zufall sein, aber da dieser abgebildete Moment während einer typischen Website-Nutzung in der Regel nur sehr kurz ausfällt, ist es viel naheliegender, dass die Website extra für das Aufnehmen des Screenshots aufgerufen wurde, dass es also Ziel der Bildschirmaufnahme war, einen Screenshot von **Spiegel Online** anzufertigen.

Zu welchem Zweck wurde der Screenshot gemacht? Aus dem Bildinhalt lassen sich nun nicht viel mehr Informationen ablesen. Hätte ich

den 3:1-Sieg des brasilianischen Herren-Teams im Eröffnungsspiel der Fußball-WM 2014 dokumentieren wollen, so wäre ich vermutlich zunächst auf die Seite des entsprechenden Artikels gewechselt. In diesem Fall kann die digitale Datei selbst weiter Aufschluss geben: Die Datei – ich habe sie bewusst so belassen, wie ich sie kürzlich auf meiner Festplatte vorgefunden habe – heißt [IdasGeburt\\_SpiegelOnline\\_01.png](#).

Tatsächlich habe ich diese Bildschirmaufnahme am Tag der Geburt meiner Tochter gemacht, um das Zeitgeschehen an diesem Tag zu dokumentieren. Der Gedanke geht auf die Idee zurück, eine Tageszeitung vom Geburtstag des Kindes aufzuheben, um sie ihm später im Erwachsenenalter zu schenken. Von meinem eigenen Geburtstag existiert eine solche Tageszeitung und auch ich hatte mir vorgenommen, für mein Kind ein solches Zeitdokument aufzuheben. Als der Tag dann gekommen war, konnte ich nicht absehen, ob ich es schaffen würde eine Tageszeitung zu kaufen und ich entschied mich kurzerhand und in Anbetracht des neugeborenen Digital Natives, das zeitgemäße Pendant zur Tageszeitung aufzuheben: die Frontseite von **Spiegel Online**.

Die Dokumentation bestand letztlich aus drei weiteren Screenshots, die die Website jeweils etwas weiter heruntergescrollt zeigen. So kann man die wichtigsten Nachrichten zum damaligen Zeitpunkt ablesen ([→ Abb. 3](#)). Unterseiten – einzelne Artikel also – habe ich allerdings nicht archiviert; vermutlich genügte mir für den Einsatzzweck ein schneller Überblick der damaligen Nachrichtenlage.

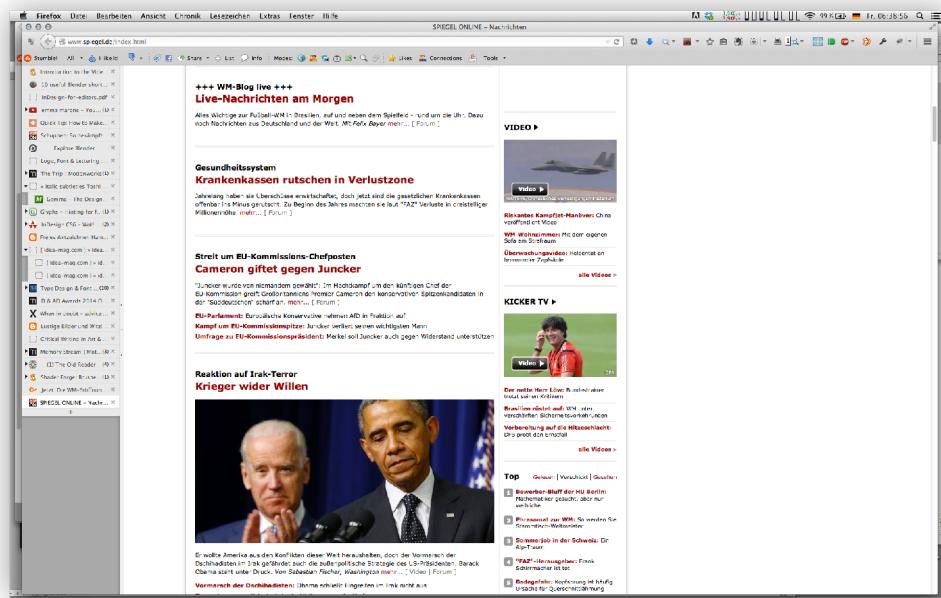


Abb. 3: spiegel.de (Screenshots, 13.06.2014) (1/3)

## Bestandsaufnahme

Nach dieser Untersuchung des Archivierungsvorgangs selbst lohnt es sich nun einen Blick darauf zu werfen, welche Web-Inhalte hier eigent-

lich archiviert wurden und inwiefern sich aus diesen Informationen über die [spiegel.de](#)-Website zum Zeitpunkt des 13.06.2014 ablese lassen. Ruft man [spiegel.de](#) heutzutage (Stand 02.02.2021) auf, so stellt man gleich einige Unterschiede zur archivierten Version fest: Die augenfälligste Änderung ist dabei die visuelle Gestaltung der Seite, die sich seither geändert hat (→ Abb. 4). Zudem stellt man fest, dass die Seite heute nicht mehr unter dem Titel **Spiegel Online** läuft, sondern inzwischen – analog zum Printmagazin – schlicht **Der Spiegel** heißt. Diese Änderung – die auch das neue Layout mit sich brachte – wurde infolge der Zusammenlegung der Online- und Printredaktionen am 8. Januar 2020 vorgenommen ([MEEDIA, 2020](#))<sup>12</sup>.

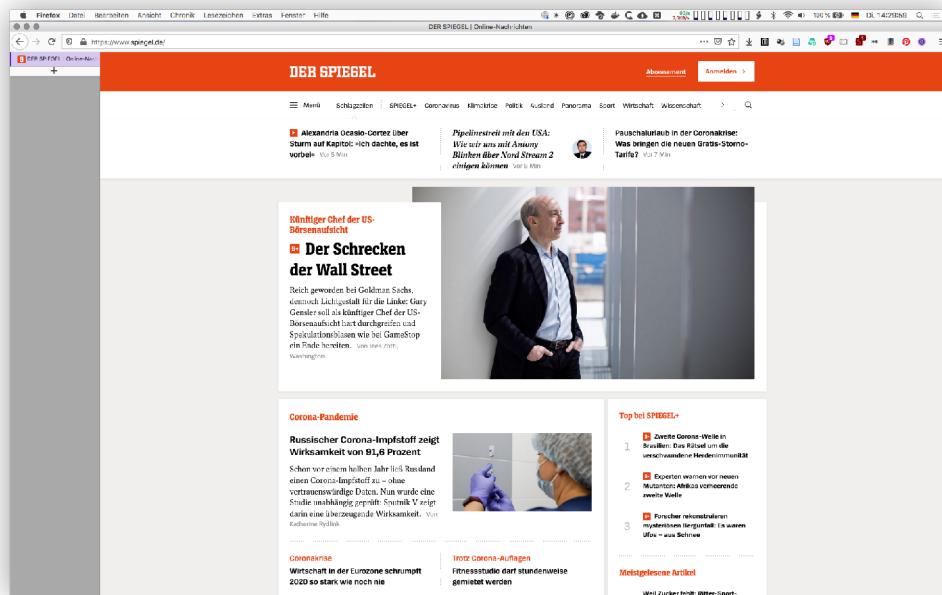


Abb. 4: [spiegel.de](#) – Aktuelles Layout (Screenshot, 02.02.2021)

Allein schon diese Umstellung sorgt dafür, dass der archivierte Screenshot, wenn auch nur ein paar Jahre alt, bereits jetzt „historisch“ ist: Die dargestellte Seite lässt sich in dieser Form heute nicht mehr aufsuchen. Nachrichtenportale wie [spiegel.de](#) sind allerdings ohnehin keine statischen Seiten, sondern der Inhalt ihrer Frontseite ändert sich permanent, so dass ein Screenshot immer nur eine Momentaufnahme ist, selbst, wenn er erst vor ein paar Minuten erstellt wurde.

Wie aber verhält es sich mit den Inhalten? Sucht man heute nach den entsprechenden Artikel-Überschriften, so finden sich problemlos alle Artikel der Hauptspalte (z.B. [Brasiliens Glücksfall](#), [Cameron giftet gegen Juncker](#), [Diese herrliche Lust am Untergang](#), alle abgerufen 02.02.2021), einzig der „WM-Blog live“ scheint verschwunden. Ebenso fehlen die Kommentare der Leser\*innen unter den Artikeln, obwohl die entsprechende Funktion augenscheinlich noch da ist: Im Zuge der Umstellung hatte [spiegel.de](#) angekündigt, alte Kommentare zu löschen und durch ein neues Kommentarsystem zu ersetzen ([Der Spiegel, 2020](#))<sup>13</sup>.

Ob die Artikel, wenn ich sie aufrufe, seit dem Archivierungszeitpunkt unverändert sind, lässt sich nicht zweifelsfrei feststellen. Data Scientist

**David Kriesel** hat in seinem Vortrag **SpiegelMining – Reverse Engineering von Spiegel-Online** ([Kriesel, 2016](#)) darauf hingewiesen, dass viele Artikel auf **spiegel.de** kurz nach Erscheinen noch mehrere Male geändert werden. Wäre es mir also um historische Korrektheit gegangen, so hätte ich auch die Unterartikel selbst archivieren müssen.

Nichtsdestotrotz: Obwohl die Archivierung auf sehr schnelle und einfache Weise erfolgte, lässt sich aus dem Ergebnis bereits eine große Menge an Informationen ablesen. Wenn ich die Screenshots in dieser Form eines Tages an meine Tochter übergebe, wird sie vermutlich in der Lage sein, die Unterartikel ebenfalls aufzufinden und sich somit über die Nachrichten am Tag ihrer Geburt zu informieren. Sie erhält zudem einen schnellen visuellen Eindruck vom damals größten Nachrichtenportal und kann herausfinden, welcher sprachliche Stil und welche visuelle Bildästhetik damals typisch waren. In Bezug auf diese Punkte ist die Form der Archivierung also eine gutes Äquivalent zu meiner Zeitung. Auch diese kann ich heute betrachten und mir die gleichen Eindrücke verschaffen.

In einem wichtigen Aspekt stehen die Screenshots aber hinter der Zeitung zurück: Die Zeitung kann ich in die Hand nehmen, durchblättern und durchlesen. Die ursprüngliche Nutzungsform ist also vollumfänglich erhalten. Das lässt sich von der archivierten Website nicht behaupten: Ich kann sie nicht durchscrollen, ich kann auf keine Verlinkungen klicken, ich habe keine Ladezeiten, ich kann keinen Text markieren oder kopieren und nicht die Größe meines Browserfensters ändern. Kurz: Das Nutzungserlebnis, das mit dem Betrachten einer Website stets einhergeht, ist vollends verloren gegangen. Der Screenshot vermittelt nur eine entfernte Idee vom ursprünglichen Nutzungserlebenis des Betrachtens der Website.

Ob mich diese Lücken in der Aufzeichnung stören, oder ob ich bereit bin sie in Kauf zu nehmen, hängt vom Ziel meiner Archivierung ab. Im vorliegenden Fall funktioniert diese Form gut, in anderen kann sie unzureichend sein: Will ich nicht das Zeitgeschehen, sondern primär die Website selbst archivieren, dann kann es wichtig sein, die Navigationsstruktur oder interaktive Elemente mitzuerhalten. Mögliche Ansätze, um auch diese Aspekte zu archivieren, werden – gemeinsam mit einer Betrachtung der Nachteile – weiter unten erörtert.

## Hat es diese Website je gegeben?

Trotz der zuvor beschriebenen Unzulänglichkeiten bei der Archivierung mittels Screenshot könnte man zunächst denken, dass zumindest das Aussehen der Website in einer „authentischen“ und damit „neutralen“ oder „objektiven“ Weise festgehalten wurde. Ein Screenshot gibt schließlich den Bildschirminhalt exakt so wieder, wie er zum Zeitpunkt der Aufnahme ausgesehen hat.

Gegen diese Feststellung ist zunächst einmal nichts einzuwenden. Allerdings muss man sich bewusst machen, dass das, was auf dem Bildschirm zu sehen war, eine einzigartige Konfiguration der Website ist, die durch vielerlei Faktoren beeinflusst ist und bei anderen Betrachter\*innen, die die Website an diesem Morgen aufgerufen haben, anders ausgesehen haben könnte.

Zum exakten Zeitpunkt des Aufrufs standen bestimmte Artikel auf der Frontseite. Wenige Augenblicke zuvor oder später könnten schon andere Artikel an dieser Stelle gelistet worden sein. Je nachdem, in welchem Browser und unter welchem Betriebssystem man die Seite betrachtet, können sich Schriftart, Schriftgröße und Abstände im Layout unterscheiden. Bei unterschiedlichen Größen des Browserfensters wird sich das Layout entsprechend anpassen.

Auf der Seite ist rechts der Hauptspalte ein großer Weißraum zu sehen (→ Abb.2). Normalerweise schaltet **spiegel.de** hier großflächige, bewegte Werbeanzeigen, die bei meinem Screenshot allerdings durch einen Werbeblocker blockiert sind und somit nicht dargestellt werden. Man könnte also überspitzt formulieren, die Website habe es in dieser Form nie gegeben, da sie nachträglich von meinem Browser geändert wurde. Auch wenn man entgegenhält, dass ja genau das die Aufgabe des Browsers ist, den Code einer Website unter Berücksichtigung aller Browser-Spezifika und installierter Plugins zu rendern, also in ein visuelles Layout zu übersetzen und dieses dann darzustellen, so muss man doch zumindest festhalten, dass diese spezifische Art der Darstellung – ohne Werbeanzeigen – von den Betreibern der Website so nicht vorgesehen war, wie diese ihrer Leser\*innenschaft auch durch Pop-Ups und entsprechende Artikel in aller Deutlichkeit mitteilen ([Der Spiegel, o. J.](#))<sup>2</sup>.

Ein Screenshot ist also immer nur ein eingefrorener Moment von einem ganz bestimmten Blick- und Zeitpunkt aus, der zwar Neutralität suggerieren mag, dieses Versprechen aber nicht einlösen kann. Jede Website, selbst wenn sie sehr einfach sein sollte, kann abhängig von ihrer digitalen Umgebung und sonstigen Umständen jeweils sehr unterschiedlich aussehen. Wenn sich aber das Aussehen einer Website nicht hundertprozentig festlegen lässt: Was ist dann eigentlich eine Website? Was macht sie aus, woraus setzt sie sich zusammen? Und gibt es dann vielleicht bessere Alternativen zur Archivierung als einen Screenshot?

## **Existiert eine Website auch dann, wenn niemand hinsieht?**

Eine Website ist eine Stelle (englisch *site*: Ort, Stelle, Stätte) im World Wide Web, auf der – meist auf mehreren Unterseiten, den sogenannten Webpages – Inhalte präsentiert werden; man spricht auch von einem Webauftritt. Zu unterscheiden sind die Begriffe *Website* und *Web-*

*seite*, die aus deutschsprachiger Sicht vermeintlich das Gleiche bezeichnen. *Webseite* bezeichnet aber nur eine einzelne Seite des Webauftritts, ist also synonym zum Begriff *Webpage*. *Website* hingegen steht für die Gesamtheit aller Seiten eines Webauftritts, eben die „Stätte“ im World Wide Web, wo diese Inhalte unter einer bestimmten Domain zu finden sind.

Eine Website „besteht“ aus mehreren digitalen Dokumenten, die auf einem Server<sup>2</sup> gespeichert sind. Diese Dateien enthalten zum einen die eigentlichen Inhalte und zum anderen Anweisungen, wie diese Inhalte in einem Internet-Browser dargestellt werden sollen. Die Umwandlung der Anweisungen in die Darstellung erfolgt dabei meist auf dem Rechner der Nutzer\*innen, dem sogenannten *Client*, nachdem die digitalen Dokumente vom Server abgerufen wurden. Wäre die Erklärung hier abgeschlossen, so wäre das Konservieren einer Website nicht sonderlich schwierig: Man müsste lediglich die digitalen Dokumente, die auf dem Server liegen, abspeichern und könnte dann – kompatible Hard- und Software vorausgesetzt – die Website auch in Zukunft im Browser öffnen, selbst wenn die eigentliche Website im World Wide Web irgendwann offline gehen sollte. Man hätte dann eine digitale 1:1-Kopie der Website und könnte diese Kopie dann entweder auf einen anderen Server spielen und von diesem abrufen oder sie sogar lokal auf dem eigenen Rechner ablegen und dort öffnen.

Bei komplexeren Websites werden allerdings meist auch Daten schon auf dem Server vorberechnet, so dass nur noch der letzte Teil – die Überführung von Daten in eine Darstellung – auf dem Client stattfindet. Häufig werden Inhalte zudem aus Datenbanken abgerufen: Der Client fragt dabei ganz bestimmte Daten vom Server ab, dieser filtert die dazu passenden Einträge aus einer Datenbank heraus und sendet sie maßgeschneidert zurück an den Client, wo sie dann in die Inhalte der restlichen Website mit eingefügt werden. Die Datenbanken können dabei auch extern auf anderen Servern gelagert sein, so dass viele verschiedene Websites auf ein und dieselbe Datenbank zugreifen können.

Damit ist nur ein kleiner Teil aller möglichen Konfigurationsoptionen von Websites beschrieben. Es lässt sich allerdings schnell erkennen, dass sich Websites häufig nicht – wie zuvor beschrieben – einfach abspeichern lassen. Viele digitale Dokumente der Website sind vom Client aus gar nicht erst abrufbar: Programmcode, der für die eben beschriebenen Vorkompilierungen zuständig ist, Datenbankeinträge oder Inhalte, deren Abruf ein gewisses Maß an Authentifizierung, also zum Beispiel ein Einloggen, erfordert.

Berücksichtigt man nun dieses komplexe Zusammenspiel von Server, Client, digitalen Dateien und gerenderter Darstellung, so verkompliziert sich die Frage, was denn eine Website ist, erheblich. Der visuelle Part der Website manifestiert sich immer nur in dem Moment, in dem

ein\*e Betrachter\*in die Website aufruft und besteht dann so lange, bis das Browserfenster wieder geschlossen wird. Dennoch besteht die Website natürlich auch zwischen diesen Momenten fort, sie liegt dann auf dem Server und ist einsatzbereit für die nächste Betrachtung.

Neben der Darstellung und der Hard- und Software-Infrastruktur gibt es noch weitere Teile, die man als konstitutiv für eine Website erachten könnte: Viele Seiten haben beispielsweise ein *Backend*, einen Bereich, der nur für User\*innen mit Administrationsrechten zugänglich ist, und wo diese Medien- und Textinhalte einpflegen können, die dann im *Frontend* der Website erscheinen, also in dem Teil, den die Nutzer\*innen zu Gesicht bekommen. Habe ich eine Website vollständig archiviert, wenn ich diesen Teil nicht mitgespeichert habe?

Und letztlich bietet eine Website vor allem auch ein Nutzungserlebnis: In der Nutzung erfüllt sich der eigentliche Zweck der Website. Ein\*e Betrachter\*in besucht die Seite, rezipiert die angebotenen Informationen oder Medieninhalte und tritt über interaktive Funktionalität in ein performatives Wechselspiel mit der Website. Die Bandbreite dieser Nutzungserfahrungen reicht dabei von einer kurzen rein informativen Nutzung bis hin zu höchst immersivem, emotionalem „Erleben“ der Website. Wie kann man aber ein solches Erleben festhalten?

Aus dem Bereich der darstellenden Künste ist das Problem performativer Arbeiten festzuhalten hinlänglich bekannt: Jede Performance ist einzigartig, durch den Ablauf der Aufführung selbst, aber auch durch die Aufführenden, durch das Publikum und die Umgebung. Es ist unmöglich die Performance selbst zu erhalten, man muss sich also anders behelfen: Man kann sie medial dokumentieren, man kann sie per Notation festhalten, man kann sie mündlich, schriftlich oder per Vermittlung tradieren – oder man kann sich entscheiden auf eine Archivierung der Performance ganz zu verzichten und ihr Schicksal somit dem Zufall zu überlassen.

Bei Websites steht uns eine ähnliche Bandbreite an Handlungsmöglichkeiten zur Verfügung. Grundsätzlich muss man jedoch zunächst zwei Grundprinzipien unterscheiden. Das oben besprochene Archivieren der **spiegel.de**-Website ist ein Beispiel, wie man eine Seite per Screenshot dokumentieren kann.<sup>3</sup> Andere Strategien sehen vor eine Website zu *konservieren*, das heißt, sie in einem möglichst authentischen Zustand weiter nutzbar zu halten. Das kann entweder dadurch geschehen, dass man die Website selbst im World Wide Web am Leben erhält oder indem man die Seite vor ihrem „Ableben“ möglichst umfangreich speichert, um sie so später an anderer Stelle wieder betreiben zu können.

Wie *konserviere* ich also eine Website? Grundsätzlich müssen dazu, wie soeben erklärt, vor allem die Code-Bestandteile sowie Medien- und Textinhalte – soweit man ihrer habhaft werden kann – abgespeichert werden. Man könnte versuchen das manuell zu machen, indem

man beispielsweise integrierte Browser-Funktionen nutzt, um die einzelnen Dateien besuchter Websites abzuspeichern. Da heutzutage aber selbst einfache Websites oft aus sehr vielen Einzeldateien bestehen, ist diese Aufgabe schnell überfordernd. Alternativ könnte man sich eigene Scripts<sup>4</sup> schreiben, die diese Aufgabe automatisieren und damit nicht nur den Aufwand, sondern auch die Fehleranfälligkeit des Unterfangens erheblich reduzieren. Das Schreiben solcher Scripts erfordert allerdings ein Mindestmaß an Programmierkenntnissen und ist somit für den größten Teil der Nutzer\*innen keine realistische Option. Als dritte und komfortabelste Möglichkeit kann man auf bestimmte Webplattformen zurückgreifen, die Werkzeuge genau für die Aufgabe Webseiten abzuspeichern bereitstellen.

Die vielleicht bekannteste dieser Plattformen ist das **Internet Archive**, ein gemeinnütziges Projekt mit Sitz in San Francisco. Es versteht sich als umfassende digitale Bibliothek und hat sich zur Aufgabe gemacht, das Internet in größtmöglichem Umfang zu archivieren ([vgl. Internet Archive, o. J.](#))<sup>5</sup>. Bis heute (Stand: März 2021) hat die Plattform dabei 475 Milliarden Webseiten archiviert. Dreh- und Angelpunkt der Plattform ist die sogenannte **Wayback Machine**, die als Tool zur Speicherung einer Webseite dient, und gleichzeitig ein Interface bietet, um bereits gespeicherte Seiten aufzusuchen. Die Idee dabei ist, dass die Wayback Machine „Snapshots“ mit Zeitstempel anlegt, die eine bestimmte Webseite zu einem bestimmten Zeitpunkt archivieren. Diese Snapshots sind keine Screenshots, sondern eine gespeicherte Version der Webseite selbst. Als Referenzpunkte dieser Speicherung dienen einerseits die URL, also die Internetadresse, und andererseits der Zeitpunkt der Archivierung. Bei vielen in der Wayback Machine gespeicherten Seiten ist es möglich, zu verschiedenen vergangenen Zeitpunkten zurückzuspringen und so im Vergleich dieser Momentaufnahmen die Entwicklung der Webseite punktuell nachzuverfolgen ([vgl. Dekker, 2020, S. 50](#)).

