

Digital Humanities

Transnationale Printkultur des 19. Jahrhunderts im digitalen Raum

Die Untersuchung von Zeitungen als Daten

Marc Priebe

1. Einleitung

Die Periodikaforschung in den Industrieländern ist aktuell in der glücklichen Lage, auf eine stetig wachsende Anzahl digitalisierter Zeitungen und Zeitschriften zugreifen zu