Jeden Tag werden dem Archiv 750 Millionen Webseiten hinzugefügt ([Jessen, 2020](#)). Das System speist sich dabei aus zwei Hauptquellen: Zum einen sind permanent hunderte sogenannte Webcrawler<sup>6</sup> automatisiert im World Wide Web unterwegs und „wandern“ dabei nach bestimmten vorgegebenen Regeln von Seite zu Seite, wobei jede so besuchte Seite archiviert wird. Die Crawler werden teils vom Internet Archive selbst betrieben, teils aber auch von Drittanbietern, die die so gesammelten Daten dann an das Internet Archive spenden. Zum anderen können Webseiten von Freiwilligen zum Internet Archive hinzugefügt werden. Wann immer man das Bedürfnis hat, eine Webseite für die Nachwelt zu erhalten, kann man ihre URL in ein [Webformular](#) eingeben und sie wird dann vollautomatisch mit zahlreichen Metadaten im Archiv abgespeichert. Alternativ kann man auch eine Browser-Erweiterung installieren, die diesen Prozess weiter vereinfacht: Man kann so in seiner alltäglichen Webnutzung mittels eines einzigen Knopfdrucks die gerade besuchte Webseite speichern, eine Vorge-

hensweise, zu der das Internet Archive auf einer seiner Anleitungsseiten mit der Formel „If You See Something, Save Something“ ermuntert ([Rossi, 2017](#))<sup>15</sup>. Die gleiche Erweiterung erlaubt einem auch das schnelle Auffinden der archivierten Versionen einer besuchten – und eventuell nicht mehr – auffindbaren Website (→ Abb. 5).



Abb. 5: Übersicht der mit der Wayback Machine archivierten Versionen von spiegel.de im Jahre 2014 (Screenshot, 19.04.2021)

Das Internet Archive versucht Webseiten so zu speichern, dass sie nach Möglichkeit so nutzbar sind, wie dies zum Zeitpunkt der Speicherung der Fall war. Das gelingt mal mehr, mal weniger gut. Es ist nicht unüblich, auf gespeicherte Seiten zu stoßen, deren Layout offensichtlich zerstört ist oder bei denen beispielsweise Medieninhalte fehlen. Dynamisch geladene Inhalte, wie sie etwa nach Aufruf der Website von Nutzer\*innen über interaktive Funktionen anfragt werden, sind im Regelfall auch nicht mehr abrufbar, das heißt, wenn ich auf der archivierten Website versuche die gleiche Funktionalität zu nutzen, passiert oft nichts. Bezeichnend ist aber vor allem, dass häufig eines der wesentlichsten Grundelemente des World Wide Web nicht mehr funktioniert: der Hyperlink. Links, die auf externe Websites oder aber auf andere Webpages der archivierten Website verweisen, führen häufig ins Leere, da die verlinkten Seiten nicht ebenfalls mitarchiviert wurden. Dieses Phänomen ist so weit verbreitet, dass sich ein eigener Begriff dafür etabliert hat: *link rot* (englisch für „Linkverrottung“). Wurden die verlinkten Seiten doch mitarchiviert, so wurden die entsprechenden Snapshots häufig an einem späteren Zeitpunkt angefertigt (da nicht alles exakt gleichzeitig abgespeichert werden kann) oder der Hyperlink führt sogar zur aktuellen Version der verlinkten Seite. Das Ergebnis entspricht also oft kaum der Webseite, wie sie in der Vergangenheit

ausgesehen haben mag, sondern ist ein nur in Teilen funktionsfähiges Archivstück und enthält zeitliche oder technische Inkonsistenzen (vgl. Dekker, 2020, S. 50).

Über all diesen beschriebenen Schwierigkeiten steht für mich aber ein Problem, dass meinem Eindruck nach in Besprechungen archivierter Seiten kaum Erwähnung findet. Verdeutlichen lässt es sich am besten an einem Beispiel. Über eine Websuche kann man relativ unproblematisch die älteste im Internet Archive gespeicherte Website herausfinden: ein Snapshot der Seite **infoseek.com** vom 12. Mai 1996. Infoseek wurde 1994 gegründet und war in den 1990er-Jahren eine vielgenutzte Suchmaschine ([„Infoseek“, 2020](#)). Als eine der frühen Suchmaschinen in einem Prä-Google-Zeitalter hat sie damit eine beträchtliche historische Bedeutung. Besucht man heutzutage [die archivierte Version dieser Website](#), so sieht man in etwa Folgendes:

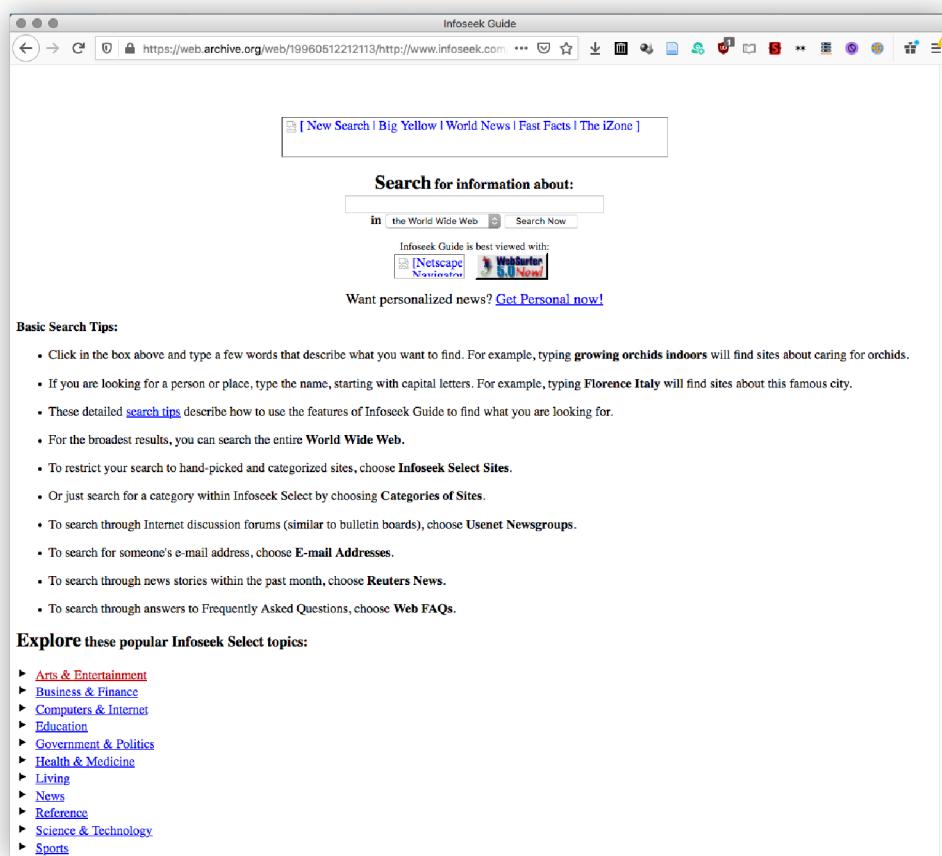


Abb. 6: Im Internet Archive archivierte Version von infoseek.com, ursprünglich archiviert am 12.05.1996 (Screenshot, 03.03.2021)

Neben den offenkundig nicht mehr verlinkten Bildern ist mein Problem vor allem dieses: Ich sehe zwar eine Webseite, ich habe aber keinerlei Bezugsrahmen, ob diese Seite auch wirklich so ausgesehen hat. Sah das Layout 1996 tatsächlich so aus? Oder wird es mir durch den vielleicht unzulänglichen Archivierungsprozess jetzt verfälscht angezeigt? Ich könnte mich daranmachen, den Code der Seite zu analysieren, um so herauszufinden, wie die Webseite eigentlich dargestellt werden soll. Aber selbst wenn ich soweit käme, blieben viele Ungewissheiten: Stellt mein moderner Browser die Seite nicht vielleicht ganz anders dar, als es ein zeitgemäßer Browser 1996 getan hat? Browser werden stetig

weiterentwickelt und so ändern sich auch die integrierten Regeln, wie bestimmte HTML-Elemente dargestellt werden sollen. Ebenfalls kann ich nicht wissen, ob die digitalen Dateien der Website beim Abspeichern in irgendeiner Form angepasst wurden, um sie mit dem Archivierungssystem kompatibel zu machen.

Kurz: Ich kann die archivierte Seite so hinnehmen, wie ich sie vorfinde, ich habe aber keinerlei Gewissheit, wie sie eigentlich ausgesehen hat. Bei einem solch alten Beispiel mag dieses Problem sehr offensichtlich sein. Aber auch bei allen neueren archivierten Webseiten ergeben sich die gleichen Probleme: Es dürften zwar in der Regel deutlich mehr Layoutstrukturen sichtbar sein, aber auch hier kann keine letztliche Gewissheit herrschen, dass die Webseiten originalgetreu dargestellt werden. Mehr noch: Dadurch, dass diese potenziellen Layoutabweichungen nicht so offensichtlich sind wie im obigen Beispiel, wird einem ein vermeintlich authentischer Rückblick auf eine frühere Version suggeriert, ein Umstand, über den man leicht hinwegsehen könnte.

Bei allen Optionen, die einem erhalten bleiben – der interaktiven Nutzung, der Möglichkeit „[to] study single pages in a website over time“ (Dekker, 2020, S. 50), dem Schaffen bleibender Referenzpunkte zum späteren Wiederauffinden – vermag das Internet Archive eine wichtige Funktion also nicht zu erfüllen: zu dokumentieren, wie die Seite genau ausgesehen hat. Das vermag hingegen ein einfacher Screenshot.

## If a long-dead site is restored, does it still count?

Diese schöne Frage eröffnet einen Artikel von **Paul Koerbin**, Web Archiving Manager der National Library of Australia ([Koerbin, 2013](#)). Er beschreibt darin ein Projekt im Jahre 2013, bei dem – 20 Jahre nach ihrer Veröffentlichung im World Wide Web – die weltweit erste Website wieder im Originalzustand hergestellt werden sollte. Dazu wurden alle Dateien der ursprünglichen Seite zusammengetragen und sogar dafür gesorgt, dass der Name des Servers, die IP-Adresse<sup>7</sup> und die URL mit denjenigen der Original-Website übereinstimmen. Die Webseite ist seither wieder unter ihrer ursprünglichen URL [info.cern.ch/hypertext/WWW/TheProject.html](http://info.cern.ch/hypertext/WWW/TheProject.html) erreichbar, nachdem sie vor ihrer Restauration viele Jahre offline war ([Noyes, 2013](#)).

Die saloppe Frage „Does it still count?“ – „Gilt das?“ – mag als wissenschaftliches Kriterium unnütz erscheinen, bringt die Fragen der Authentizität und des Originals, mit der das Thema Webarchivierung einhergeht, allerdings treffend auf den Punkt. Koerbin führt dazu aus:

„[...] is this resurrection of the original files, machine names and IP addresses really *the actual* first website? What does the re-creation of a virtual artefact mean? There is evidence to suggest this site disappeared, was indeed lost, years ago. [...]“

So, perhaps the site just went missing for a while, the best part of a couple of decades [...]. When it returns however, is it the original site or a replica, or a recreation or a new site – well I guess, arguably, all the above in parts.“

(*Koerbin, 2013*) 

Welche Voraussetzungen müssten gegeben sein, um noch von der Original-Website sprechen zu können? Muss die Website durchgehend online gewesen sein? Reicht es, wenn Kopien der damaligen Dateien jetzt wieder auf Servern liegen und die Website damit wieder läuft? Ist die Frage nach Authentizität schon beim Archivieren nicht-digitaler Objekte schwierig genug, so gewinnt sie im digitalen Raum noch erheblich an Komplexität: Mit dem Aufkommen digitaler Dokumente und damit der digitalen Kopien und der Möglichkeit, diese Dokumente immer weiter zu ändern, ist es schlicht unmöglich geworden, an der Idee des Originals in ihrer ursprünglichen Form festzuhalten.

In Bezug auf den berühmt gewordenen Aufsatz „Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit“ des Philosophen **Walter Benjamin** und seine darin beschriebene Idee der „Aura“, die einem Original innewohne und die in technischen Reproduktionen (z. B. Fotografien) dieses Originals nicht transportiert werden könne (vgl. Benjamin, 1977, S. 15ff.), bemerkt der Web-Künstler **Rafael Lozano-Hemmer**:

„Most [art collectors] do not understand that an original file is identical to a copy. And if they do, they are so completely absorbed with the aura of authenticity that I have heard of artists having to destroy a digital file once they print copies of a digital picture. This is absolutely absurd and unnecessary [...].“

(*Grau/Hoth/Wandl-Vogt (Hrsg.), 2019, S. 113*)

Dem Gedanken, dass es absurd und unnötig sei, digitale Kopien zu zerstören kann ich mich anschließen. Die Frage, ob eine Kopie tatsächlich identisch mit dem Original ist, verdient allerdings eine nähere Betrachtung, die ich im Abschnitt [Networks of Care](#) vornehme.

Digitale Dokumente sind nicht nur frei kopierbar, sondern auch immer weiter modifizierbar und daher potenziell nie in einem fertigen Zustand, sie befinden sich in einem permanenten *Prozess*. Es gibt inzwischen zahlreiche Websites, die seit mehr als zwei Jahrzehnten dauerhaft online sind. Im Regelfall unterliegt eine Website innerhalb dieser Lebensdauer zahlreichen Änderungen: Die Seite wird visuell, strukturell und inhaltlich an die jeweils aktuellen Bedürfnisse angepasst. Was ist bei einer so lang bestehenden Seite die Original-Website? Die Seite im „Ursprungszustand“, so wie sie einst vor vielen Jahren online ging? Oder die aktuell online verfügbare Seite, die eventuell auf dem

Ursprungszustand basiert und von diesem über die Jahre Schritt für Schritt weiterentwickelt wurde? Es wird schnell offensichtlich, dass es *das eine Original* im Digitalen nicht geben kann.

Die Theoretikerin und Kuratorin **Annet Dekker** schlägt daher vor, dass man ein digitales Werk in *Versionen* unterteilt.<sup>8</sup> Eine Version ist ihr zu folge eine Variante eines Werks, die einerseits von der vorherigen Version abstammt und mit dieser verbunden ist und daher nicht als eigenständiges, neues Werk gesehen werden kann, und die andererseits genug Unterschiede zur vorherigen Version aufweist, um sich von dieser wahrnehmbar zu unterscheiden.<sup>9</sup> (vgl. Dekker, 2020, S. 30f.) Da die unterschiedlichen Versionen einer Arbeit also Weiterentwicklungen einer vorherigen Version sind, sind sie laut Dekker eben nicht nur Kopien oder Reproduktionen dieser Arbeit, sondern haben als vollwertige neue Version jeweils eine eigene „Daseinsberechtigung“, weshalb sie durchaus auch eine authentische Qualität besitzen (ebd., S. 73).

Die Möglichkeit einer Vielzahl an Versionen ist also integraler Bestandteil von Websites und muss entsprechend bei der Archivierung berücksichtigt werden. Dies wird schließlich zur kuratorischen Aufgabe: „[...] a conservator must choose which version(s) to save [...].“ (Dekker, 2020, S. 87). Weil man schlicht nicht alles speichern kann, muss ausgewählt werden, welche Zustände, welche Versionen speichernswert sind. Da dies nicht immer offensichtlich ist, kommt dem\*der Konservator\*in also in vielen Fällen die Aufgabe zu, überhaupt erst zu entscheiden, was bei einer bestimmten Website eine neue Version ist und was nicht. Wie diese Kuration erfolgen kann, wird in [Teil 3: Archive. Fünf Fallbeispiele](#), anhand der Betrachtung einiger Beispiele genauer besprochen.

Aber selbst wenn man die Prozesshaftigkeit und Wandelbarkeit als Eigenschaft von Websites einmal ausblendet, so bleibt das Authentische weiterhin ein problematischer Begriff. Hierzu bietet es sich an Websites zu untersuchen, die schon seit langer Zeit in unveränderter Form online sind. Ein bekanntes Beispiel einer solchen Seite ist die Website des US-amerikanischen Spielfilms **Space Jam** aus dem Jahre 1996. Sie hat nicht zuletzt deshalb einen so hohen Bekanntheitsgrad erlangt, weil sie als eines der letzten Relikte aus einer längst vergangenen Internet-Ära gilt und als solches in zahlreichen Online-Artikeln Erwähnung findet (eine lesenswerte Erzählung ihrer Entstehungsgeschichte findet sich z. B. unter: [‘Space Jam’ Forever: The Website That Wouldn’t Die](#)). Die Seite geriet nach dem Kinostart des Films zunächst in Vergessenheit, blieb offenkundig aber dennoch online und wurde schließlich 2010 in den sozialen Netzwerken wiederentdeckt ([Malinowski, 2015](#)).<sup>10</sup> Nur 14 Jahre nach ihrer Entstehung war sie visuell und technisch dermaßen aus der Zeit gefallen, dass sie schlagartig zur Internetsensation avancierte. Wenn man sie heute unter [spacejam.com](#) besucht, erkennt man die Ästhetik früher Internet-Tage auf den ersten Blick (→ Abb. 7).



Abb. 7: spacejam.com Homepage (Screenshot, 07.03.2021)

Obwohl sich die Seite also – was ihren Inhalt und ihr Aussehen betrifft – noch im Originalzustand befindet, ist es unmöglich, beim Besuch der Website von einem authentischen Erlebnis zu sprechen. Das liegt daran, dass sich Kontext und Umfeld der Nutzung nun komplett verändert haben: In den 1990er-Jahren haben die Nutzer\*innen die Website auf den damals üblichen Röhrenmonitoren mit extrem niedriger Auflösung betrachtet, der Seitenaufruf erfolgte erheblich langsamer, als das heute der Fall ist. Der Browser, in dem die Seite damals aufgerufen wurde, war ein ganz anderer und gab der Website auch visuell einen völlig anderen Rahmen. Wenn ich eine solche alte Website also heute auf einem modernen Retina-Flachbildschirm aufrufe und diese – dank schneller Internetverbindung – augenblicklich in meinem modernen Browser oder gar auf meinem Smartphone dargestellt wird, dann entspricht das in keiner Weise dem Nutzungserlebnis, das man im Jahre 1996 hatte. Auch die interaktiven Spielereien, die die Seite bereithält, boten den damaligen Nutzer\*innen ein vermutlich deutlich neuartigeres Nutzungserlebnis, als es die Website heute bieten kann, wo wir andernorts täglich mit deutlich komplexeren und ausgereifteren Web-Apps zu tun haben.

Zudem hat sich die inhaltliche Bedeutung der Website völlig verschoben: Hätte ich mir damals die Website angesehen, um mich über den Film zu informieren oder um mit diesen damals innovativen interaktiven Features zu spielen, so schaue ich sie mir heute vor allem unter internet-historischen Gesichtspunkten an. Ich wurde vermutlich von Blogartikeln, Social-Media-Posts oder Theorietexten auf die Website geleitet, im Jahre 1996 sind Besucher\*innen möglicherweise noch hauptsächlich über das Abtippen der URL von Promo-Material auf der Seite gelandet. Auch ist die Website durch ihre Sonderstellung in Bezug auf die Geschichte des Internets heute natürlich viel bedeutender, als sie es früher je war.<sup>10</sup>

Kurz: Auch wenn sich die Seite selbst inhaltlich und visuell nicht geändert hat, so hat sich doch der Kontext ihrer Nutzung heute grundlegend gewandelt. Wenn man also eine Website nicht nur als die Summe der ihr zugrundeliegenden Dateien und ihrer elektronischen Infrastruktur versteht, sondern auch Konzept und Nutzung zu den konstitutiven Bestandteilen einer Website zählt, so muss man zu dem Schluss kommen, dass sich auch die Website selbst geändert hat. Auch Dekker spricht davon, dass eine neue Version einer Website nicht notwendigerweise durch Änderungen des Codes entstehen muss, sondern gegebenenfalls auch – unter Umständen völlig ohne Zutun ihrer Urheber\*innen – durch eine Verschiebung des Kontexts entsteht: „Or versions are culturally induced, or incidental, and are therefore not necessarily linked to authorization“ (Dekker, 2020, S. 30). Man könnte also im Falle der Space-Jam-Website von mindestens zwei Versionen sprechen: Der ursprünglichen Website in den 1990ern und der historisch aufgeladenen, wiederentdeckten Website nach 2010.

Die Untersuchung zeigt: Ob man eine Website als verschiedene authentische Versionen der gleichen Website sieht oder als authentliches oder nicht-authentliches Original oder als authentische oder nicht-authentische Kopie oder aber als völlig neue Website, hängt stark davon ab, in welchem Kontext man diese Einteilung vornimmt und auf welchen Aspekt man Wert legt: Ein auf das Nutzungserlebnis gerichteter Blick mag zu einem anderen Ergebnis kommen als ein Blick, der sich nur auf den Code einer Website konzentriert. Die Ausgangs-Frage „Does it count?“ bleibt daher in Bezug auf Archivierung legitim und ist zugleich ein guter Gradmesser: „Zählt“ es, wenn ich eine Website auf diese oder jene Art und Weise archiviere? Eine allgemeingültige Antwort und damit Archivierungspraxis kann es nicht geben, es hängt letztendlich davon ab, welcher Aspekt der Website mir bei der Archivierung wichtig erscheint, und ob ich diesen Aspekt dann bei meiner archivierten Version in angemessener Weise berücksichtigt sehe.

## 2. Digitales Erinnern und Vergessen

Bei der Diskussion über das Archivieren von Websites und all den damit verbundenen Schwierigkeiten rückt irgendwann unweigerlich eine Frage in den Mittelpunkt: Wozu das alles? Warum sammeln wir überhaupt Websites oder andere digitale Artefakte? Könnten wir nicht einfach gänzlich darauf verzichten und die digitale Alltagskultur dem stetigen Wandel und damit dem Lauf der Dinge überlassen?

So unterschiedlich die Zielsetzungen hinter einzelnen Sammel- und Archivierungsbemühungen auch sein mögen, so steht hinter allen ein gemeinsames Grundmotiv: das Erinnern. Dinge – digitale wie analoge – werden zusammengetragen und bewahrt, damit sie nicht in Vergessenheit geraten. Hätten wir ein perfektes, fotografisches Gedächtnis,

so müssten wir auch nichts archivieren, sondern könnten jederzeit selbst unser Wissen wieder abrufen.

Die Kulturwissenschaftlerin **Aleida Assmann** hat viel zum Erinnern und Vergessen geforscht. Sie beschreibt das Vergessen als „Grundmodus menschlichen und gesellschaftlichen Lebens“ (Assmann, 2016, S. 30). Nur wenn man diesem Grundmodus die Anstrengung des Erinnerns entgegenseetzt, kann man Dinge vor dem Vergessen bewahren. Wir schaffen so auf persönlicher Ebene ein *individuelles* Gedächtnis, im größeren Maßstab auf gesellschaftlicher Ebene aber auch ein *kollektives* bzw. *kulturelles* Gedächtnis. Erinnern geschieht, indem wir einkommende Informationen mit unserer Perspektive und unseren Wertvorstellungen abgleichen und so entscheiden, was relevant und was irrelevant ist und dadurch (meist unbewusst) eine Auswahl treffen, welche Informationen wir im Gedächtnis behalten und welche wir vergessen. Mit dem Produzieren von Erinnerungen versichern wir uns so unserer eigenen Identität im individuellen wie auch im gesellschaftlichen Rahmen.

Mit Erfindung der Schrift wurde erstmals eine Speichertechnik geschaffen, die das Tradieren von Texten und kulturellen Praktiken stabilisiert (ebd., S. 220), so dass dieses Behalten und Weitergeben nicht mehr nur auf das menschliche Gedächtnis beschränkt war. Dieses erste Speichermedium hat sich weiterentwickelt, bis hin zu den heutigen digitalen Medien, „als eine Schrift, die nicht mehr nur Sprache codiert, sondern auch Bilder und Töne aufzeichnen [kann]“ ([Schülein, 2021](#))<sup>2</sup>.

Unser kulturelles Gedächtnis „beruht [...] auf Medien, die dieses Gedächtnis formen und definieren“ (Assmann, 2016, S. 212). Mit dem Aufkommen der digitalen Medien ist auch erstmals ein Kapazitätsproblem der Vergangenheit überwunden: Zuvor waren Informationen aus Platz- und Logistikgründen nicht unbegrenzt speicherbar, verwaltbar und abrufbar und mussten daher regelmäßig von „Altläden“ befreit werden. Informationen, die nicht (mehr) in den Wertekanon einer Gesellschaft passten, waren es nicht wert gespeichert zu werden, verblassten so und gerieten schließlich in Vergessenheit. Nun allerdings, da Speicherplatz praktisch unbegrenzt verfügbar ist, hat sich ein grundlegender Wandel der Medien und damit der Möglichkeiten zur Formung des kulturellen Gedächtnisses vollzogen.

Assmann bezeichnet diese Wende als „Post-Verknappungs-Kultur“, in der „immer mehr Information für immer mehr Nutzer ohne besonderen Aufwand erreichbar, speicherbar, übertragbar“ (ebd., S. 213) ist. Im Widerspruch zu anderen Theoretiker\*innen verweist sie jedoch auch darauf, dass damit nicht Medien und Gedächtnis gleichzusetzen sind (vgl. ebd.): Zur Formung des Gedächtnisses besteht weiterhin „die Notwendigkeit der Auswahl, die wiederum durch die Verbindung des Gedächtnisses mit einer Identität und einer Perspektive gegeben ist“

(ebd., S. 217). Der Unterschied zu vorher besteht darin, dass wir nun auf einen viel größeren Pool an Informationen zurückgreifen können, um (meist unbewusst) diese persönliche Auswahl zu treffen, die uns im Gedächtnis bleibt und damit unsere Identität bestimmt.

Dieser riesige Pool an Daten, dem wir gegenüberstehen, ist allerdings nicht mehr direkt von uns lesbar, sondern nur noch von Maschinen. Die Vermittlung zwischen den gespeicherten Daten und dem Menschen findet daher über Maschinen und deren Algorithmen statt ([vgl. Schülein, 2021](#)). Assmann unterscheidet dabei zwei Formen der *Aufmerksamkeit*:

„Die Aufmerksamkeit der Maschine funktioniert quantitativ. Sie hierarchisiert Informationen statistisch nach der Menge von Klicks, sie strukturiert somit auch die menschliche Aufmerksamkeit vor und entscheidet letztlich darüber, *was gefunden werden kann und was nicht*. Die menschliche Aufmerksamkeit funktioniert dagegen qualitativ. Aus ihr entspringt der Suchbegriff, der eingegeben wird, und der Pfad, den man sich durch das Gewebe der Links bahnt.“

(Assmann, 2016, S. 216, Hervorhebung T.R.)

Diese Vorstrukturierung verfügbarer Informationen durch Algorithmen definiert und limitiert unseren Handlungsspielraum in der Arbeit mit großen Archiven und ist daher eines ihrer zentralen Probleme, auf das ich in [Teil 4: Archive und Wirklichkeit](#) noch näher eingehen werde.

In Bezug auf die Dynamiken rund um die Bildung des kulturellen Gedächtnisses spricht Assmann von zwei Polen, dem *Kanon* und dem *Archiv*:

„Man kann diesen Zusammenhang anschaulich im Museum erkennen: Dort gibt es kuratierte Ausstellungen, die öffentliche Aufmerksamkeit bekommen. Unten im Keller lagern dagegen Artefakte, die womöglich nie gezeigt werden. [...] Es gibt den Kanon als enge Auswahl der derzeit als bedeutend angesehenen Werke und Ereignisse. Und es gibt das sehr viel umfangreichere Archiv, in dem alle Dokumente schlummern, die in der Gegenwart nicht erinnert werden.“

([Schülein, 2021](#))

Bei beiden Ausprägungen – Ausstellung und Archiv – handelt es sich also um Sammlungen, wobei der kanonischen Sammlung ein Abgleich mit den kulturell dominanten Werten und damit ein Kurationsschritt vorausgeht, wohingegen die Archive oft mehr eine „Ansammlung“ an Dingen sind, die das Potenzial haben irgendwann einmal wieder relevant zu werden und dann wieder in den Kanon aufgenommen zu werden, beispielsweise, wenn sich das Wertegerüst einer Gesellschaft

verschiebt. Assmann spricht beim Sammeln bzw. beim Archiv von einer „bewahrenden Vorhaltung“ (Assmann, 2016, S. 20) oder in Bezug auf das Vergessen von einem „Verwahrensvergessen“ (ebd., S. 38). Die so ins Vergessen geratenen Informationen sind nicht vollständig vergessen, sondern können wieder reaktiviert werden, in dem Moment, wo sie relevant werden. Dabei kann es zum Beispiel zu einer Wiederentdeckung bestimmter gesellschaftlicher Akteur\*innen oder Diskurse kommen, die zuvor keine Rolle (mehr) gespielt haben.

„Bibliotheken und Archive sichern die Quellenbasis der historischen Geisteswissenschaften und bilden einen wichtigen Fundus für Arbeiten von Schriftstellern, Künstlern und Filmemachern. Damit erfüllen sie eine ganz wichtige Aufgabe, denn sie bilden die Voraussetzung für das, was in Zukunft noch über eine Gegenwart gesagt werden kann, wenn diese zur Vergangenheit geworden sein wird.“

(Assmann, 2016, S. 38)

Das Internet selbst könnte allerdings auch als ein riesiges Archiv verstanden werden. Warum müssen Websites also überhaupt gesammelt werden? Man könnte doch in dem Moment, wo man sie benötigt, einfach danach suchen und sie dann aufrufen? Wie die bisherige Diskussion aber gezeigt hat, taugt das Internet selbst als beständige Sammlung nur bedingt. Laut Assmann hat dieses „digitale Mega-Archiv die Eigenschaft, von allen Seiten blitzschnell und mühelos alles Mögliche aufzunehmen, aber nichts zu vergessen – und auch nichts verlässlich zu behalten“ (ebd., S. 203f.). Dass nichts verlässlich behalten wird, haben wir bereits gesehen. In Bezug darauf, dass in diesem Mega-Archiv nichts vergessen wird, möchte ich Assmann aber entschieden widersprechen: Es mag zwar stimmen, dass sich immer größere Mengen an Daten ansammeln. Sind die Daten aber nach einiger Zeit nicht mehr relevant und werden daher nicht mehr aufgesucht, sind sie in der Folge irgendwann auch nicht mehr auffindbar: Die entsprechenden Links landen im Nichts oder die Websites sind durch neuere Versionen der Seite ersetzt worden. Wenn die Social-Media-Community Kopf steht, weil plötzlich eine verloren geglaubte 14 Jahre alte Website wieder auftaucht und diese fortan als eine der wenigen noch bestehenden Seiten aus einer eigentlich noch jungen Internetvergangenheit gilt, dann ist das ein deutlicher Beleg dafür, dass das Vergessen auch im Internet der Grundmodus ist – zumindest solange nichts unternommen wird, um diesen Zerfallsprozess zu stoppen.

Das Internet Archive ist ein gutes Beispiel einer solchen Gegenmaßnahme: Es ermöglicht (und praktiziert) das Archivieren im größtmöglichen Maßstab und bildet somit einen großen Pool an Websites, die bei Bedarf irgendwann wieder reaktiviert und damit in unseren kulturellen Kanon aufgenommen werden können. So erklärt sich auch das scheinbar exzessive Sammeln solch ungeheurer Datenmengen inner-

halb dieses Projekts: Man kann vorab nie wissen, welche der dorthin verschwundenen Seiten plötzlich dringend wieder benötigt werden, also speichert man besser gleich so viel wie möglich.

## Screenshots als Erinnerungsträger

In Bezug auf das Erinnern verdient auch der Screenshot als Archivierungswerkzeug eine genauere Betrachtung. Die Medienwissenschaftlerin **Annekathrin Kohout** sieht eine zunehmende Tendenz Screenshots aufzunehmen und stellt diese Tätigkeit in die Tradition der Alltagsfotografie ([vgl. Kohout, 2020](#)). Nachdem das Erstellen von Screenshots ursprünglich hauptsächlich die Funktion eines Dokumentierens von Software-Fehlern hatte, wurde es mit zunehmender Verlagerung unseres Alltags in die digitale Welt mehr und mehr als ein Werkzeug begriffen, um die Bildschirmwirklichkeit festzuhalten:

„Mittlerweile wird über [das] Computer-Display mehr denn je eine Welt gesehen, die vergleichbar unbeständig, vergänglich, live ist wie jene, die uns im realen Raum umgibt. Immer mehr Einträge haben ein Ablaufdatum – besonders Storys auf Instagram, Facebook und Snapchat. Vieles wird live übertragen und ist insofern nicht weniger momentlich, augenblicklich, wie eine Situation außerhalb des Bildschirms. In dieser Bildschirm-Welt übernimmt der Screenshot deshalb auch die wichtigste Funktion, die einst der Fotografie zukam: Erinnerung.“

([Kohout, 2020](#))

Auf Smartphones und Rechnern werden also ständig digitale „Alltags-situationen“ festgehalten, um sich später an diese erinnern zu können. Wie die so gesammelten Bilder zu Erinnerungen werden können und anschließend dann zur Prägung unserer Identität beitragen, zeigt sich gut im Vergleich mit der Alltagsfotografie per Smartphone: Mit dem Smartphone produziere ich permanent Bilddaten und häufe schließlich eine riesige Ansammlung an Fotos an. Erst wenn ich beginne, diese Sammlung zu sichten und die Fotos zu löschen, die ich nicht mehr brauche, entdecke ich sie zu dem ein, was ich für die Zukunft bewahren möchte und was deshalb meine Erinnerung an die Vergangenheit tragen wird und somit meine Identität prägt. Wir alle haben – selbst bei einem ursprünglich viel größeren Pool an Bildern, der uns bereitstand – eine kleine Auswahl ganz bestimmter Bilder, an die wir uns stets erinnern können und die quasi als „kanonische Sammlung“ die Erinnerung an unser Leben repräsentieren. „Sammeln ist Suchen, Festhalten, Bewahren und stiftet damit zugleich ein Fundament individueller Identität: „Meine Sammlung – das bin ich““ (Assmann in Junge/Suber/Gerber (Hrsg.), 2008, S. 346).

Die Sammlung an Screenshots dokumentiert dabei ebenso mein Leben auf dem Bildschirm, wie die Sammlung an Smartphone-Fotos das

Leben außerhalb des Internets dokumentiert. Kohout bemerkt dazu: „Screenshots zu machen und ihnen den Status von Fotografien zu verleihen, bedeutet auch, das Internet als Alltagswirklichkeit und Lebensrealität anzuerkennen.“ Ich würde die Aussage eher umstellen: Weil das Internet längst zur Lebensrealität geworden ist, machen wir innerhalb dieser Wirklichkeit ganz selbstverständlich auch Fotos – und das sind dem Medium entsprechend eben Screenshots.

Die Screenshot-Funktionalität elektronischer Endgeräte ist das wohl simpelste Werkzeug, um digitale Erinnerungen festzuhalten: Sie ist unkompliziert, schnell nutzbar und vor allem immer verfügbar, wenn man vor dem Bildschirm sitzt – man kommt nie in die Verlegenheit seine Kamera vergessen zu haben. Darüber hinaus haben Screenshots einen großen Grad an „Wirklichkeitstreue“ ([Kohout, 2020](#)). Im Gegensatz zur Fotografie wird hier nicht im Moment der Bilderzeugung ein räumliches Motiv auf eine Fläche reduziert, sondern es wird ein bereits zweidimensionales Motiv pixelgenau in der Weise festgehalten, wie wir es schon auf dem Bildschirm sehen ([vgl. ebd.](#)). Aufgrund der niedrigschwelligen Anwendung gibt es zahlreiche Nutzer\*innen, die im digitalen Alltag eine beträchtliche Menge an Screenshots produzieren, ohne dabei notwendigerweise bewusst ein Archiv anlegen zu wollen. Genau in dieser Nutzung erfüllt sich der Ausspruch „We are all archivists“ (Dekker, 2020) für die Bildschirmwirklichkeit, wie er sich parallel für die Offline-Wirklichkeit mit der Smartphone-Fotografie erfüllt hat.

Aber natürlich kann die Nutzung der Screenshot-Funktion auch über diese unbeabsichtigte Alltags-Archivierung hinausgehen und gezielter eingesetzt werden. In meiner Arbeit als Gestalter nutze ich Screenshots häufig zum schnellen Speichern visueller Referenzen. Stoße ich im Internet-Alltag auf Inhalte, die visuell, konzeptionell oder sonstwie von Interesse sind, so halte ich sie per Screenshot fest. Dabei ist das schnelle Erstellen der Screenshots meist wichtiger als das sorgfältige Inszenieren der Aufnahme, denn oft dienen sie nur als gedankliche Stütze, um später wieder zu einem bestimmten Inhalt zurückzufinden. So können sie dann zukünftig bei der Recherche oder Konzeption eines neuen Projekts Verwendung finden. Der so entstandene und stetig anwachsende Screenshot-Ordner funktioniert verlässlich als Erinnerungswerkzeug: Bei jedem neuen Durchstöbern fördert er etwas bereits Vergessenes zutage. ([→ Abb. 8](#))



Abb. 8: Auswahl einiger Screenshots aus meinem digitalen Alltag (2019-2021) (23/23)

Obwohl der Aspekt des Erinnerns am Beispiel der Screenshots wohl am besten nachzuvollziehen ist, dienen auch alle anderen Archivierungstechniken dazu, etwas vor dem Vergessen zu bewahren. Die daraus entstehenden Archive sind somit alle mit dem Bilden einer Identität und des (kulturellen) Gedächtnisses verbunden. Es gilt, aus dem großen vergänglichen Mega-Archiv des Internets Untersammlungen anzulegen, Archive also, die Websites und damit die Digitalkultur von heute – und somit der zukünftigen Vergangenheit – sichtbar halten. Trotz dieses gemeinsamen Grundgedankens verfolgen sie vordergründig ganz unterschiedliche Zielsetzungen. Im Ergebnis hat man Archive, die sich sehr unterschiedlich positionieren und jeweils andere Strategien finden, die gesammelten Bestände zu organisieren und zu präsentieren.

Im folgenden Abschnitt werde ich diese Unterschiede genauer betrachten und fünf Website-Archive exemplarisch vorstellen. Dabei ist genau zu prüfen, nach welchen Kriterien dabei Daten gesammelt werden und inwiefern diese dann auffindbar und in Zukunft nutzbar sind.

### 3. Archive. Fünf Fallbeispiele.

#### **hallointer.net**

Auf seiner Website [hallointer.net](http://hallointer.net) betreibt Grafikdesigner **David Liebermann** seit 2014 eine „platform for contemporary internet“ ([Liebermann, o. J.](#))<sup>2</sup>, mit dem Ziel „to create a pool of impressions to get many different views on the Internet“ ([ebd.](#))<sup>2</sup>. Die Seite sammelt Websites – meist aus dem zeitgenössischen Designumfeld – und präsentiert diese dann chronologisch sortiert in linearer Abfolge. Durchschnittlich kommen pro Monat mehrere neue Websites zu dieser Sammlung hinzu.

Alle gesammelten Seiten werden mit einem Screenshot dokumentiert und präsentiert, um einen schnellen visuellen Eindruck der jeweiligen

Website zu vermitteln. Eine Besonderheit der Seite im Vergleich zu anderen Website-Archiven ist dabei der hohe Grad an Interaktivität, den sie ermöglicht: Die archivierten Websites werden nämlich nicht nur mittels dieses Screenshot gezeigt, sondern sie können (größtenteils) auch direkt innerhalb der hallointer.net-Website genutzt werden, da sie über einen Inlineframe<sup>11</sup> „live“ eingebunden sind (→ Abb. 9). Die visuelle Einrahmung jeder gesammelten Seite in einem stilisierten Rechner unterstreicht diese Nutzung einer Website innerhalb einer Website.



Abb. 9: hallointer.net (Screenshot, 19.04.2021)

Durch diese doppelte Abdeckung, den Screenshot einerseits und vor allem das interaktive Einbinden der archivierten Website selbst, werden die Vorteile verschiedener Archivierungsstrategien miteinander verbunden. Die User\*innen können sich sowohl extrem schnell einen visuellen Überblick über die archivierten Seiten verschaffen (insbesondere im Index, der verkleinerte Thumbnails in einem großen Raster präsentiert (→ Abb. 10)), als auch anfangen, die jeweiligen Seiten gleich auszuprobieren und zu nutzen. Letzterer Aspekt ist insbesondere deshalb wichtig, weil viele der gesammelten Websites gerade aufgrund ihrer außergewöhnlichen interaktiven Funktionsweise interessant sind und oft wohl gerade deshalb in die Sammlung aufgenommen wurden.

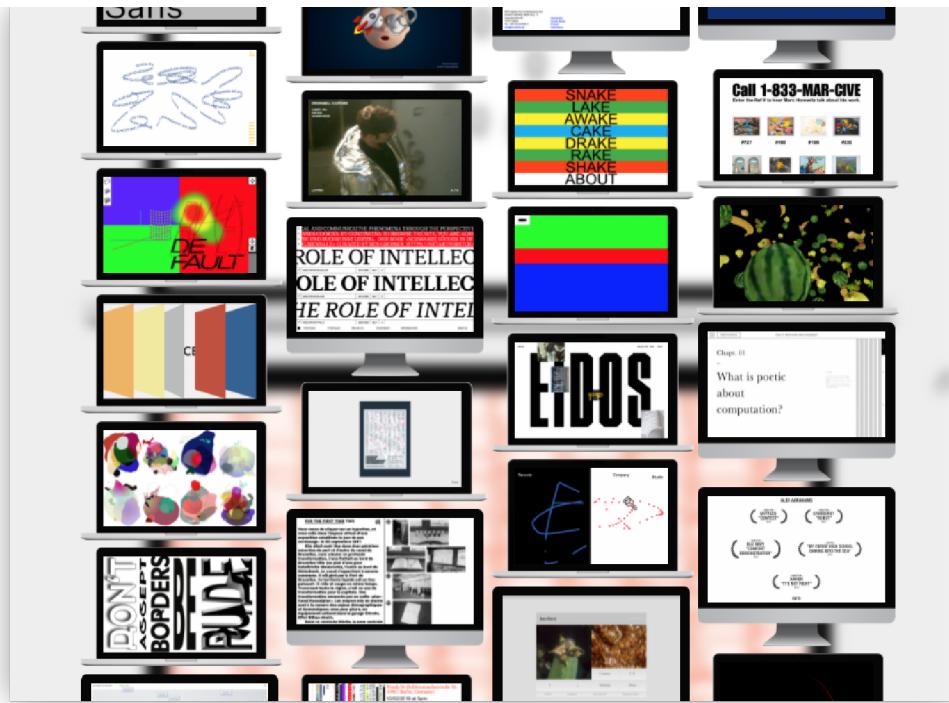


Abb. 10: hallointer.net-Index (Screenshot, 19.04.2021)

Eine Funktion, die diese Art von Archivierung allerdings nicht leisten kann, ist, diese interaktive Komponente der Websites zu fixieren, das heißt dafür zu sorgen, dass diese auch dauerhaft verfügbar und nutzbar ist. Wie im Internet üblich, kann es also vorkommen, dass sich Websites ändern oder offline gehen. So gibt es beispielsweise [Seiten](#), deren Screenshot klar erkennen lässt, dass die live eingebundene Website nichts mehr mit der Website zum Archivierungszeitpunkt zu tun hat (→ Abb. 11). Das ist bei einem so offensichtlichen Abweichen vom Screenshot schnell erkennbar, bei anderen Websites, die sich nur geringfügig geändert haben, ist aber eventuell weniger offensichtlich, dass sie nicht mehr der Version zum Archivierungszeitpunkt entsprechen.

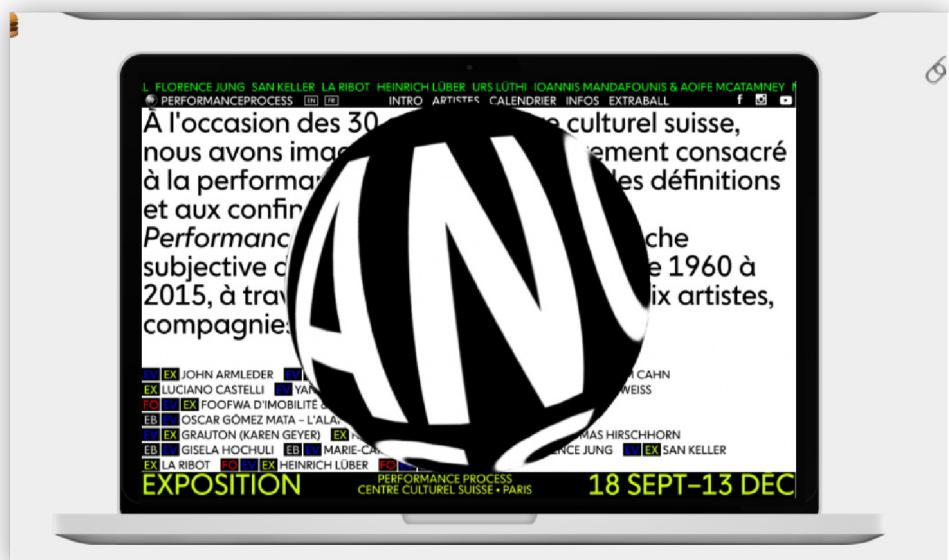


Abb. 11: hallointer.net, Screenshot vs. live eingebundene Seite (Screenshots, 19.04.2021) (1/2)

Die spezifische Machart der Seite lässt allerdings auch erkennen, dass hier nicht der Anspruch ist, ein historisch akkurate Archiv zu errichten. So wird zum Beispiel gänzlich darauf verzichtet ein Datum oder sonstige Metadaten zu den Websites anzugeben<sup>12</sup>. Der Schwerpunkt liegt vielmehr darauf, das „contemporary internet“ abzubilden und dieses dabei auch innerhalb des Mediums selbst darzustellen. Dass die entsprechenden Inhalte sich jederzeit ändern können, ist konstitutives Merkmal dieses Mediums und wird entsprechend in Kauf genommen.

Dieses Vorgehen stimmt mit einer Forderung des Medienkünstlers und Digitalarchivars **Dragan Espenschied** überein, der auf das Problem hinweist, dass Geschichtsschreibung in Bezug auf digitale Kultur meist nur außerhalb dieser stattfindet und dass diesbezüglich ein Umdenken stattfinden müsse:

„The grief I have with the creation of history in digital culture is that it is in many cases located outside of digital culture itself. Digital culture is regarded as too flimsy (or the classic “ephemeral”) to take care of itself, so conservation is done by removing artifacts from the cultural tempest they originated in and putting them into a safe place. [...] [This] doesn’t do digital culture any justice [...] because many artifacts can only blossom in their environment, in concert or contrast with a vernacular web [...]“

(Owens, 2014) ↗

Ganz im Sinne dieser Feststellung vermeidet David Liebermann durch die beschriebene Live-Einbettung diese Übersetzung digitaler Inhalte in eine traditionellere Archivform gänzlich, indem er alle Websites „live“ auf seiner Website zusammenführt und dadurch sozusagen eine eigene, kleine kuratierte Version des Internets erstellt.

## **net.artdatabase.org**

Ein weiteres Archivierungs-Projekt, das allerdings einen völlig anderen Ansatz verfolgt, ist die Website [net.artdatabase.org](http://net.artdatabase.org). Sie wurde 2011 gemeinsam von Webkünstler **Constant Dullaart** und Kunsthistoriker **Robert Sakrowski** initiiert, und hat zum Ziel Net-Art-Projekte zu dokumentieren. Grundlage dieses Vorhabens ist dabei eine eigens entwickelte Dokumentationsmethode, die diese Projekte samt ihrer Nutzungsinteraktion bis ins kleinste Detail festhalten soll:

„Dullaart and Sakrowski have developed a method to document net art that aims to move beyond the technical specifications and the interaction model of the artwork. They try to capture the reception of net art in an environment in

which it was originally perceived. As Sakrowski explains, ,the context, the private atmosphere, and the hardware interaction defines a large part of the «net art activity».““

(*Spreeuwenberg, 2012, S. 4*)<sup>12</sup>

Neben der Seite selbst sollen also auch die Interaktion mit der Website und sogar die Umgebung, in der die Website genutzt wurde, dokumentiert werden, um so das Nutzungserlebnis möglichst authentisch festzuhalten. Um diesen Anspruch zu erfüllen, wird die Seitenutzung durch zwei Videos festgehalten: eine Bildschirmaufnahme, sowie ein Video, das den\*die Nutzer\*in beim Besuchen der Website zeigt. Beide Videos werden dann im Archiv synchron im Splitscreen gezeigt (→ Abb. 12).

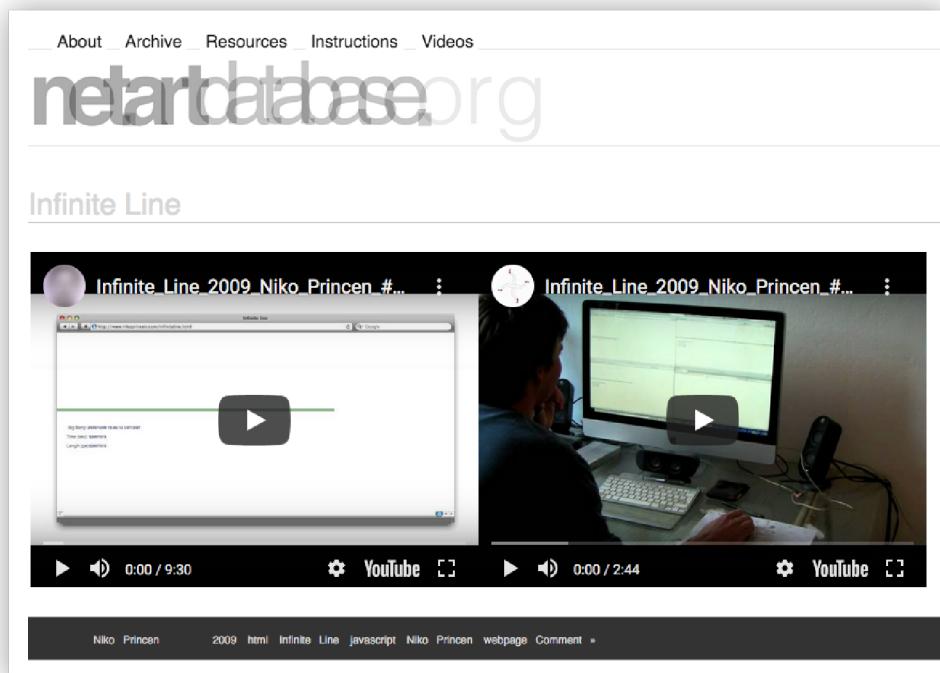


Abb. 12: net.artdatabase.org (Screenshot, 16.03.2021)

Dieses recht aufwendige Verfahren wurde in zwei Pilotversuchen getestet und ausgiebig dokumentiert, um so seine Vor- und Nachteile zu untersuchen. Dabei stellte sich schnell heraus, dass die Methode zwar geeignet ist, die Interaktion zwischen Nutzer\*innen und Website festzuhalten, dass sie allerdings gleichzeitig eine wirklich authentische Nutzungssituation fast vollständig unterbindet: Da die Nutzer\*innen sich während der Aufnahme der externen Beobachtungssituation nur allzu bewusst sind, nutzen sie die Website mitunter völlig anders, als sie es vielleicht tun würden, wenn dies unbeobachtet geschähe. Im anschließenden Interview äußerte eine Probandin: „I don't know how it is supposed to react, did I see everything? I just clicked around.“ (*ebd., S. 14*)<sup>13</sup>. Dieses „supposed to“ zeigt, dass die Nutzung vor allem durch eine vermeintliche äußere Erwartungshaltung beeinflusst wurde. Das Bedürfnis, alles von der Website gesehen zu haben, entspringt dabei wohl weniger einem eigenen Wunsch, als vielmehr dem Wunsch, zum Zweck der Aufnahme einmal die gesamte Seite abzubilden.

Andere Entscheidungen, wie zum Beispiel die Umgebung oder die Auswahl der Websites und der Nutzer\*innen selbst sind natürlich ebenfalls in dem Wissen getroffen worden, dass eine ziemlich aufwendige Form der Aufzeichnung stattfinden wird, weshalb die Dokumentation nicht repräsentativ ist, sondern eben nur *eine* mögliche Nutzung unter extrem künstlichen Bedingungen zeigt.

Zudem besteht das Problem des verhältnismäßig hohen Aufwands dieser Art der Dokumentation. Eine Kamera samt Stativ muss so eingerichtet werden, dass sie das Geschehen über die Schulter der Nutzer\*innen filmt, man muss eine Bildschirmaufnahme erstellen, dann im Anschluss die beiden Aufnahmen zusammenbringen und synchronisieren und sie schließlich ins Archiv einstellen. Ein Blick in das heutige Archiv scheint diese Aufwendigkeit des Verfahrens zu bestätigen: Nach den zwei Pilotversuchen scheint noch eine Reihe an weiteren Websites dokumentiert worden zu sein, allerdings stehen insgesamt nur elf solcherart archivierter Arbeiten online (von denen zudem fast alle nicht mehr funktionsfähig sind).

Das steht quantitativ im krassen Gegensatz zum zuvor diskutierten Archiv von David Liebermann, das im siebten Jahr seines Bestehens bereits 596 Websites archiviert hat (Stand 16.03.2021). Zufällig saß ich zu Beginn meines Studiums einmal neben David, als er mit Handy am Ohr und seinem Laptop auf dem Schoß eine neue Website auf seiner Plattform online stellte. Der ganze Prozess – Screenshot aufnehmen, hochladen, Eintragen der Metadaten – dauerte trotz des parallel laufenden Telefonats nur wenige Minuten.

Auf der Seite der net.artdatabase befindet sich auch eine [Anleitung](#), um selbst eine Seite zum Archiv beizusteuern. Das dort verfügbare Formular ist allerdings nicht mehr nutzbar und auch die vorgeschlagenen Software-Tools und Workflows zum Erstellen der Videos ([vgl. ebd., S. 10f.](#))<sup>2</sup> sind inzwischen veraltet. Eine weniger starre Anleitung, die die benötigten Komponenten eher in ihrer Funktion beschrieben hätte, anstatt sich in technischen Details zu verlieren, hätte vermutlich zu einer höheren Partizipation führen können.

In einem E-Mail-Wechsel mit Constant Dullaart habe ich ihn deshalb nach seiner Einschätzung gefragt, ob sich die vorgeschlagene Methode für das Erstellen von Web-Archiven überhaupt eignen würde, oder ob sie für diesen Zweck zu aufwendig sei. Er erklärte, dass er sie tatsächlich für zu aufwendig für das kollaborative Erstellen einfacher Web-Archive halte und er beim Entwerfen seiner Methode eher die Anwendung in einem institutionellen Kontext im Sinn gehabt hätte, wenn es also beispielsweise darum ginge, im Kunstmuseum Net-Art-Arbeiten zu archivieren. Für diesen Einsatzzweck habe sich die Methode dann doch als eine verhältnismäßig unkomplizierte Alternative zu den sonst gebräuchlichen, übermäßig komplexen institutionellen Archivierungsprozessen erwiesen. Die Methode komme laut Dullaart

in abgewandelter Form bei der kommerziellen Archivierungsdienstleistung **ArtHost** zum Einsatz, die er gemeinsam mit dem Amsterdamer Institut für Medienkunst **LIMA** entwickelt hat ([vgl. LIMA, o. J.](#)).<sup>12</sup> Dullaart selbst verwendet zum Archivieren seiner eigenen Net-Art-Arbeiten eine Kombination aus dieser Methode, aus **Conifer** (das ich in einem [kommenden Abschnitt](#) bespreche) und Scraping<sup>13</sup> (Dullaart, persönliche E-Mail, 16.03.2021).

Trotz der beschriebenen Nachteile muss man der Archivierungsmethode zugutehalten, dass sie es als eine der wenigen leistet, die tatsächliche Nutzung der Website samt Umgebung zu dokumentieren. Würde man heute also auf ein gut gepflegtes Archiv stoßen, das sich dieser Methode bedient, dann hätte dieses Archiv durchaus einen gewissen historischen Wert und würde Aufschluss über die Digitalkultur vergangener Tage geben.<sup>14</sup> Realistisch betrachtet erscheint die Methode aber so aufwendig, dass sie eher im Kunstkontext Anwendung findet und nicht beim Archivieren konventioneller Websites.

## Are.na

Neben den Plattformen, die vornehmlich von Einzelpersonen kuratiert werden, gibt es auch kollaborative Plattformen, auf denen man – meist nach Anmeldung – selbst Websites in Sammlungen organisieren kann. Eine solche Plattform ist die Website [Are.na](#). Die Seite ist als kollaboratives Sammelwerkzeug für digitale Inhalte konzipiert, das die Möglichkeit bietet, die so zusammengetragenen Inhalte mit denen anderer Nutzer\*innen zu verbinden und so immer neue Untersammlungen zu bilden und andere zusammenhängende Inhalte zu finden. Laut Eigenbeschreibung ist die Plattform „a place to structure your ideas and build new forms of knowledge together“ ([Are.na, o. J.](#)). Aufgrund dieser vorgeblichen Eignung für den Inspirationsprozess, für die Ideenfindung oder die visuelle Recherche ist die Plattform vor allem im kreativen Umfeld beliebt ([vgl. Gotthardt, 2018](#)).

Es lassen sich auf Are.na verschiedene Medien-Inhalte wie Bild-, Audio- und Videodateien, Textschnipsel und eben auch Websites sammeln und diese können alle gemeinsam in einem „Channel“ gespeichert und dann auf der Plattform geteilt werden. Jeder Inhalt wird als quadratische digitale Karteikarte – im Are.na-Jargon „Block“ – dargestellt, so dass alle Blocks hierarchisch gleichwertig nebeneinanderstehen (→ Abb. 13). Das Abspeichern einer Website erfolgt schlicht über das Eingeben ihrer URL. Nach dem Anlegen eines solchen Blocks wird dann automatisch ein Screenshot der Website erstellt, der fortan als visuelle Präsentation der so gesammelten Website dient. Neben dem Screenshot wird eine Verlinkung zur Website angelegt und es gibt einen Bereich, in dem Metadaten und Kommentare angezeigt werden, sowie Verbindungen zu anderen Sammlungen gelistet sind (→ Abb. 14).

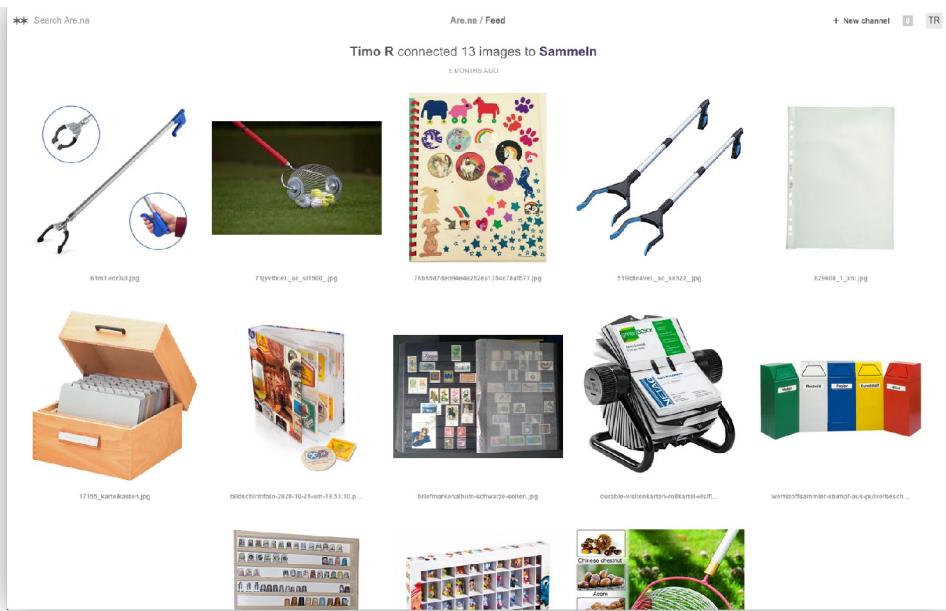


Abb. 13: Are.na-Blocks (Screenshot, 19.04.2021)

Abb. 14: Arena-Screenshot und Metadaten (Screenshot, 19.04.2021)

Die Plattform bietet die Möglichkeit, sich mit anderen Nutzer\*innen in Gruppen zusammenzuschließen oder diesen zu folgen und, wie bereits erwähnt, ihre Inhalte mit den eigenen zu verbinden. Sie beinhaltet also durchaus Komponenten, die typisch für soziale Netzwerke sind. Allerdings verweist Mitbegründer **Charles Broskoski** darauf, dass bewusst darauf verzichtet wurde einen Like-Button einzubauen. Die einzige Möglichkeit, einen Inhalt zu „liken“ sei daher, diesen mit einer eigenen Sammlung zu verknüpfen, was ein Nachdenken darüber erfordere, wo genau eine Sache inhaltlich einzuordnen ist: „This is a much harder action to take and takes a cognitive load to think about where something belongs“ ([Broskoski, 2019](#)). Ziel ist es, so ein großes Netzwerk an verknüpften Informationen entstehen zu lassen, das dann potenziell für die kreative Recherche nutzbar ist.

Was auf den ersten Blick wie eine einfache Möglichkeit der Website-Archivierung erscheint, kommt auf den zweiten Blick mit einigen Problemen daher. Meinem persönlichen Eindruck nach hat das User Interface einen deutlichen Spielraum für Verbesserungen: Es ist ausgesprochen unübersichtlich, beim Navigieren der Seite verliert man schnell den Überblick auf welcher Ebene man sich gerade befindet: Bin ich in einem „Channel“, bin ich in einer „Group“, bin ich in meinem „Feed“, warum sehe ich in der aktuellen Ansicht eine Mischung aus „Blocks“ und „Channels“? (→ Abb. 15) Für eine Plattform, deren Anspruch es ist Ideen zu verknüpfen, ist die Suchfunktion auch deutlich verbesserungswürdig: Der Suchalgorithmus scheint nach dem Suchbegriff ausschließlich in den Titeln der abgelegten Inhalte oder Channels zu suchen und häufig bekommt man auf der Ergebnis-Seite eine Aufzählung riesiger, weißer Kacheln präsentiert, die allesamt Channels mit fast identischen Namen darstellen. Es gibt keinerlei visuellen Hinweis, welche Inhalte sich in den jeweiligen Sammlungen befinden. Zudem scheinen die Channels nicht einmal nach Relevanz sortiert zu sein, so dass man häufig einen vollkommen leeren Channel an erster Stelle der Suchergebnisse findet (→ Abb. 16). Der Vorteil, dass hier ein großer Wissenspool mit vielen, manuell verknüpften Inhalten entsteht, wird so von der Plattform nicht richtig ausgespielt. So ist es oftmals eher Glückssache, ob man zu einem gesuchten Begriff eine passende Sammlung findet, die auch wirklich einen Mehrwert bietet.

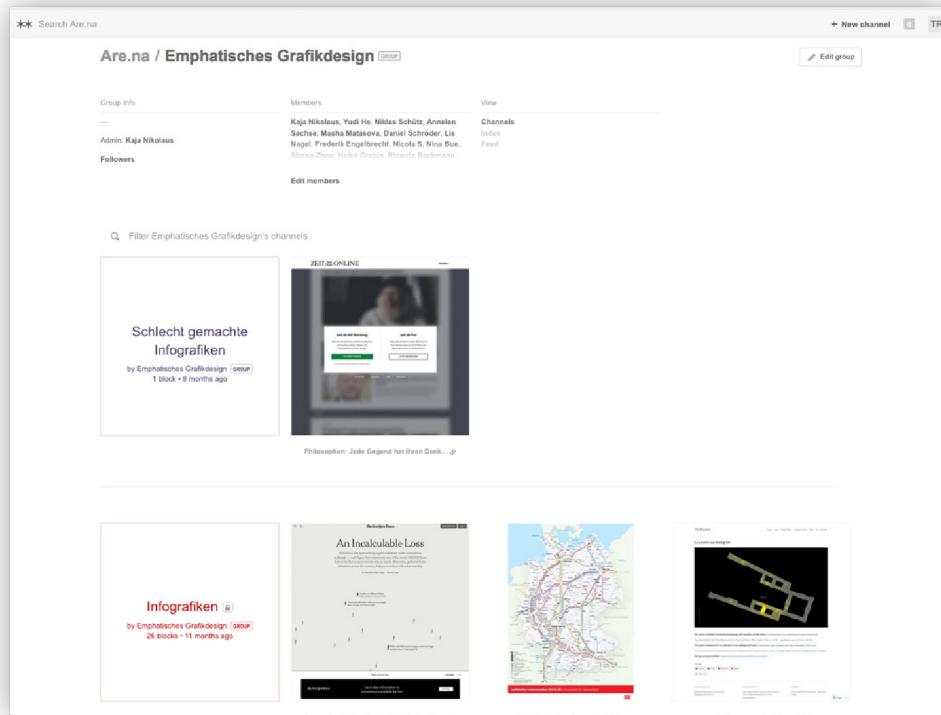


Abb. 15: Are.na, Channels und Blocks in einer Ansicht (Screenshot, 19.04.2021)

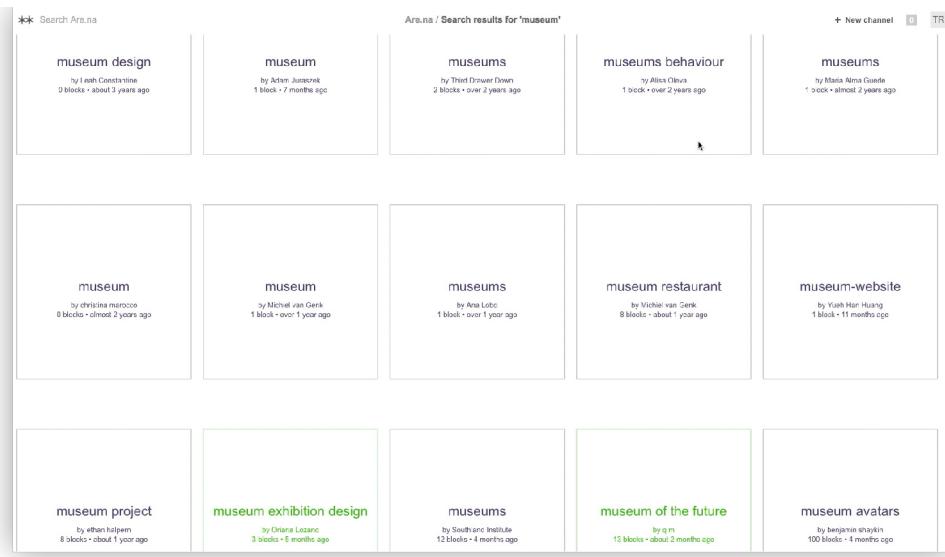


Abb. 16: Are.na, Suchergebnis zum Schlagwort „Museum“ (Screenshot, 22.03.2021)

Überprüft man, was die Plattform leisten kann, wenn es um das Archivieren von Websites geht, fällt das Ergebnis ähnlich ernüchternd aus: Es werden im Grunde nur die URL und der automatisierte Screenshot gespeichert, das heißt es handelt sich eher um eine Sammlung visualisierter Bookmarks als um ein richtiges Archiv gespeicherter Inhalte. Vor dem Verschwinden der Webinhalte ist man allenfalls durch den Screenshot abgesichert, er ist das Einzige, was bleibt, sobald die zugehörige Website nicht mehr online ist. Da diese Screenshots aber automatisiert erstellt werden, kommt es häufig vor, dass sie wenig hilfreich sind: Oft werden die Seiten im Moment dieser Screenshot-Aufnahme von Cookie-Popups oder Adblocker-Hinweisen überlagert, so dass am Ende dieses Vorgangs dann ein Thumbnail entstanden ist, der bei der visuellen Zuordnung der gesammelten Website in keinster Weise hilfreich ist (→ Abb. 17). Are.na erlaubt zwar im Rahmen einer kostenpflichtigen Premium-Mitgliedschaft einen eigenen Thumbnail für jeden Block festzulegen, da aber nur ein Bruchteil der Mitglieder eine solche Mitgliedschaft hat, gibt es zahlreiche Sammlungen, deren Thumbnails von diesem Problem betroffen sind, was die entsprechenden Sammlungen für alle Nutzer\*innen schwerer nutzbar macht.

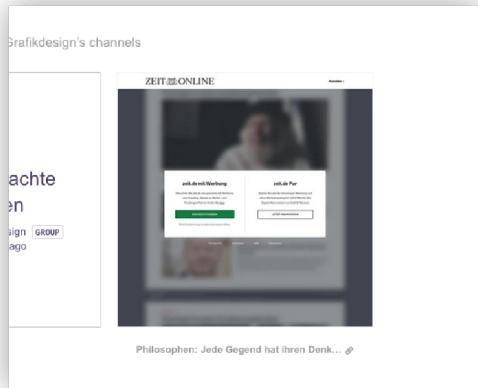


Abb. 17: Are.na, Thumbnail einer Website einer Infografik, überlagert von einem Adblocker-Banner (Screenshot, 19.04.2021)

Die Idee, mit der die Plattform laut Broskoski angetreten ist, „to make a coherent network of information“ ([ebd.](#))<sup>14</sup> und „[to] use the stuff to build new ideas, new knowledge“ ([ebd.](#))<sup>15</sup>, bleibt auf halbem Wege stehen. Es ist zwar durchaus möglich, beim Herumstöbern auf neue Inhalte zu stoßen und diese dann mit den eigenen zu verbinden, allerdings steht sich die Plattform durch das kommerzielle Modell und das geschlossene Format, sowie durch die verbesserungswürdige Nutzbarkeit letztendlich selbst im Weg. Websites lassen sich zwar sammeln, sind aber im unübersichtlichen User Interface nur schwierig zu organisieren und noch schwerer aufzufinden und sichern darüber hinaus nicht vor dem eventuellen Verschwinden der Original-Seite ab. Es ist sogar davon auszugehen, dass mit fortschreitender Zeit der Anteil an „toten“ URLs innerhalb der Sammlungen steigen wird und damit der Nutzen dieser Art der Speicherung noch weiter abnimmt.

## Conifer

Die amerikanische Non-Profit-Organisation **Rhizome**, deren Aufgabenfeld im Ankaufen, Präsentieren und Archivieren von Net Art liegt ([vgl. Rhizome, o. J.](#))<sup>16</sup>, entwickelt seit 2015 das digitale Archivierungswerkzeug **Conifer** und stellt dieses kostenlos unter einer Open-Source-Lizenz zur Verfügung. Das Tool, das vormals unter dem deskriptiven Namen **Webrecorder** verfügbar war, ermöglicht Nutzer\*innen Websites bei deren Nutzung „aufzunehmen“ und sie somit dauerhaft zu speichern. Diese Aufnahmen können entweder über das Online-Tool auf der [Conifer-Plattform](#) erstellt werden oder alternativ mit einer Desktop-App, die wiederum **ArchiveWeb.page** heißt.<sup>17</sup>

Im Gegensatz zu Archivierungswerkzeugen wie der Wayback Machine des Internet Archive werden die gesammelten Websites nicht in ein großes Gesamtarchiv aufgenommen, sondern werden zunächst einmal nur in einer persönlichen Sammlung innerhalb des Nutzer\*innen-Accounts gespeichert. Der\*die Nutzer\*in kann dann selbst entscheiden, ob diese Sammlung privat bleibt, oder ob sie veröffentlicht wird, so dass auch alle anderen auf die gesammelten Websites zugreifen können.

Das Tool funktioniert so, dass man eine „Session“ aufnimmt, man gibt also eine URL an und betätigt dann einen digitalen Aufnahmeknopf, der diese Aufnahmesession startet. Nun kann man innerhalb der entsprechenden Website navigieren und beliebig Inhalte laden oder sogar zu anderen Websites wechseln. Die Besonderheit des Tools – und der große Unterschied zu anderen Archivierungswerkzeugen wie der Wayback Machine – liegt darin, wie und welche Daten abgespeichert werden.

Archiviere ich beispielsweise eine Website in der Wayback Machine, gebe ich die URL der Seite ein und auf den Servern der Plattform sorgen Archivierungstools vollautomatisiert dafür, dass die Seite aufgeru-

fen wird und dann im Archiv abgespeichert wird. Dabei wird die Seite aber in der Regel einfach nur geladen, ohne dass weitere Interaktionen ausgeführt werden. Es werden also beispielsweise keine eingebetteten Videos gestartet oder es wird nicht automatisiert die Seite heruntergescrollt, um so zum Beispiel auch Inhalte zu laden, die erst beim Runterscrollen nachgeladen werden, wie es heute bei zahlreichen Websites der Fall ist. Auch scheitern Werkzeuge wie die Wayback Machine am sinnvollen Abspeichern von Seiten wie **Facebook**, **Instagram** oder **Twitter**. Um überhaupt die Inhalte dieser Plattformen zu sehen, muss man dort eingeloggt sein. Aufgrund der fehlenden Interaktion kann sich die Wayback Machine aber gar nicht auf diesen Seiten einloggen und eine solche Funktion würde auch gar keinen Sinn ergeben: Diese Portale zeigen personalisierte Content-Streams für die eingeloggten User\*innen an und können daher nicht in dieser Form stellvertretend für die Allgemeinheit archiviert werden. Wenn man in der Wayback Machine solche URLs aufsucht, findet man höchstens eine archivierte Version der Login-Page. Sämtliche Seiten, deren Betrachtung ein Login benötigt, sind also von den Archivierungsmechanismen der Wayback Machine ausgeschlossen.

Conifer löst dieses Problem, indem das Werkzeug direkt die Daten, die vom Browser abgerufen werden, abspeichert. Wenn ein\*e Nutzer\*in also ein eingebettetes Video auf der Seite aufruft oder sich durch eine Slideshow klickt und dabei Inhalte nachlädt, die beim anfänglichen Seitenaufruf noch gar nicht mitgeladen wurden, so speichert Conifer diese neugeladenen Daten ebenfalls ab undbettet sie im Archiv derart in den Gesamtkontext mit ein, dass sie beim Aufruf der archivierten Version ganz normal abgerufen werden können, so als wäre man gerade zum Zeitpunkt online, an dem die Seite archiviert wurde. Damit umgeht Conifer einen der großen Schwachpunkte vieler anderer Tools, ein Problem, das in einem Rhizome-Blogartikel – in Anlehnung an den zuvor erwähnten *link rot* – als *embed rot* bezeichnet wird ([vgl. Kreymer/Espenschied, 2019](#)): Das „Verrotten“ eingebetteter Inhalte, die sonst Gefahr laufen würden verloren zu gehen. Selbst Websites, die ein Login benötigen, sind so archivierbar, wodurch es dann auch möglich wird die persönlichen Timelines der Social-Media-Netzwerke zu speichern.

Auch Conifer kann am Speichern komplexerer, interaktiver Inhalte scheitern, allerdings ist die Erfolgschance solche Inhalte zu erhalten deutlich höher als bei anderen Archivlösungen:

„On the technical side, Conifer focuses on “high fidelity” web archiving. Items relying on complex scripting, such as embedded videos, fancy navigation, or 3D graphics have a much higher success rate for capture with Conifer than with traditional web archives.“

([Conifer, o. J.](#))

Aufgrund des höheren Grads an Personalisierung der Inhalte werden die Sammlungen standardmäßig zunächst privat angelegt und man muss sie veröffentlichen, um sie dann per Link mit anderen Nutzer\*innen zu teilen. Deshalb bietet Conifer keine durchsuchbare Plattform, wie man es von einer Archivlösung vermutlich erwarten würde, einzig durch geteilte Links erhält man also Zugang zu anderen Sammlungen.

Auch wenn das beschriebene Vorgehen deutlich erfolgreicher darin ist Seiten samt Inhalten komplett zu speichern, so ist dennoch nachvollziehbar, warum es in dieser komplexen Form bei größeren Archiven wie dem Internet Archive nicht zum Einsatz kommen kann: Die Menge der Inhalte und Daten, die auf manchen Seiten abgerufen werden kann, ist potenziell endlos, wie beispielsweise bei Seiten, die viele Videos einbinden, oder bei den vielen Blogs, die sich heutzutage mittels „Infinite Scrolling“ immer weiter scrollen lassen. Auch bei komplexeren Navigationsstrukturen der Websites ist nicht immer klar erkennbar oder festgelegt, in welcher Abfolge man durch die Seite navigieren soll, so dass eine generelle Automatisierung hier unmöglich ist und letztlich immer noch die Nutzer\*innen selbst diese Navigation steuern müssen.

Beim Abspeichern individueller Nutzungs-Sessions oder bei ähnlichen Szenarien, an denen andere Werkzeuge scheitern, eignet sich Conifer allerdings ganz hervorragend zum Archivieren von Webinhalten. Als offene Plattform erlaubt es zudem das Exportieren der so archivierten Daten in offene Standardformate, so dass eine spätere Nutzung außerhalb der Plattform möglich bleibt.

## One Terabyte of Kilobyte Age

Die Netzaktivistin **Olia Lialina** und ihr Partner **Dragen Espenschied** arbeiten seit 2010 mit einem riesigen Archiv alter GeoCities-Websites. **GeoCities.com** war in der zweiten Hälfte der 1990er-Jahre ein sogenannter Freehoster, ein kommerzieller Anbieter also, der Nutzer\*innen kostenloses Webhosting anbot und damit in den Anfangstagen des Internets vielen erstmals ermöglichte eine eigene Homepage zu entwerfen und diese online zu stellen. Unter diesen Voraussetzungen entstanden – Jahre vor den sozialen Netzwerken – eine große Internet-community und ein kleines Universum an Amateur-Homepages. Lialina bezeichnet diese Zeit als ein „Amateur Web“ im positiven Sinne: Das Internet war noch nicht durch Professionalisierung vereinnahmt und Nutzer\*innen teilten ihren Content nicht vornehmlich auf vorstrukturierten Social-Media-Plattformen, sondern ließen eben meist auf eigenen kleinen Homepages – oft zu banalen oder obskuren Themen – ihre eigenen kleinen Welten entstehen. 1999 wurde die Plattform schließlich von **Yahoo!** gekauft und der Blick auf diese Websites änderte sich: Was zuvor noch als Möglichkeit des individuellen Ausdrucks im jungen World Wide Web gegolten hatte, wurde plötzlich ne-

gativ betrachtet: „Having a page on there became a synonym for diletantism and bad taste.“ ([GeoCities Research Institute, 2011](#)) Nachdem die Zahl der aktiven Nutzer\*innen stetig abnahm, die Plattform langsam in der Bedeutunglosigkeit verschwand und die Homepages durch Social-Media-Profile ersetzt wurden, kündigte Yahoo! im April 2009 schließlich die baldige Schließung der Plattform an.

Diese Ankündigung wiederum rief bestimmte Teile der Netzgemeinde auf den Plan, die einen wichtigen Teil der Internetgeschichte von der Zerstörung bedroht sahen. So beschloss das **Archive Team**, ein zweckgebundenes Kollektiv an engagierten Freiwilligen, sich der Sache anzunehmen. Das Kollektiv hatte sich kurz zuvor mit dem Ziel formiert Internet-Inhalte vor ihrem Verschwinden zu bewahren, und proklamierte dies mit seinem Slogan „We Are Going to Rescue Your Shit!“ ([Archive Team, 2019](#)). Diese „Rettungsmissionen“ digitaler Inhalte fanden zunächst meist anlassbezogen statt, wenn ein bestimmter Datenbestand laut Ankündigung kurz vor der Löschung stand, später ging man auch mehr und mehr dazu über, digitale Inhalte präventiv zu sichern, für den Fall, dass diese unangekündigt irgendwann verschwinden oder von neuen Inhalten überschrieben werden würden. Die GeoCities-Archivierung wurde zur ersten großen Rettungsaktion des Kollektivs. Die Beweggründe, gerade diese Websites zu retten sind im Archive-Team-Wiki festgehalten:

„While the natural urge by some would be to let GeoCities sink into obscurity and death, [...] the fact remains that GeoCities was for millions of people the first experience dealing with the low-cost, full-color, world-accessible website and all the possibilities this contained. To not at least have the option of browsing these old sites would be a loss of the very history of the web from the side of the people who came to know it, not the designers who descended upon it. For that reason, Archive Team thinks GeoCities is worth saving.“

(„[GeoCities](#)“, 2020)

In einem großen, koordinierten Bemühen gelang es schließlich mithilfe automatisierter Tools einen Großteil der alten GeoCities-Websites zu sichern, bevor Yahoo! diese schließlich am 27. Oktober 2009 offline nahm. Ein Jahr später wurde dieses Archiv, 641 Gigabyte groß, als Torrent<sup>16</sup> veröffentlicht und kann seither – die nötige Geduld und Festplattenkapazität vorausgesetzt – von allen Internetnutzer\*innen heruntergeladen werden. Es beinhaltet über 1,2 Millionen Profile, wobei nicht mehr zu rekonstruieren ist, welcher Prozentsatz der ursprünglichen Daten damit erhalten ist ([vgl. Lialina, 2012](#)).

Lialina und Espenschied begannen 2010 ebenfalls mit dem Download und machten sich dann an die Arbeit, das Archiv zu untersuchen und es künstlerisch zu nutzen. Unter dem Titel **One Terabyte of Kilobyte**

**Age** starteten sie einen [Blog](#), in dem sie ihre Funde aus dem Archiv analysieren und in ihre künstlerische Praxis einbinden. Zudem betreiben sie parallel unter der Bezeichnung **One Terabyte of Kilobyte Age Photo Op** einen [weiteren Blog](#), in dem durchgehend in 20-Minuten-Abständen ein neuer Screenshot einer Website aus dem Archiv veröffentlicht wird. Der erste Blogeintrag dieses Projekts stammt vom 7. Februar 2013 und laut Espenschied ist genug Material vorhanden, dass noch bis ins Jahr 2027 hinein neue Posts erfolgen werden ([vgl. ebd.](#)). Die Websites werden als Screenshot präsentiert, jeweils umrandet vom Rahmen eines zeitgemäßen Browsers. Daneben stehen der Veröffentlichungszeitpunkt innerhalb des Blogs, Schlagwörter, das letzte Änderungsdatum der archivierten Website, sowie die Original-URL ([→ Abb. 18](#)). Diese ist natürlich seit der Löschaktion von Yahoo! nicht mehr aufrufbar, sie hilft aber in vielen Fällen dabei, eine archivierte Version der Seite über das Internet Archive aufzurufen.



Abb. 18: One Terabyte of Kilobyte Age Photo Op, archivierte GeoCities-Website (Screenshot, 19.04.2021)

Das von Liliana und Espenschied verwaltete Archiv fand zudem Eingang in zahlreiche andere ihrer Projekte und Ausschnitte davon wurden in Ausstellungen gezeigt. Liliana präsentiert auf ihrem [Twitter-Account](#) – gemischt mit anderen Inhalten – ebenfalls regelmäßig Websites aus dem Archiv ([→ Abb. 19](#)). Das Besondere an dieser Präsentationsform ist, dass sie die Screenshots nicht unkommentiert postet und damit für sich stehen und wirken lässt, sondern dass sie Zitate aus den GeoCities-Homepages hervorhebt und so die ehemaligen Web-

site-Betreiber\*innen aus der Vergangenheit sprechen lässt oder durch Kommentierungen mit ihnen in einen fiktiven Dialog tritt. Sie gibt damit den längst vergangenen Gedanken, Wünschen und Sehnsüchten dieser Betreiber\*innen eine Plattform und macht dabei gleichzeitig durch ihr Wirken viele dieser Websites wohl wichtiger, als sie es früher je waren. Liliana selbst reflektiert allerdings diesen stark intervenierenden Umgang mit dem Archiv und den so präsentierten Websites in zahlreichen Blog-Artikeln, Essays und Vorträgen ausführlich und weist auch kritisch auf die daraus resultierenden Schwierigkeiten hin.

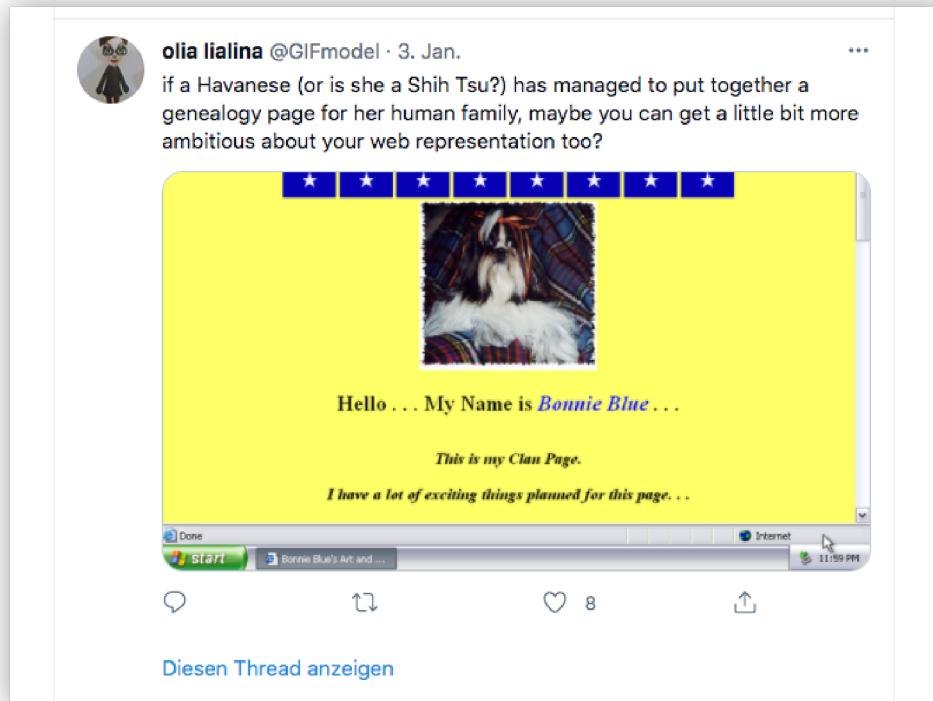


Abb. 19: Olia Lialina kommentiert auf Twitter einen Archivfund von GeoCities (Screenshot, 19.04.2021)

Viele der GeoCities-Homepages sind unvollständig, es fehlen beispielsweise verlinkte Grafiken. Es ist im Rückblick nicht mehr nachvollziehbar, ob diese Grafiken schon zu „Lebzeiten“ der Website fehlten, ob sie im Archivierungsprozess abhanden kamen oder vielleicht deshalb nicht mehr vorhanden sind, weil sie auf eine externe Quelle verlinkten, die diese Grafik jetzt nicht mehr bereitstellt. Lialina bezeichnet diese unvollständigen Seiten als „Web-Ruinen“ und hält fest, dass es sogar zu erwarten sei, dass diese Seiten so aussehen würden: „[T]he amount of destruction seems normal and even right to a certain extent: profiles that have been torn away from the web, never maintained, killed off, are not supposed to look any other way.“ ([ebd.](#))<sup>25</sup>. Diese Websites – ohne verlinkte Grafiken – müsse man daher nicht unbedingt abschreiben, sondern sie können in ihrer neuen Erscheinungsform eine ganz neue ästhetische Relevanz entwickeln (→ Abb. 20).

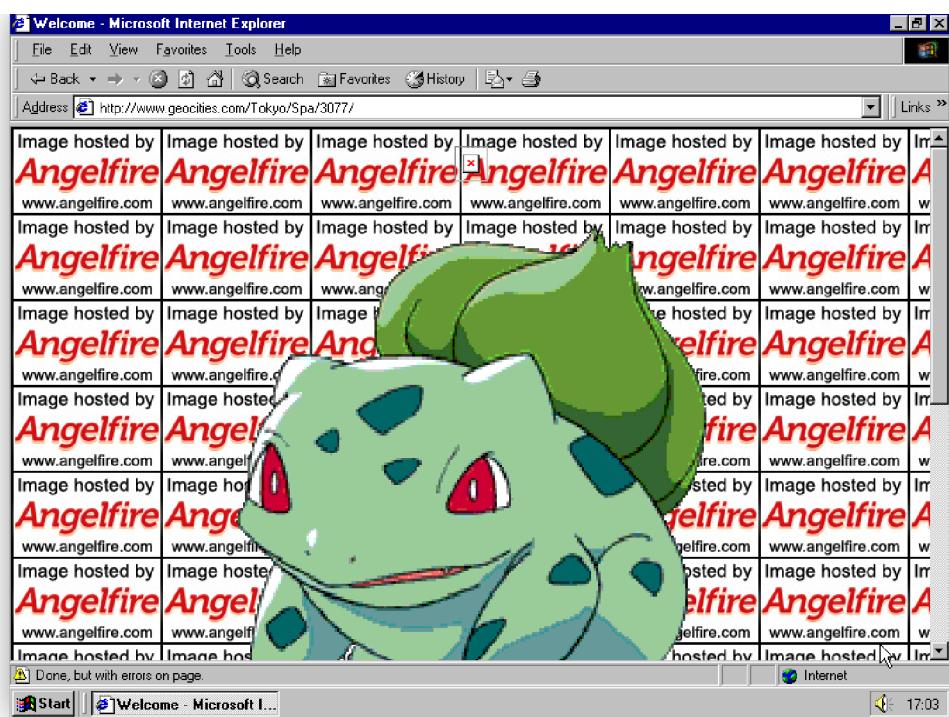


Abb. 20: Screenshot archivierter GeoCities-Website mit fehlenden Hintergrundgrafiken  
(Quelle: <https://pad.profolia.org/ruins>)

Lialina verweist zudem darauf, dass dieser Verfallsprozess andauernd sei. In einem Online-Vortrag zum Thema präsentierte sie den Screenshot (→ Abb. 21) einer Website, die 2009, als das GeoCities-Archiv erstellt wurde, wohl noch vollständig war. 2017 stellte allerdings der Imagehoster **Photobucket**, auf dem die Grafik der Website gehostet war, sein Gebührenmodell um, so dass sämtliche dort gehosteten Bilder plötzlich nicht mehr abrufbar waren. Der Screenshot der Website wurde erst nach 2017 erstellt, so dass das Bild darauf dann nicht mehr zu sehen war und die Website laut Lialina „einen zweiten Tod gestorben ist“ (vgl. Lialina, 2021) <sup>15</sup>.

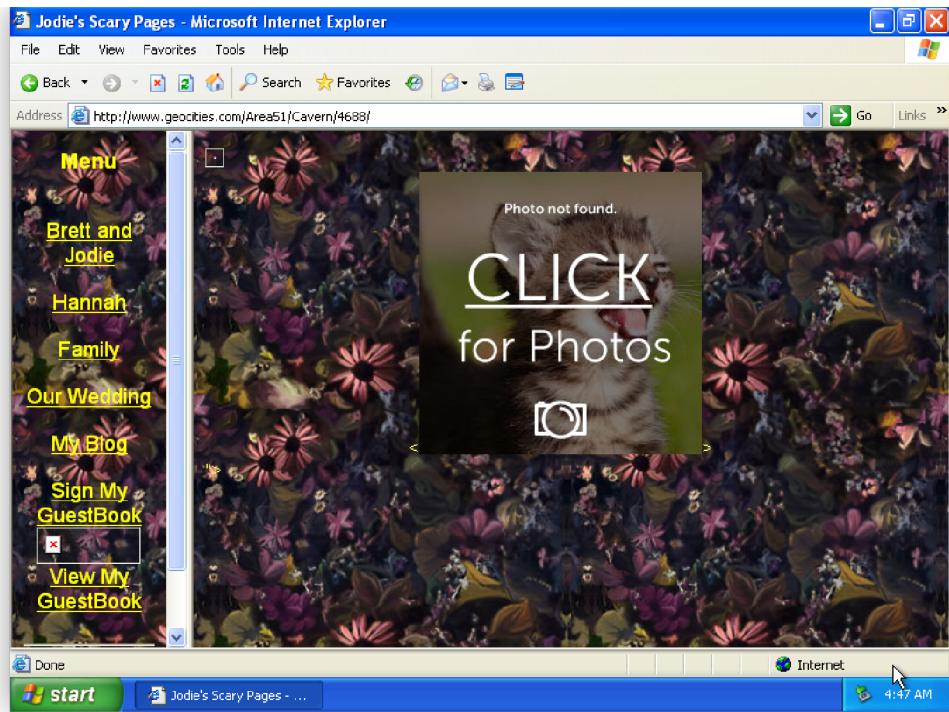


Abb. 21: Archivierte GeoCities-Website mit fehlender Photobucket-Grafik  
(Quelle: <https://pad.profolia.org/ruins>)

Mehr noch, manche der archivierten Seiten lassen laut Lialina klar erkennen, dass sie noch nicht fertig waren. Es handelt sich dabei zum Beispiel um die für das Web 1.0 so typischen „Under Construction“-Seiten, die eine kommende Seite ankündigen (→ Abb. 22). Ist eine solche Seite also jetzt in dieser Form im Archiv eingefroren, so sieht man nicht einmal die eigentliche Website – diese hat es nämlich offenkundig nie gegeben – sondern man sieht nur die Idee, die *Intention* einer Website. All die zuvor diskutierten Fragestellungen zur Authentizität, zu Versionen und zur Frage, was eine Website ist, werden in der fort dauernden Auseinandersetzung mit diesem Archiv in besonderer Weise sichtbar.

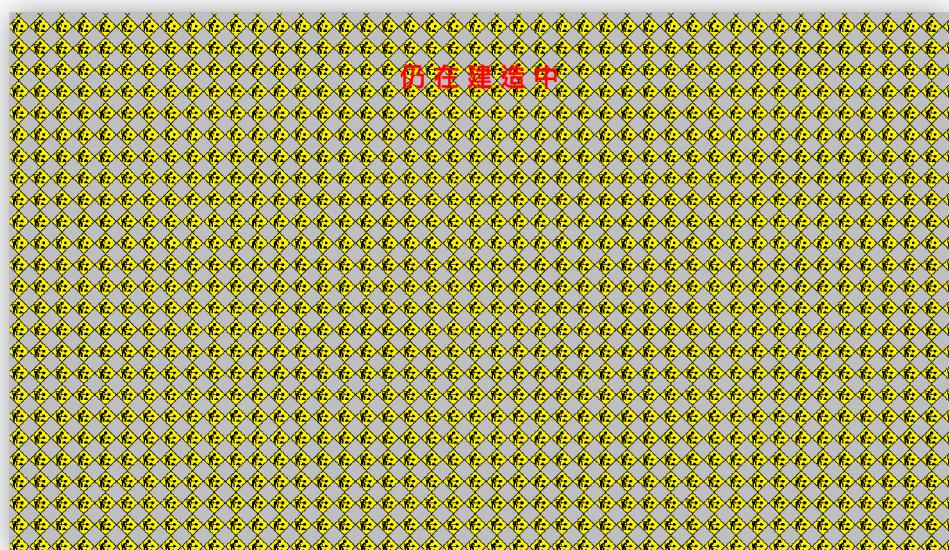


Abb. 22: Under-Construction-GeoCities-Website, aufgerufen über die Wayback Machine  
(Screenshot, 19.04.2021)

Im selben Vortrag ging Lialina auch auf das Problem des Erinnerns ein: Die Erinnerung derer, die GeoCities zum Zeitpunkt seines Bestehens noch live erlebt haben, verblasst langsam und wird nach und nach durch die Bilder ersetzt, die heute das gedankliche Bild von GeoCities prägen: die zahlreichen Screenshots nämlich, die aus dem großen Archivbestand entstanden sind. Was fehlt, sind die Animationen, die Sounddateien, die interaktiven Elemente und vor allem natürlich das eigentliche Erlebnis die Website zu nutzen – und das, obwohl das Archiv ja potenziell die Möglichkeit birgt, die Seiten auch heute noch – häufig freilich als Ruine – ans Laufen zu bringen. Die Screenshots, die sich in Umlauf befinden, sind aber so dominant, dass Lialina resümiert: „The snapshot replaces the actual memory.“ ([ebd.](#))<sup>12</sup> Das geht häufig mit einem gewissen Maß an rückblickender Verklärung einher und führt dabei nicht nur zu einem falschen Eindruck der Websites im Einzelnen, sondern im größeren Maßstab auch dazu, dass das Internet der 1990er-Jahre selbst zunehmend mit der visuellen Hinterlassenschaft dieses GeoCities-Archiv gleichgesetzt wird.

Angesichts all dieser Schwierigkeiten arbeiten Lialina und Espenschied künstlerisch und forschend mit dem Archiv weiter, *gerade um* die interessierte Netzöffentlichkeit auf diese Fragestellungen aufmerksam zu machen und mit ihr in einen Diskurs darüber zu treten.

## 4. Archive und Wirklichkeit

### Dark Archives

Wie im Abschnitt über das Erinnern gezeigt, ist das Anlegen von Sammlungen und damit das Bewahren von Daten vor dem Verfall nur der erste wichtige Aspekt in der Arbeit mit Archiven. Der zweite ist die Nutzbarkeit dieser Archive: Um mit den archivierten Daten überhaupt etwas anfangen zu können, müssen diese Daten auch wieder aufgefunden werden können. Das extreme Anwachsen der Speicherkapazität ist nur dann von Nutzen, wenn die Daten nicht in der Versenkung verschwinden, sondern genau in dem Moment wieder abrufbar sind, wenn sie uns nützlich werden. In Bezug auf dieses Kriterium gibt es Archive, die uns entgegenkommen und solche, die sich uns eher verschließen.

Die Künstlerin **Erica Scourti** hat sich in ihrer Arbeit mit dem Begriff *Dark Archives* auseinandergesetzt. In den Archiwissenschaften wird damit normalerweise eine parallel zum Hauptarchiv gepflegte Kopie eines digitalen Archivs bezeichnet, die im Falle einer Beschädigung der Hauptdatenbank als Ausfallsicherung dienen kann (Grau/Hoth/Wandl-Vogt (Hrsg.), 2019, S. 137). Dieses Dark Archive erlaubt im Regelfall keinen Zugriff und hält nicht die eigentlichen Daten des Archivs vor, sondern nur die Metadaten, die benötigt werden, um die Struktur des eigentlichen Archivs wiederherzustellen.

Scourtī erweitert nun den Begriff des Dark Archives, indem sie ihn auf solche Archive bezieht, innerhalb derer Daten schwer aufzufinden sind ([vgl. Scourtī, 2016](#)). Der Gegenbegriff dazu ist das *Light Archive*. Als Beispiel führt sie die Website des Internethändlers **Amazon** als ein *Light Archive* an, da die Artikel im Amazon-Katalog so vorgehalten werden, dass sie möglichst gut auffindbar sind: Es soll bestenfalls keine falsch benannten Artikel, keine Doppelungen, keinen Spam etc. geben, damit die Kund\*innen auf der Seite genau das finden, was sie suchen, die Seite weiter nutzen und sich nicht frustriert nach besseren Anbietern umsehen ([vgl. ebd.](#)). Unter einem Dark Archive versteht Scourtī entsprechend ein Archiv, dessen Daten *nicht* gut auffindbar sind und daher im Verborgenen bleiben. Das können Archive sein, die nur einzelnen Nutzer\*innen zugänglich sind oder auch Archive, die trotz einer Suchfunktion nicht ermöglichen die eigentlichen Daten „ans Licht“ zu befördern.

Erstaunlicherweise kann man dazu auch solche Archive zählen, deren Hauptaufgabe eigentlich das Aufsuchen von Daten ist, nämlich Suchmaschinen. Wenn ich in die Google-Suche einen beliebten Begriff, wie zum Beispiel „Sammlung“, eingebe, so wird mir zunächst eine hohe Anzahl an Suchergebnissen angezeigt, in diesem Fall „Ungefähr 95.300.000 Ergebnisse“ ([→ Abb. 23](#)). Wenn ich nun versuche, mich zu den hinteren Seiten der Suchergebnisse durchzuklicken, komme ich lediglich bis auf Seite 47 bzw. Suchergebnis Nummer 464 ([→ Abb. 24](#)), die Listung der Suchergebnisse hört dort unvermittelt auf. Das heißt Google mag zwar 95 Millionen Websites kennen, die das Wort Sammlung beinhalten, tatsächlich habe ich aber durch meine Suche nur direkten Zugang zu einem Bruchteil davon. Mir wird eine Verfügbarkeit von Informationen vorgetäuscht, die in dieser Form gar nicht gegeben ist. Der Rest der Informationen verbleibt im dunklen Archiv.

The screenshot shows a Google search results page for the query "Sammlung". The search bar at the top contains the word "Sammlung". Below the search bar, it says "Ungefähr 95.300.000 Ergebnisse (0,37 Sekunden)". A sidebar on the right is titled "Ergebnisse für Sammlung" and includes a snippet: "Ein Museum zeichnet sich durch eine Sammlung oft ...". The main search results area has a heading "Wörterbuch" and a search bar "Nach einem Begriff suchen". Below this, there's a section titled "Sammlung" with a definition: "Substantiv, feminin [die] 1. das Sammeln (1c) 'die Sammlung [er]brachte, ergab eine stattliche Summe'" and "2a Gesamtheit gesammelter (1b) Gegenstände 'eine bedeutende, reiche, reichhaltige, kostbare, wertvolle Sammlung [von Gemälden, Münzen, Waffen]'". There are also sections for "Ähnlich" with terms like "Hilfsaktion", "Sammelaktion", "Spendenaktion", "Anhäufung", "Ansammlung", "Auslese", "Auswahl", "Kollektion", and "Sortiment". Below these are links to "de.wikipedia.org", "www.dwds.de", and "www.welt.de".

Abb. 23: Google-Suche mit 95.300.000 Ergebnissen (Screenshot, 19.03.2021)

This screenshot shows the same Google search results page for "Sammlung" as in Abb. 23, but on a different page. The search bar and sidebar are identical. The main results area now displays "Seite 47 von ungefähr 464 Ergebnissen (1,68 Sekunden)". The results listed are: "Leihgaben: Wenn private Kunstsammler mehr besitzen als ...", "Theatralische Sammlung", "Staatliche Kunstsammlungen Dresden - Auf der ... - Facebook", "Vollständige Sammlung aller in den Jahren 1816, 1817, 1818 ...", and "Staatliche Kunstsammlungen Dresden - Auf der ... - Facebook". The results from the previous page have disappeared.

Abb. 24: Die selbe Google-Suche zeigt auf Seite 47 plötzlich nur noch 464 Suchergebnisse (Screenshot, 19.03.2021)

Eine bekannte Internet-Weisheit, die in Bezug auf Suchmaschinenoptimierung oft zitiert wird, lautet: „The best place to hide a dead body is page 2 of Google.“ (→ Abb. 25) Was auf der zweiten Seite von Googles Suchergebnissen steht, wird statistisch gesehen kaum betrachtet und ist damit faktisch vom Algorithmus versteckt, was noch weiter hinten steht, wird gar nicht mehr gesehen. Um an eines der 95.300.000 Suchergebnisse zu kommen, müsste man die Suchanfrage genauer eingrenzen, dazu müsste man aber genau wissen, in wel-

cher Form. Kurz, man befindet sich bereits mitten im von Aleida Assmann beschriebenen Dilemma, dass die Algorithmen darüber entscheiden, was gefunden werden kann und was nicht (Assmann, 2016, S. 216).

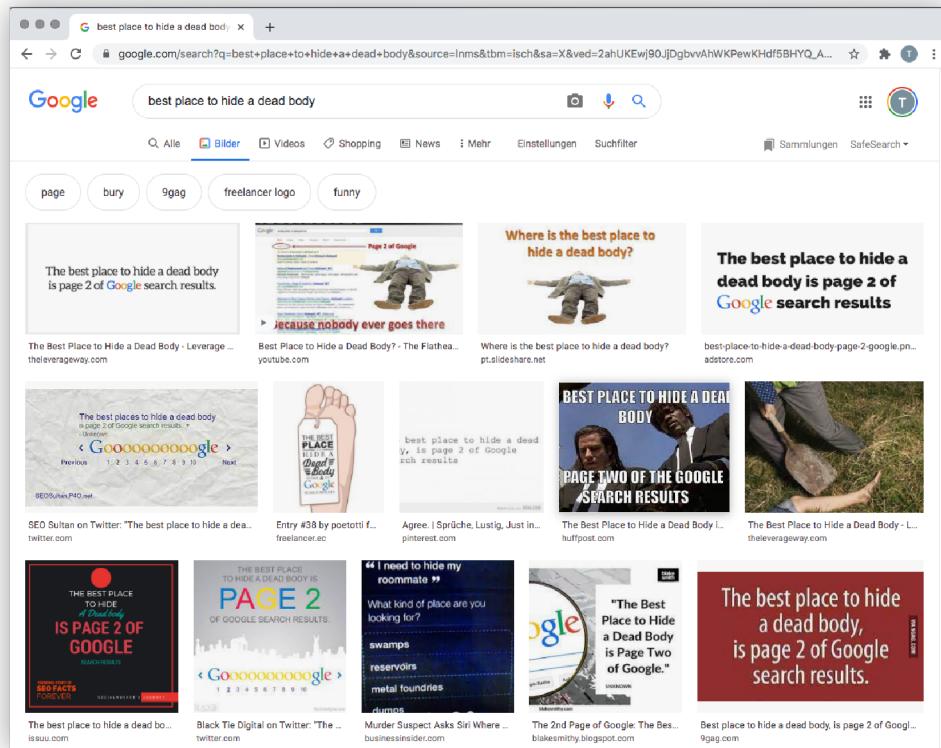


Abb. 25: Google-Suche „best place to hide a dead body“ (Screenshot, 19.03.2021)

Auch in den sozialen Medien ist man zunehmend mit dem Problem mangelnder Auffindbarkeit konfrontiert. Die Netzkünstlerin **Giselle Beiguelman** bringt die Situation auf den Punkt: „There have never been so many records, and yet it has never been so difficult to access our recent past.“ (Grau/Hoth/Wandl-Vogt (Hrsg.), 2019, S. 119) War es in den Anfangstagen dieser Plattformen noch problemlos möglich bestimmte Inhalte wieder aufzufinden, die man zuvor gesehen hatte, so wird dies heute nicht nur von der wachsenden Menge an Inhalten erschwert, sondern wiederum von den Algorithmen: Die Funktionsweise der Plattformen hat sich dahin verlagert, den Nutzer\*innen einen nie enden wollenden Stream neuer Inhalte zu präsentieren, um sie so möglichst lange bei Laune und auf der eigenen Plattform zu halten. Ältere Inhalte spielen in diesem Szenario keine tragende Rolle mehr.

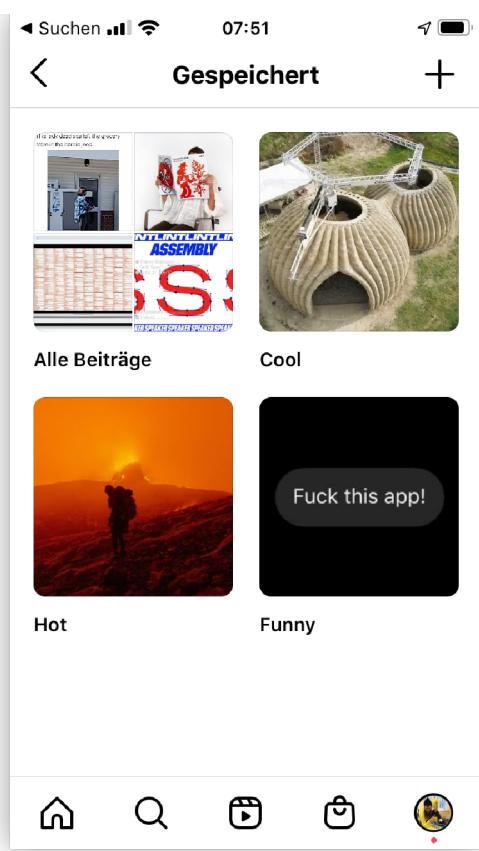


Abb. 26: Instagram-Collections  
(Screenshot, 19.04.2021)

Zwar bieten manche Tools, wie beispielsweise Instagram eine Speicher-Funktion, mit der man Inhalte in Sammlungen ablegen kann ([→ Abb. 26](#)), diese Funktion erlaubt das Speichern aber nicht für alle Inhalte und man hat im Ergebnis zudem eine Sammlung an Sammlungen, die einzig innerhalb der Instagram-Infrastruktur betrachtet und genutzt werden kann, es gibt keine eingebaute Möglichkeit des Exports. Vor allem aber sind eben Inhalte, die ich vorab nicht zufällig so gespeichert habe, nach wie vor nicht einfach auffindbar, es gibt also kein spontanes Zurück. Beigelman weist auf dieses Problem hin und setzt es ebenfalls in Bezug zum individuellen und kulturellen Gedächtnis:

„One of the adages of the contemporary world is „the Internet does not forget.“ But social media does not *let us* remember. The information architecture for these floating data spaces does not favor retrospective queries. This does not mean that the data is not there. On the contrary, it is. The data is simply not accessible by search engines. [...] This amount of data production all day, every day is part of our effective memory and this memory is becoming something that we cannot retrieve anymore.“

(Grau/Hoth/Wandl-Vogt (Hrsg.), 2019, S. 88)

Sie stellt fest, dass unsere Erinnerung zum „corporate issue“ (ebd.) wird, da wir ohne effektive Gegenmaßnahmen gänzlich der kommerzi-

ellen Funktionslogik sozialer Netzwerke und anderer Plattformen ausgeliefert sind. Selbst innerhalb meiner eigenen Profile auf diesen Plattformen hält sich diese Logik aufrecht: Theoretisch sollte ich an dieser Stelle Zugriff auf alle von mir erstellten Inhalte haben, aber auch hier übernimmt der Algorithmus die Narration. In Bildverwaltungsprogrammen werden Bilder nach undurchsichtigen Kriterien zusammengestellt und mir als Videozusammenschnitte ereignisreicher Tage präsentiert, auf Facebook werden unaufgefordert jahrealte Momente hervorgeholt und zum erneuten Teilen auf meiner Timeline vorgeschlagen, „intelligente Suchen“ legen mir bestimmte Suchparameter nahe und präsentieren mir dann wiederum Ergebnisse, denen ich nicht ansehen kann, ob sie in Bezug auf meine Suchkriterien wirklich vollständig sind.

Je größer die Menge an Daten wird, die wir zu verwalten haben, desto verlockender ist es, digitale Entscheidungsarbeit an Algorithmen abzugeben. In dem Moment, wo wir dies tun, verlieren wir aber Zugriff auf all jene Daten und Erinnerungen, die durch das algorithmische Raster fallen und daher im Dunkeln dieser Dark Archives verborgen bleiben. Der Initiator des zuvor erwähnten Archive Teams, **Jason Scott**, spricht sich daher dafür aus, sich mit digitalen Plattformen, die nicht das Exportieren der eigenen Daten ermöglichen, gar nicht erst abzugeben ([vgl. Lialina, 2012](#))<sup>12</sup>.

Olia Lialina formuliert ihre Kritik am algorithmischen Wirken der sozialen Netzwerke noch drastischer: Die digitale Erzählung, die „Story“, die wir innerhalb dieser Netzwerke für andere sichtbar auf unserer Timeline inszenieren, ist nur auf den ersten Blick durch uns selbstbestimmt und damit auch unsere eigene Story. In Wirklichkeit folgt sie den Regeln, Strukturen und Ideen, die das Netzwerk vorgibt und favorisiert, und ist daher, selbst wenn wir sie irgendwie exportieren könnten, nie wirklich unsere eigene Story:

„However, in my opinion, what users of Facebook, for example, would want to export is not the same as what they uploaded. They uploaded data, but want to get back the narration. To export a story is an effort on completely different scale, conceptually and programming wise. Additionally, the story that you thought you were telling [...], is in fact told, designed, scripted, staged by the service. You can hit an export button till you die, it will not become yours.“

([Lialina, 2012](#))<sup>13</sup>

## Networks of Care

Damit die Nutzer\*innen sich aus dieser Abhängigkeit von kommerziellen Funktionsmechanismen lösen können, müssen jene digitalen Inhalte, über die sie selbst bestimmen wollen, außerhalb der kommerziellen Plattformen gespeichert werden, um sie dort, unabhängig von

den Interessen dieser Anbieter selbst verwalten und nutzen zu können. Diese Verlagerung der Inhalte nach außen kann im kleinen Rahmen durch individuelle Bemühungen – wie das Erstellen von Screenshots, das Anlegen von Bookmarks etc. – erfolgen, wie im [Teil 2: Digitales Erinnern und Vergessen](#) bereits ausgeführt. Für das Bewahren größerer Datenbestände hat es sich aber als effektive Strategie erwiesen, sich für dieses Vorhaben zusammenzuschließen und koordiniert zu agieren.

Das zuvor geschilderte Beispiel von GeoCities zeigt anschaulich, wie sehr das Überdauern von Datenbeständen abhängig ist von der Instanz, die sie verwaltet oder von anderen Faktoren, auf die die Nutzer\*innen selbst keinen Einfluss haben. In diesem Beispiel hat Yahoo! kurzerhand entschieden eine riesige Informationssammlung zu löschen, die vielen Nutzer\*innen zu irgendeinem Zeitpunkt relevant erschien. Selbst bei den heute etabliertesten Unternehmen ist also nie mit Sicherheit abzusehen, ob die Daten, die wir ihnen anvertrauen, in ein paar Jahren noch verfügbar sein werden. Sie können entweder – wie in diesem Fall – gänzlich gelöscht werden oder durch andere strategische Entscheidungen ganz oder in Teilen unzugänglich werden: Zuvor öffentlich verfügbare Online-Artikel sind plötzlich nur noch über kostenpflichtige Abonnements abrufbar, Rechteinhaber\*innen bestimmter Medieninhalte lassen Videos sperren, Social-Media-Accounts werden gesperrt und ihre Inhalte damit unzugänglich oder Inhalte werden zensiert, weil sie teils gegen Gesetze, häufiger aber gegen interne Konzernrichtlinien verstossen ([→ Abb. 27](#)). Diese Zensur findet oftmals auch mit größerem zeitlichen Abstand zum ursprünglichen Erstellen der Inhalte statt.



Abb. 27: Sperrhinweis auf YouTube  
(Quelle: <https://www.planetbackpack.de/youtube-laendersperre-umgehen>)

Die Blog-Plattform **Tumblr** geriet 2018 in die Kritik, als sie ihre Richtlinien zu erlaubten Inhalten dahingehend änderte, dass das Zeigen an-

stößiger Inhalte fortan nicht mehr erlaubt war. Als einer der Gründe für diese Änderung wurden **Apples** engen Vorgaben für erlaubte Inhalte innerhalb des **App Store** vermutet, über den auch Tumblr eine App zur Nutzung seiner Plattform vertreibt ([vgl. „Tumblr“, 2021](#)). Die Nutzer\*innen, die dort teils seit Jahren eigene Blogs mit Inhalten betrieben hatten, die zuvor unproblematisch waren und nun aber plötzlich mit den neuen Richtlinien in Konflikt standen, wurden nur zwei Wochen vor dem Löschtermin von dieser Änderung in Kenntnis gesetzt. Die Tumblr-Community identifizierte dabei auch zahlreiche Inhalte, die fälschlicherweise von der Zensur betroffen waren und dadurch schließlich dennoch der Löschaktion zum Opfer fielen ([vgl. ebd.](#)).

Auch in diesem Fall wurde das Archive Team aktiv, um vor der Schließung möglichst viele der betroffenen Inhalte zu retten. Dazu wurden innerhalb kürzester Zeit Scripts und Automatisierungslösungen programmiert, um diese Inhalte zu identifizieren und in großem Stil zu sichern ([vgl. Archive Team, 2021](#)).

Diese Gefährdung von Inhalten durch (firmen-)politische Entscheidungen hat Giselle Beiguelman – in Abgrenzung zur technisch bedingten technologischen Obsoleszenz – als *ideologische Obsoleszenz* bezeichnet (Grau/Hoth/Wandl-Vogt (Hrsg.), 2019, S. 128). Besonders problematisch sei dabei, dass sich ein Großteil digitaler Inhalte auf wenige große Internet-Konzerne konzentriere und diese dann im Alleingang die Spielregeln bestimmen können: „Anyone can participate in them, but only according to the rules prescribed by previously programmed algorithms.“ (ebd., S. 127) Dabei haben beispielsweise unabhängige Apps, die offene Standards nutzen und außerhalb der engen Regeln dieser Plattformen agieren, häufig einen schweren Stand. Sie sind von bestimmten Zertifizierungen ausgeschlossen und können deshalb auf bestimmten Plattformen, beispielsweise auf Apples mobilem Betriebssystem, nicht genutzt werden (vgl. ebd., S. 128).

In dieser digitalen Großlage erweist sich das Vorgehen, welches das Archive Team durch sein Agieren aufzeigt, als ein effektive Gegenstrategie, um sich aus der Abhängigkeit der Großkonzerne zu lösen: Man vertraut von Anfang an nicht darauf, dass Yahoo! und Co. als datenverwaltende Instanzen unsere Daten bei Bedarf schon herausrücken werden, sondern beginnt präventiv damit, die Inhalte im großen Stil nach außen zu tragen. Die Besonderheit des Archive Teams besteht darin, dass sich dieses Kollektiv ausschließlich aus Freiwilligen zusammensetzt, sich dabei komplett online koordiniert und zudem ohne nennenswerte finanzielle Unterstützung agiert.

Annet Dekker bezeichnet solche digitalen Zweckgemeinschaften, die sich der Bewahrung bestimmter Inhalte verschrieben haben, als *Networks of Care*. Sie bilden sich – oft spontan – um einen digitalen Datenbestand herum und sorgen durch ihr Wirken dafür, dass dieser vor dem Verschwinden bewahrt wird. Mit großem Engagement entwickeln

diese Freiwilligen gemeinsam digitale Tools, die bestimmte Archivierungsaufgaben ausführen oder tauschen sich in Online-Foren über das stetige Erweitern der Speicherkapazitäten ihrer Homeserver aus (→ Abb. 28). Fälle wie dieser zeigen, dass sich aus der Not eines drohenden Datenverlusts ein großes produktives Potenzial entfalten kann.



Abb. 28: Ein „Data Hoarder“ teilt ein Bild seines Home-Servers mit 370 TB Speicherkapazität online. (Quelle: <https://www.reddit.com/r/DataHoarder>)

Um ein solches Network of Care auf den Plan zu rufen, muss es sich beim drohenden Datenverlust nicht notwendigerweise um die *Lösung* von Datenbeständen handeln, sondern es kann auch darum gehen, bestimmte nicht zugängliche Daten für die Öffentlichkeit überhaupt erst sichtbar zu machen, also darum, die dunklen Daten eines Dark Archive ans Licht zu bringen. So gibt es zahlreiche Initiativen und Archive, die sich zum Ziel gemacht haben, kulturell relevante Medien zu erhalten und einem größeren Publikum zugänglich zu machen. Die Netzkünstlerin **Cornelia Sollfrank** sieht Projekte wie [UbuWeb](#), [Memory of the World](#), [Monoskop](#) und [aaaaarg.fail](#) neben zahlreichen weiteren in dieser Tradition, betont deren selbstorganisierte, experimentelle und offene Struktur und bezeichnet sie als sogenannte „Schattenbibliotheken“ (vgl. Griesser-Sternscheg/Sternfeld/Ziaja (Hrsg.), 2020, S. 85). Damit meint sie Bibliotheken, die oft aus der Anonymität heraus operieren und der „Idee des radikalen freien Zugangs“ (ebd., S. 89) folgen, somit also auch urheberrechtlich geschützte Werke für alle frei verfügbar machen (vgl. ebd., S. 86). Dabei folgen sie der von Kritiker\*innen des Urheberrechts oft vertretenen und im Internet vielzitierten Ansicht „Copying is not theft“ (→ Abb. 29), der Feststellung also, dass Kopieren kein Diebstahl sei, da nach dem Kopiervorgang niemandem etwas fehle. Auch viele Unterprojekte des Internet Archives beziehen sich beispielsweise auf diese These.



Abb. 29: „Copying Is Not Theft“-Video auf YouTube (Screenshot: 19.04.2021)

Sollfrank verortet diese Projekte – wie auch vergleichbare mit wissenschaftlichen, historischen oder zeitgeschichtlichen Inhalten – im Kontext des [Guerilla Open Access Manifestos](#), das der politische Hacker **Aaron Swartz** im Jahr 2008 veröffentlichte. Im Anschluss an die Grundthese, dass Information Macht sei, ruft er zum Handeln auf, um solche Wissensbestände für alle zugänglich zu machen:

„We need to take information, wherever it is stored, make our copies and share them with the world. We need to take stuff that's out of copyright and add it to the archive. We need to buy secret databases and put them on the Web. We need to download scientific journals and upload them to file sharing networks. We need to fight for Guerilla Open Access.“

(*Swartz, 2008*)<sup>12</sup>

Auch im zeitgeschichtlichen Kontext ergeben sich Situationen, in denen sich spontane Networks of Care bilden können, um so Daten vor dem Verschwinden zu bewahren. So formten sich in den Wirren rund um die Erstürmung des US-Kapitols am 6. Januar 2021 abermals spontan solche digitalen Netzwerke, mit dem Ziel, Livestreams, Videos und Bilder des Ereignisses zu sichern, bevor diese vorhersehbarweise von den Ersteller\*innen wieder von den Plattformen gelöscht werden würden ([vgl. Cole, 2021](#))<sup>13</sup>. Das Statement eines an der Archivierung mitwirkenden Internetnutzers zeugt davon, wie sehr sich die Beteiligten der Dringlichkeit ihrer Bemühungen in diesem Moment bewusst waren und wie sehr sie inzwischen um die Vergänglichkeit von Internetinhalten wissen:

„Literally from the moment I heard this was happening I knew I needed to start backing it up. [...] Especially the livestreams and POV of the people breaking into Capitol. I remember from previous protests and riots just how fast websites will remove content.“

Das so gesicherte Material dient nicht nur der (auch von den Strafverfolgungsbehörden genutzten) unmittelbaren Beweissicherung, sondern prägt auch unser Bild des Ereignisses selbst. Wie bei vielen anderen inzwischen vornehmlich durch Social-Media-Streams dokumentierten Unruhen ist ungewiss, ob wir ein vergleichbares Bild dieser Geschehnisse hätten, wären sie nicht auf diese Weise spontan archiviert worden, sondern nur mittels offizieller Medienkanäle.

Im gleichen politischen Kontext ist auch der Twitter-Account des damaligen US-Präsidenten **Donald Trump** zu nennen, welcher sich im Laufe seiner Präsidentschaft zu einem seiner wichtigsten politischen Werkzeuge und zum Sprachrohr seiner politischen Agenda entwickelte. Kurz vor Ende seiner Präsidentschaft – zwei Tage nach der Erstürmung des Kapitols – wurde ihm dieser Kommunikationskanal bekanntlich abgedreht: Trumps Twitter-Account wurde gesperrt. Auch wenn Zweck der Sperrung war, Trump an der Verbreitung weiterer wahlbeeinflussender Fehlinformationen zu hindern, so wurde damit gleichzeitig ein wichtiges historisches Zeugnis zensiert, nämlich Trumps Twitter-Timeline mit ihren zahlreichen Tweets, die seine Präsidentschaft bezeugten, aber fortan nicht mehr sichtbar waren. Ruft man heute einen Link zu einem dieser Tweets auf, so findet man nur einen Hinweis, dass der (nicht angezeigte) Tweet von einem gesperrten Account stammt, ohne weiteren Hinweis dazu, um wessen Account es sich dabei handelt (→ Abb. 30).



Abb. 30: Gesperrter Trump-Tweet  
(Screenshot, 11.04.2021)

Allerdings hatten sich auch um diesen Datensatz herum vorausschauenderweise Networks of Care gebildet, die frühzeitig damit begannen, Trumps Tweets zu archivieren, so dass diese heute leicht auffindbar für alle abrufbar sind, beispielsweise im [Trump Twitter Archive](#). Normalerweise ist in den Vereinigten Staaten gesetzlich geregelt, dass die gesamte präsidiale Kommunikation behördlich durch die **National Archive and Records Administration** dokumentiert und archiviert wird. Seit der Obama-Administration werden im Zuge dessen auch Social-Media-Accounts archiviert. Zum Zeitpunkt des neuerlichen Administrationswechsels im Januar 2021 war aber aufgrund der zahlreichen Personalwechsel unter Trump unklar, ob dieser Aufgabe weiter nachgegangen worden war (vgl. [Lörchner, 2021](#)) ↗. Zudem sind die so

von offizieller Seite angelegten Archive nur für bestimmte sehr eingegrenzte Personengruppen einsehbar. Das Wirken der Networks of Care schafft hier also eine Absicherung von öffentlicher Seite her, indem es die Daten selbst bewahrt und sie für alle frei zugänglich macht.

Warum ein solches Handeln unerlässlich ist, liegt gerade in Bezug auf solch intensive ideologischen Grabenkämpfe, wie sie im Zuge von Trumps Präsidentschaft geführt wurden, auf der Hand: In Zeiten von einfachst manipulierbaren Medien und – tatsächlicher oder vermeintlicher – Fake News findet ein permanentes Ringen um Wahrheit und Deutungshoheit statt. Die Kuratorin **Annette Maechtele** hält zur Bedeutung eines fortlaufenden Dokumentierungsvorgangs fest: „Wenn man keine Geschichte schreibt und [...] nicht dokumentiert [...], dann ist alles weg, und das ermöglicht die Verbreitung von Mythen.“ ([Buurman/Wudtke, 2020, S. 3f.](#))<sup>12</sup>

## Archiv als Monument

Nachdem ein Bewusstsein für die Flüchtigkeit digitaler Daten und für die Problematik ihres Verschwindens im Dark Archive geschaffen ist, gilt es abschließend die Frage zu klären, was diese, hoffentlich durch einen solcherart informierten Prozess geschaffenen, Archive repräsentieren und was das für ihr Wirken nach außen bedeutet.

Die enorme gesellschaftliche Bedeutung von Archiven wurde zuvor schon mit Aleida Assmanns Ausführungen aufgezeigt. Je umfangreicher und je öffentlicher das Archiv, desto größer ist sein Einfluss und damit auch die Verantwortung, die mit Erhalt und Betreiben des Archivs einhergeht.

Der Medientheoretiker **Knut Ebeling** betont die Bedeutung staatlicher Archive und erklärt „die Einsehbarkeit der Archive zur Messlatte für moderne Demokratien überhaupt“ (Griesser-Sternscheg/Sternfeld/Ziaja (Hrsg.), 2020, S. 70), die Demokratien verwendeten die Institutionen gesammelter Akten und Gesetze zu ihrer Legitimation (ebd., S. 69). Auch Assmann bekräftigt diesen Zusammenhang:

„Das Erste, was in einem autoritären Staat passiert, ist der Verschluss der Archive. Denn sie sind es, die den Erfahrungs-, Wissens- und Denkraum einer Gesellschaft über die Gegenwart hinaus in die Vergangenheit und Zukunft erweitern.“

([Schülein, 2021](#))<sup>13</sup>

Wer also die Macht über das Anlegen und die Einsehbarkeit eines Archiv hat, bestimmt folglich die Narration, in deren Zusammenhang dieses Archiv steht. Das Archiv steht immer im Dienste einer Stabilisie-

rung des Systems (vgl. Griesser-Sternscheg/Sternfeld/Ziaja (Hrsg.), 2020, S. 124).

Dabei ist es wichtig festzuhalten, dass ein Archiv sich zwar aus der Realität speist, denn was es verwahrt, steht immer in Beziehung zur Wirklichkeit (vgl. ebd., S. 66). Wie von anderen dokumentarischen Formaten aber bekannt, ist auch das Archiv ein Dokumentationswerkzeug, das die dokumentierte Realität nicht einfach wirklichkeitsgetreu wiedergibt, also *reflektiert*, sondern diese durch seine spezifische Ausgestaltung in einer bestimmten Weise *codiert*. Da naturgemäß in einem Archiv immer nur ein Bruchteil der Wirklichkeit gesammelt und damit gezeigt werden kann, es aber dennoch die (vermeintliche) Wirklichkeit repräsentieren soll, bezeichnet Ebeling das Archiv als einen „Realitätsraffer“ (ebd.), mit ihm werde eine Wirklichkeit präsentiert, die „realer als die Realität“ (ebd.) sei. Das Archiv sei damit „nicht Dokument, sondern Monument“ (ebd.).

So ist nachzuvollziehen, dass der Archivierungsprozess nicht nur in die gegenwärtige Realität hinein- und zurückwirkt, sondern eine vergangene Realität oftmals erst konstruiert, sie damit wiederum zur kulturellen Erinnerung werden lässt und somit schließlich Einfluss auf unsere Identität nimmt. Erica Scourti identifiziert diesem Gedanken folgend selbst Suchanfragen als „Auslöser“ spontaner Archive und damit als Vorgänge, die Wirklichkeit konstruieren:

„Perhaps every document creates (rather than describes or illustrates) the event; every search creates an archive, and every archive gives rise to a different reality. Search queries both create an archive and are potentially archival material in themselves (as the still ongoing fascination with Google’s auto-complete attests to) and [...] the archiving itself is productive of events, historical and otherwise.“

(Scourti, 2016) [\[2\]](#)

Die Verantwortung, die mit dem Schaffen und Betreiben eines Archivs einhergeht, hängt damit zusammen, dass dieses in Gegenwart und Zukunft hinein wirkt: Gespeicherte Dokumente werden nicht nur betrachtet, sondern sie werden selbst wiederum zur Quelle für Inspirations-, Recherche- und Lernprozesse (vgl. Grau/Hoth/Wandl-Vogt (Hrsg.), 2019, S. 19).

Es ist also angeraten, immer dann, wenn man mit einem Archiv zu tun hat, dessen Strukturen kritisch zu hinterfragen: Welche Intention steht hinter dem Archiv, welcher Blick soll hier eröffnet werden, welche Narration wird hier erzählt? Erst wenn man diesen Fragestellungen nachgeht, tut sich die Möglichkeit auf, sich von der Perspektive der Archivsteller\*innen zu lösen und einen eigenen Blick auf die Inhalte zu gewinnen.

Auch – und gerade – für Archive, die man selbst sammelt und verwaltet, behält der Hinweis der kritischen Hinterfragung Gültigkeit: Sammelt man dauerhaft nach bestimmten Gesichtspunkten, ohne diese jemals in Frage zu stellen und kritisch zu prüfen, verschließt sich einem das eigene Archiv, indem es nur die eine Lesart zulässt, in der es von Anfang an konzipiert war. Die Museologin **Susan Kamel** verweist darauf, dass beispielsweise im Kunstkontext viele Beteiligte oftmals unreflektiert und kanonisch nur „sich selbst sammeln“, bezogen auf Museen in Deutschland also beispielsweise nur „weiße“, westliche Kunst (vgl. Griesser-Sternscheg/Sternfeld/Ziaja (Hrsg.), 2020, S. 133). Es gebe aber im Umfeld dieser Institutionen zunehmend Bestrebungen, dem entgegenzuwirken:

„Es gibt viele Museen, die sich mit Sammlungen anlegen, viele, die Sammlungen ablegen oder gar nicht erst angelegt haben. Es gibt sie, die Museen, die ihre Sammlungen zusammenlegen, und einige mit Gegensammlungen; Archive und Sammlungen, die sich gegen Museen wehren, um sich der Macht der Institution nicht beugen zu müssen.“

(Griesser-Sternscheg/Sternfeld/Ziaja (Hrsg.), 2020, S. 133)

Der Kurator **Marcelo Rezende** plädiert deshalb dafür, sich beim Sammeln und Verwalten der Sammlung bisweilen von einer „produktiven Unbestimmtheit“ und vom Zufall leiten zu lassen statt vom institutionellen Diskurs (vgl. ebd., S. 125ff.). Nur so eröffne sich auch die Möglichkeit zu einer produktiven De-Stabilisierung des Systems und zum Aufbrechen des Narrativs. Das wiederum ist von enormer Wichtigkeit, denn dieses Vorgehen bezeugt die Existenz dessen und derer, die normalerweise aus dem Narrativ ausgeschlossen sind (vgl. ebd., S. 175).

Digitale Archive bringen im Wesentlichen alle Voraussetzungen mit, um im Sinne dieser Unbestimmtheit zu funktionieren. Die Kuratorin **Anne-Marie Duguet** beschreibt die Vorteile des digitalen Archivs gegenüber dem herkömmlichen: „no dust, no smell, no specific place, building, basement, no definitive format. By nature it can be constantly transformed, extended, indefinitely recombined and it can migrate to various kinds of supports.“ (Grau/Hoth/Wandl-Vogt (Hrsg.), 2019, S. 74) Der Medientheoretiker **Oliver Grau** bezeichnet ein solches Vorgehen als digitale *ars combinatoria*, die durch jede neue Betrachtung und jede Neukonfiguration neue Verknüpfungspunkte und Inspirationsmomente entstehen lässt (ebd., S. 83). Auch Annet Dekker hebt diese Beweglichkeit als charakteristisches Merkmal des Archivs hervor: „[T]he default state of the digital archive is re-use instead of storage, circulation rather than centrally organised memory, constant change vs. stasis.“ (ebd., S. 135) Das Versprechen dieser Wandelbarkeit und des endlosen Neu-Arrangierens erfüllt sich aber nur, wenn das digitale Archiv offen gestaltet ist und daher eine solche Arbeitsweise zulässt.

Im Realitätsabgleich mit bestehenden Archiven, besonders mit Online-Plattformen, stößt man dabei schnell an die Grenzen dieses Prinzips. Wie zuvor erläutert, übernehmen hier in großem Maße die Algorithmen die Regie und sorgen dafür, dass das System sich stabilisiert und die erwähnte Unbestimmtheit verhindert. Dass Entscheidungsfreiheit und Selbstbestimmung der Nutzer\*innen innerhalb der kommerziellen Plattformen gar nicht erst vorgesehen sind, ließ der damalige Google-CEO **Eric Schmidt** 2005 in einer bemerkenswerten Aussage im Rahmen eines Interviews durchblicken:

„When you use Google, do you get more than one answer? Of course you do. Well, that's a bug. [...] We should know what you meant. [...] We should get it exactly right and we should give it to you in your language and we should never be wrong.“

(Ferenstein, 2013) ↗

Mit dieser Aussage deklariert Schmidt kurzerhand die Möglichkeit der freien Wahl zum *Bug* und spricht den Nutzer\*innen jedes Maß an Autonomität ab. Da der Bug aber weiterhin Bestand hat, zeigt sich, dass die Nutzer\*innen ihren Anspruch auf Selbstbestimmung nicht einfach aufgeben werden. Auch Plattformen, die eigentlich den Sammlungs- und Inspirationsprozess unterstützen sollen, bleiben mit ihren algorithmischen Verheißen auf halbem Wege stehen. Wenn man bei **Are.na** oder **Pinterest** den „Empfehlungen“ folgt, den Vorschlägen also, die bereits angelegte Sammlungen ergänzen sollen, so stellt man fest, dass diese Vorschläge häufig eigentlich genau dem entsprechen, was man ohnehin schon gesammelt hat und einen in der Recherche so nicht weiterbringen. Selbst im Amazon-Katalog – zuvor noch als helles Archiv beschrieben – tun sich bisweilen dunkle Ecken auf, wie ein populärer Tweet von 2018 erkennen lässt (→ Abb. 31).



Abb. 31: Tweet abgerufen am 15.04.2021

Aufgrund dieses algorithmischen Feedbacks bleibt man auf sein eigenes Vorstellungsvermögen angewiesen, will man die Sammlungen unbestimmt erweitern und vermeiden „sich selbst zu sammeln“. Wenn selbst der Algorithmus des weltgrößten Online-Händlers mir zum ersten Klositz nur den zweiten vorzuschlagen weiß, zeigt das eindrücklich, dass wir den geschlossenen, kommerziellen Plattformen nicht trauen sollten uns die Entscheidungsarbeit abzunehmen, denn nicht

immer sind algorithmische Fehlentscheidungen so offensichtlich und damit durchschaubar und korrigierbar wie in diesem Fall.

Sammlungen müssen wir selbst anlegen und damit dies frei von algorithmischen Zwängen geschehen kann, ist es notwendig, dass dies innerhalb *offener* Systeme geschieht: Transparent und verständlich in ihrer Funktionsweise, zugänglich für alle involvierten Akteur\*innen, flexibel und niedrigschwellig in der Nutzbarkeit und mit der Möglichkeit ausgestattet, die Daten wieder aus dem Archiv zu exportieren und in andere Kontexte zu überführen. Die Daten müssen sich leicht auffinden, umsortieren und erweitern lassen, um so die oben genannten Strategien der Unbestimmtheit, des Gegensammelns und des Erschließens neuer Lesarten zuzulassen. Das Archiv muss eine dezentrale Organisationsweise erlauben, um so kollaboratives Arbeiten zu ermöglichen und durch partizipatives Sammeln einen Gegenkanon schaffen zu können (vgl. Griesser-Sternscheg/Sternfeld/Ziaja (Hrsg.), 2020, S. 136).

Der Zugewinn an Selbstbestimmung der Nutzer\*innen, der mit dem Wechsel von geschlossenen hin zu offenen Plattformen erfolgt, lässt sich hervorragend an der Entstehung des GeoCities-Archivs nachvollziehen. Olia Lialina sieht im Entreißen der Daten aus Yahoo!s Einflussbereich den Startpunkt eines ergebnisoffenen Arbeitens mit dem Archiv:

„Therein lies a curious paradox: On the one hand, the shutting down of Geocities was no doubt a barbaric act. On the other hand, all the buzz it created and the fact that Geocities was eventually rescued and archived has created this research opportunity to investigate the web as it existed in the 1990s in a more meaningful way than was ever possible before. This means looking at the sheer quantity of web profiles collected in one location, *the freedom of not being dependent on search engine algorithms, the popularity of some of the page's authors [and] their desire to link to other authors [...]*.“

(Lialina, 2012, Hervorhebung T.R.)<sup>12</sup>

Erst wenn es möglich ist, unabhängig von den genannten Einflüssen mit den Daten zu arbeiten, kann dieses Arbeiten außerhalb des ursprünglichen Narrativs produktiv werden und wahrhaft neue Lesarten zulassen. In diesem emanzipativen Moment dürfen wir aber nicht die Verantwortung vergessen, die uns durch das selbstbestimmte Betreiben des Archivs und durch dessen Wirken nach außen zufällt, wie noch einmal mit einer Mahnung Kurt Ebelings betont sei: „Sammlungen und Sammeln sind nicht neutral, nicht unschuldig, nicht nett. Indem man etwas sammelt, repräsentiert man nicht (nur) eine Vergangenheit [...], man codiert eine Zukunft“ (Griesser-Sternscheg/Sternfeld/Ziaja (Hrsg.), 2020, S. 66).

## Fazit: Die Macht der Screenshots

Mit solch großer Verantwortung beladen, ist es für uns als Sammelnde und Archivierende wichtig, trotzdem unsere Handlungsfähigkeit zu erhalten. Was ist also zu tun beim Bewahren, Sammeln und Archivieren von Websites?

Die vom Internet Archive geäußerte Aufforderung „If you see something, save something“ dient als ein guter Startpunkt für den eigenen Umgang mit dem Bewahren von Websites. Das Prinzip ist dabei selbstregulierend: Alles, was ich online gesehen habe, ist zumindest für einen Moment in irgendeiner Weise relevant für mich gewesen und daher auch speichernswert, weil es irgenwann einmal das Bedürfnis geben könnte, zu diesen Inhalten zurückzukehren. Was ich nicht sehe, ist hingegen nicht relevant, und es gibt keinen Grund es zu speichern. Deshalb sind partizipative Formate, bei denen Archivierungsentscheidungen von Nutzer\*innen getroffen werden, deutlich wertvoller als solche, die beispielsweise mittels automatisierter Scripts Listen abarbeiten und in großem Stil Inhalte speichern.

Damit das jedoch überhaupt funktionieren kann, ist es wichtig, dass für das partizipative Mitwirken an Archivierungsvorhaben ein niedrigschwelliger Einstieg besteht, so dass diese Lösungen auch wirklich in der Breite genutzt werden können. Denn auch mit der nützlichsten Archivstruktur ist mir nicht geholfen, wenn der Weg diese zu befüllen so mühsam ist, dass dies dann in der Alltagsnutzung kaum geschieht.

Als ein Beispiel eines solchen Tools sei abermals die Browser-Erweiterung der Wayback Machine genannt: Nach einmaligem Installieren hat man am Rechner jederzeit ein Werkzeug zur Hand, das mit einem Knopfdruck die aktuell betrachtete Website archiviert. Dass dieses Archiv nicht perfekt ist, wurde oben erläutert. Allerdings ist allein durch das Festhalten schon viel gewonnen. Wenn später einmal Links innerhalb der archivierten Website nicht mehr funktionieren sollten, hat man doch immer noch den Inhalt, aus dem sich viele Informationen rekonstruieren lassen.

Ein solches System bewirkt ein wechselseitiges Profitieren: Ich möchte aus bestimmten Gründen eine Seite speichern, das Internet Archive und andere Dienste stellen mir dazu die nötige digitale Infrastruktur zur Verfügung und ich steuere im Gegenzug ein für mich relevantes Stück digitaler Alltagskultur bei. Mit dem Speichern erhalte ich mir auch ein gewisses Maß an Handlungsfähigkeit: Ich bin auf bestimmte Websites – zur Informationsbeschaffung, zur Inspiration etc. – angewiesen und sorge so dafür, dass sie mir weiter verfügbar bleiben.

Ist abzusehen, dass man einmal eine akkurate Aufzeichnung einer Website benötigt, so lohnt es sich, einen etwas aufwendigeren Weg zu gehen und ein Werkzeug wie Conifer zu verwenden, um damit Websites samt ihrer Medieninhalte „aufzunehmen“. Auch bei Zugängen,

die ein Login erfordern, ist diese Methode zu empfehlen. Bei interaktiven Inhalten kann man im Zweifelsfall zunächst ausprobieren, ob die Archivierung in der Wayback Machine ausreichend ist, um die Funktionalität zu erhalten und andernfalls die gleiche Seite noch einmal mit Conifer aufnehmen.

Den beiden genannten Archivierungslösungen ist gemein, dass sie offene Plattformen sind. Sie sind für alle frei nutzbar und entsprechende Lizenzen garantieren, dass die Daten der öffentlichen Archive dauerhaft für alle zugänglich sind. Daten können zudem aus den Archiven herausbefördert werden, entweder über eine öffentlich verfügbare Programmierschnittstelle oder eine Exportmöglichkeit. So ist sichergestellt, dass man stets eigene Untersammlungen und Gegensammlungen außerhalb dieser Plattformen anlegen kann und man nicht innerhalb der Infrastruktur der Plattformen eingeschlossen ist, wie das beispielsweise bei Are.na oder erst recht bei den Sozialen Netzwerken wie Facebook oder Instagram der Fall ist.

Trotz meiner eigenen Nutzung und uneingeschränkten Empfehlung dieser offenen Lösungen führt kein Weg vorbei am effektivsten Archivierungswerkzeug, dem Screenshot. Kein anderes Tool hat eine vergleichbare Verfügbarkeit und keines eine so einfach Nutzung wie er. Dazu bietet er eine pixelgenaue Exaktheit dessen, was ich auf dem Bildschirm sehe. Er ist schnell aufgenommen und nach Erstellen in meinem eigenen digitalen Besitz, so dass ich dann frei darüber verfügen kann, frei von algorithmischen Zwängen oder vorgegebenen Strukturen. Nichtsdestotrotz ist er leicht zu teilen und in Umlauf zu bringen, da es sich schlicht um eine digitale Bilddatei handelt.

Dragan Espenschied sieht in dieser Möglichkeit des Zirkulierens den entscheidenden Vorteil des Screenshots als Archivierungsformat: „[T]o be useful an artifact has to work as a ‚post‘, it has to become impartial and be brought into a format that is accepted everywhere. And that is a screenshot.“ ([Owens, 2014](#))<sup>12</sup>

All diese Eigenschaften des Screenshots zusammengenommen überwiegen den Nachteil der fehlenden Interaktivität häufig bei Weitem. Die oben diskutierten Beispiele zeigen, dass sich auch aus einem jahrealten, statischen Screenshot eine große Menge an Informationen rekonstruieren lässt und man dadurch häufig auch zu den zugehörigen interaktiven Inhalten zurückfinden kann.

Deshalb möchte ich das Vorgehen stark machen, den digitalen Alltag großzügig mit Screenshots festzuhalten. Da dieser Ansatz ohnehin schon weit verbreitet ist, möchte ich dazu anregen, ihn auch als dokumentarische Praxis ernst zu nehmen. Was wir jetzt nebenbei als Alltagsmoment festhalten, ist immer gleich auch Dokument und kann später zum Monument innerhalb eines Kanons werden oder auch einen Gegenkanon bestärken.

Wenn wir Sammlungen anlegen, für uns selbst oder andere, sollten wir dabei stets kritisch im Blick behalten, wie die Sammlung wirkt, ob die Sammlung funktioniert, ob wir sie nicht unnötig eingrenzen oder in eine bestimmte Richtung lenken. Das Wissen um Funktions- und Wirkungsweise sowie um Wirkmacht von Archiven kann uns informierend begleiten, ohne uns dabei im Weg stehen zu müssen. Die Sammel- und Rettungsaktionen im ideologischen Spannungsfeld zeigen, dass es manchmal vor allem schnell gehen muss. Solange wir aber das Format offen halten, bleibt uns auch die Möglichkeit, jederzeit von der Sammlung zurückzutreten, einen neuen Blick auf das Gesammelte zu gewinnen und die Stellschrauben der Sammlung gegebenenfalls nachzusteuern.

Wir wissen nicht, welche Archive und Plattformen die Zukunft bringt. Wenn wir aber in diesem Sinne offen sammeln, sind wir gut darauf vorbereitet.

1. Eine Domain ist der frei wählbare Name, unter dem eine Website erreichbar ist.  
Mit der Bezeichnung beschreibt man in der Regel die Zusammensetzung aus diesem Namen und der Toplevel-Domain, wie z.B. .com oder .de. Die Domain der Website der Klasse Digitale Grafik lautet beispielsweise digitale-grafik.com. ↵
2. Ein Server ist ein spezieller Computer, der rund um die Uhr läuft, und so jederzeit in der Lage ist auf Anfragen zu antworten und dann Dateien bereitzustellen oder Berechnungen durchzuführen, wie dies beispielsweise beim Aufrufen einer Website geschieht. Webserver befinden sich meist in größeren Rechenzentren kommerzieller Anbieter, man kann theoretisch aber auch privat zu Hause einen eigenen Homeserver betreiben, der – in der Regel mit weniger Absicherung gegen Ausfälle – den gleichen Zweck erfüllt. ↵
3. Unter Dokumentation versteht man das Aufzeichnen und Organisieren von Daten rund um die zu dokumentierende Arbeit in möglichst objektiver Form. Dabei kann der Inhalt der Arbeit dokumentiert werden, wie auch die Entstehung, der Kontext und alle anderen Aspekte, die in Verbindung zur Arbeit stehen. (vgl. Dekker, 2020, S. 34) ↵
4. Ein Script ist ein kleines Computerprogramm ohne User Interface, das oft nur aus wenigen Zeilen Programmcode besteht und meist eine ganz bestimmte, sehr eng eingegrenzte Aufgabe ausführt. Scripts können entweder von Nutzer\*innen manuell ausgeführt werden oder beispielsweise auf Servern automatisiert auf verschiedene Ereignisse reagieren oder zu verschiedenen Uhrzeiten laufen. ↵
5. Darüber hinaus sammelt das Internet Archive in Zusammenarbeit mit öffentlichen und privaten Institutionen allerdings auch andere Medien, wie Bücher, Audiodateien, Bilder, Videos etc., mit dem Ziel, diese zu digitalisieren und für die Allgemeinheit im Internet verfügbar zu machen. ↵
6. Als Webcrawler bezeichnet man Scripts, die – ohne User Interface – automatisiert und bestimmten Regeln folgend, zahlreiche Websites besuchen und von diesen bestimmte Informationen bzw. gleich die ganze Website abspeichern und dann gesammelt an einer Stelle, meist auf einem Server, ablegen. Beispielsweise erstellt Google seinen Suchindex mittels Webcrawlern. Websitebetreiber\*innen können auf ihren Websites festlegen, dass sie nicht von solchen Webcrawlern „besucht“ werden wollen. In diesem Fall werden sie dann von den meisten Webcrawlern ignoriert. ↵

7. Eine IP-Adresse ist eine Zahlenabfolge, die als Adresse innerhalb eines Netzwerks funktioniert. Jedem Gerät innerhalb eines Netzwerks – beispielsweise des Internets – ist eine solche eindeutige Zahlenfolge zugeordnet, so dass es über diese Zahlenfolge mit anderen Geräten oder Servern kommunizieren kann. Die IP-Adresse der gerade besprochenen ersten Website lautete beispielsweise 128.141.201.74. ↵
8. Es sei angemerkt, dass sich sämtliche von Annet Dekker erwähnten Überlegungen auf *Net Art* beziehen, wie der Titel ihres Buches „Collecting and Conserving Net Art“ (2020) unschwer erkennen lässt. Die Stellen, die ich von ihr zitiere, lassen sich meines Erachtens aber auch auf nicht-künstlerische Websites und Projekte beziehen, weshalb ich sie hier als Belege für das Archivieren von Websites grundsätzlich anführe. ↵
9. Das Prinzip der *Versionierung* stammt zwar ursprünglich aus der Welt der Software-Entwicklung, ist uns aber allen geläufig, da es über Formulierungen wie *Web 2.0* oder *Industrie 4.0* längst Einzug in unseren alltäglichen Sprachgebrauch gehalten hat. ↵
10. Ein Umstand, den sich geschäftstüchtige Internetnutzer\*innen zunutze machen wollten: Angesichts der Tatsache, dass der Nachfolgefim Space Jam 2 im Sommer 2021 in die Kinos kommt, haben sie sich die Domain [spacejam2.com](http://spacejam2.com) gesichert, unter der sie nicht nur ihre „Weltraum-Marmelade“ („Space Jam“) verkaufen, sondern für entsprechend zahlungswillige Kundschaft auch gleich die Domain selbst, in der Hoffnung, dass Warner Bros. Pictures den neuen Film nicht einfach unter [spacejam.com](http://spacejam.com) vermarkten würde, da dazu die Original-Website geändert werden müsste. Die Hoffnung hat sich inzwischen als vergebens erwiesen, da Warner Bros. Pictures mittlerweile auf der Hauptdomain [spacejam.com](http://spacejam.com) den kommenden Nachfolgefim bewirbt und die Original-Website auf die Unter-Domain [spacejam.com/1996](http://spacejam.com/1996) verschoben hat. Es bleibt abzuwarten, ob die Website nach der erfolgten Space-Jam-2-Kampagne wieder in den Ursprungszustand zurückversetzt wird. Diese Entscheidung macht so abermals sehr deutlich, dass im Internet nichts ewig währt. ↵
11. Ein Inlineframe ist ein spezielles HTML-Element, das das Einbinden fremder Websites innerhalb einer Website erlaubt. So kann man beispielsweise mithilfe eines Inlineframes eine ganze Website innerhalb eines kleinen Rahmens auf der eigenen Website platzieren. Dabei bleibt die Interaktivität dieser eingebundenen Website komplett erhalten, so dass sie ganz normal nutzbar ist. Manche Websites unterbinden allerdings auch das Einbinden in andere Websites, so dass man nicht jede beliebige Website auf seiner eigenen Website platzieren kann. ↵
12. Der Erstellungszeitpunkt eines Screenshots lässt sich allerdings aus dessen Dateinamen ableSEN, wenn man sich die Mühe macht, diesen mithilfe der entsprechenden Browserwerkzeuge abzufragen ↵
13. „Scraping“ (englisch *scrape*: abschaben) bezeichnet das automatisierte Herunterladen bestimmter oder aller Inhalte von einer oder mehreren Websites. So kann man mittels eines Scripts beispielsweise alle Bilder einer bestimmten Auswahl an Websites „scrapen“, um sie so auf seiner Festplatte zu sammeln. ↵
14. Beim Recherchieren für diese Masterthesis ist mir erst aufgefallen, wie schwierig es ist, Bildmaterial alter Websites in ihrer Umgebung – also beispielsweise auf einem alten Röhrenmonitor – zu finden. Das ist also eine Leerstelle, die ein solches Archiv zu füllen imstande wäre. ↵
15. 🎉 ↵
16. Ein Torrent ist ein sogenanntes Filesharing-Protokoll, das es erlaubt große Dateien schnell unter zahlreichen Nutzer\*innen zu verteilen. Die Dateien liegen dabei im Regelfall auf den Rechnern der Nutzer\*innen und werden nach und nach von diesen heruntergeladen. Diese Verteilung funktioniert deshalb so schnell, weil die Technik dafür sorgt, dass die entsprechenden Dateien – oder Bestandteile davon – auch parallel von den Rechnern verschiedener Nutzer\*innen heruntergeladen werden können. Im Prozess des Herunterladens stellt man die Dateien

dabei auch gleichzeitig den anderen zur Verfügung. Bekanntheit erlangte diese Technik vor allem im Rahmen illegaler digitaler Tauschbörsen Ende der 1990er-Jahre. Sie wird heute aber auch vielfach als rechtlich völlig einwandfreie Methode eingesetzt, wenn es darum geht, große Dateien zu verteilen. ↪

## Literatur

- ASSMANN, ALEIDA: *Erinnerungsräume. Formen und Wandlungen des kulturellen Gedächtnisses*. 5. Auflage. München: C. H. Beck, 2010.
- ASSMANN, ALEIDA: *Formen des Vergessens*. Göttingen: Wallstein, 2016.
- BAUDRILLARD, JEAN: *Das System der Dinge. Über unser Verhältnis zu den alltäglichen Gegenständen*. 3. Auflage. Frankfurt / Main: Campus, 2007.
- BENJAMIN, WALTER: *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit*. Frankfurt am Main: Suhrkamp, 1977.
- DEKKER, ANNET: *Collecting and Conserving Net Art. Moving Beyond Conventional Methods*. Oxfordshire: Routledge, 2020.
- GRAU, OLIVER (Hrsg.) / HOTH, JANINA (Hrsg.) / WANDL-VOGT, EVELINE (Hrsg.): *Digital Art through the Looking Glass. New strategies for archiving, collecting and preserving in Digital Humanities*. Krems: Edition Donau-Universität, 2019.
- GRIESSESTERMSCHEG, MARTINA (Hrsg.) / STERNFELD, NORA (Hrsg.) / ZIAJA, LUISA (Hrsg.): *Sich mit Sammlungen anlegen. Gemeinsame Dinge und alternative Archive*. Berlin: De Gruyter, 2020.
- JUNGE, KAY (Hrsg.) / SUBER, DANIEL (Hrsg.) / GERBER, GEROLD (Hrsg.): *Erleben, Erleiden, Erfahren. Die Konstitution sozialen Sinns jenseits instrumenteller Vernunft*. Bielefeld: transcript-Verlag, 2008.
- KITTNER, ALMA-ELISA: *Visuelle Autobiographien. Sammeln als Selbstentwurf bei Hannah Höch, Sophie Calle und Annette Messager*. Bielefeld: transcript, 2009.
- POMIAN, KRZYSZTOF: *Der Ursprung des Museums. Vom Sammeln*. 4. Auflage. Berlin: Wagenbach, 2013.
- SOMMER, MANFRED: *Sammeln. Ein philosophischer Versuch*. 2. Auflage. Frankfurt am Main: Suhrkamp, 2018.
- STALDER, FELIX: *Kultur der Digitalität*. 4. Auflage. Berlin: Suhrkamp, 2019.

## Onlinequellen

- ARCHIVETEAM WIKI CONTRIBUTORS (2020). *GeoCities - Archiveteam*. Abgerufen am 20.03.2021, von <https://wiki.archiveteam.org/index.php?title=GeoCities&oldid=45435>.
- ARCHIVETEAM WIKI CONTRIBUTORS (2019). *We are going to rescue your shit*. Abgerufen am 20.03.2021, von [https://wiki.archiveteam.org/index.php?title=We\\_are\\_going\\_to\\_rescue\\_your\\_shit&redirect=no](https://wiki.archiveteam.org/index.php?title=We_are_going_to_rescue_your_shit&redirect=no).
- ARCHIVETEAM WIKI CONTRIBUTORS (2021). *Tumblr - Archiveteam*. Abgerufen am 20.03.2021, von <https://wiki.archiveteam.org/index.php/Tumblr>.
- GEOCITIES RESEARCH INSTITUTE (2011). *About | One Terabyte of Kilobyte Age*. Abgerufen am 20.03.2021, von <https://blog.geocities.institute/about>.
- DER SPIEGEL (o. J.). *Schalten Sie bitte den Adblocker ab!*. Abgerufen am 16.02.2021, von <https://www.spiegel.de/dienste/spiegel-online-schalten-sie-den-adblocker-bitte-ab-a-888158.html>.
- ALEXA (2021). *Alexa-Ranking*. Abgerufen am 16.01.2021, von <https://www.alexa.com/topsites/countries/DE>.

- ARE.NA (o. J.). *About*. Abgerufen am 17.03.2021, von <https://www.are.na/about>.
- LIMA AMSTERDAM (o. J.). *ArtHost*. Abgerufen am 16.03.2021, von <https://www.li-ma.nl/lima/article/arthost>.
- RHIZOME (o. J.). *About*. Abgerufen am 14.04.2021, von [https://conifer.rhizome.org/\\_faq](https://conifer.rhizome.org/_faq).
- INTERNET ARCHIVE (o. J.). *About the Internet Archive*. Abgerufen am 03.03.2021, von <https://archive.org/about>.
- DER SPIEGEL (2020). *Was der neue Kommentarbereich für unsere Nutzerinnen und Nutzer bedeutet*. Abgerufen am 16.02.2021, von <https://www.spiegel.de/dienste/in-eigener-sache-informationen-zur-einstellung-des-spon-forums-a-1303596.html>.
- RHIZOME (o. J.). *Rhizome About*. Abgerufen am 14.04.2021, von <https://rhizome.org/about>.
- MEEDIA (2020). *Alles neu beim „Spiegel“ im Web – so will Chefredakteur Klusmann neue Digitalabonnenten gewinnen*. Abgerufen am 16.02.2021, von <https://meedia.de/2020/01/08/alles-neu-beim-spiegel-im-web-so-will-chefredakteur-klusmann-neue-digitalabonnenten-gewinnen/>.
- BROSKOSKI, CHARLES (2019). *Creating the New Social Web*. Abgerufen am 18.04.2021, von <https://www.youtube.com/watch?v=q1K-OTILIVc>.
- BUURMAN, NANNE / WUDTKE, INA (2020). *Er-Haltung und das Archiv als Are-na*. Abgerufen am 14.12.2020, von [https://documenta-studien.de/media/1/documenta\\_studien\\_10\\_CapriCare\\_DE\\_2.pdf](https://documenta-studien.de/media/1/documenta_studien_10_CapriCare_DE_2.pdf).
- COLE, SAMANTHA (2021). *Archivists Are Preserving Capitol Hill Riot Livestreams Before They're Deleted*. Abgerufen am 02.04.2021, von <https://www.vice.com/en/article/3an5e3/archivists-are-preserving-capitol-hill-riot-livestreams-before-theyre-deleted>.
- FERENSTEIN, GREGORY (2013). *An Old Eric Schmidt Interview Reveals Google's End-Game For Search And Competition*. Abgerufen am 07.04.2021, von <https://techcrunch.com/2013/01/04/an-old-eric-schmidt-interview-reveals-googles-end-game-for-search-and-competition/?guccounter=1>.
- GOTTHARDT, ALEXXA (2018). *Creatives Are Flocking to This Artist-Designed Social Network*. Abgerufen am 17.03.2021, von <https://www.artsy.net/article/artsy-editorial-creatives-flocking-artist-designed-social-network>.
- JESSEN, JENICA (2020). *Looking Back on 2020 - Internet Archive Blogs*. Abgerufen am 03.03.2021, von <http://blog.archive.org/2020/12/19/looking-back-on-2020/>.
- KOERBIN, PAUL (2013). *What is the oldest website?*. Abgerufen am 04.03.2021, von <https://www.nla.gov.au/stories/blog/web-archiving/2013/05/03/what-is-the-oldest-website>.
- KOHOUT, ANNEKATHRIN (2020). *Screenshots – Die Knipserfotografie des Internets*. Abgerufen am 26.03.2021, von <https://sofrischsogut.com/2020/09/01/screenshots-die-knipserfotografie-des-internets/>.
- KREYMER, ILYA / ESPENSCHIED, DRAGAN (2019). *Self-Hosted Archival Embeds*. Abgerufen am 18.03.2021, von <https://blog.conifer.rhizome.org/2019/11/06/self-hosted-archival-embeds.html>.
- KRIESEL, DAVID (2016). *SpiegelMining – Reverse Engineering von Spiegel-Online*. Abgerufen am 17.02.2021, von [https://media.ccc.de/v/33c3-7912-spiegelmining\\_reverse\\_engineering\\_von\\_spiegel-online](https://media.ccc.de/v/33c3-7912-spiegelmining_reverse_engineering_von_spiegel-online).
- LIALINA, OLIA (2012). *There is no Export Function*. Abgerufen am 23.03.2021, von <https://blog.geocities.institute/archives/3647>.
- LIALINA, OLIA (2012). *Still There*. Abgerufen am 09.03.2021, von <http://contemporary-home-computing.org/still-there/geocities.html>.

- LIALINA, OLIA (2021). *Documenting the Ruins of the Web*. Abgerufen am 18.03.2021, von <https://thephotographersgallery.org.uk/whats-on/talks-and-events/documenting-ruins-web>.
- LIEBERMANN, DAVID (o. J.). *hallointer.net*. Abgerufen am 10.03.2021, von <http://hallointer.net/info>.
- LÖRCHNER, JASMIN (2021). *Sorry, wir haben den Präsidenten gelöscht*. Abgerufen am 25.01.2021, von <https://www.spiegel.de/geschichte/donald-trumps-twitter-erbe-sorry-wir-haben-den-praesidenten-geloescht-a-6a2bd454-e72c-436b-88b7-68bd2dc392ba>.
- MALINOWSKI, ERIK (2015). *'Space Jam' Forever: The Website That Wouldn't Die*. Abgerufen am 09.03.2021, von <https://www.rollingstone.com/movies/movie-news/space-jam-forever-the-website-that-wouldnt-die-70507/>.
- NOYES, DAN (2013). *First URL active once more*. Abgerufen am 04.03.2021, von <https://first-website.web.cern.ch/first-website/node/16.html>.
- OWENS, TREVOR (2014). *Digital Culture is Mass Culture*. Abgerufen am 16.02.2021, von <https://blogs.loc.gov/thesignal/2014/03/digital-culture-is-mass-culture-an-interview-with-digital-conservator-dragan-espenschied/>.
- ROSSI, ALEXIS (2017). *If You See Something, Save Something*. Abgerufen am 03.03.2021, von <https://blog.archive.org/2017/01/25/see-something-save-something/>.
- SCHÜLEIN, SELMAR (2021). „Was nicht umkopiert wird, geht verloren“. Abgerufen am 16.03.2021, von <https://www.cicero.de/kultur/erinnern-mit-der-digitalisierung-was-nicht-umkopiert-wird-geht-verloren/plus>.
- SCOURTI, ERICA (2016). *Dark Archives – Interview*. Abgerufen am 10.03.2021, von <https://wrongdreams.com/2016/03/16/dark-archives-interview/>.
- SPREEUWENBERG, KIMBERLEY (2012). *Documenting Internet-based Art: The Dullaart-Sakrowski Method*. Abgerufen am 15.03.2021, von [http://aaaan.net/wp-content/uploads/2015/09/Documenting-Internet-Based-Art\\_FINAL.pdf](http://aaaan.net/wp-content/uploads/2015/09/Documenting-Internet-Based-Art_FINAL.pdf).
- SWARTZ, AARON (2008). *Guerilla Open Access Manifesto*. Abgerufen am 07.04.2021, von <https://openaccessmanifesto.wordpress.com/guerilla-open-access-manifesto/>.
- WIKIPEDIA CONTRIBUTORS (2020). *Infoseek*. Abgerufen am 09.03.2021, von <https://en.wikipedia.org/w/index.php?title=Infoseek&oldid=995643256>.
- WIKIPEDIA CONTRIBUTORS (2021). *Tumblr*. Abgerufen am 19.04.2021, von <https://en.wikipedia.org/w/index.php?title=Tumblr&oldid=1018579089>.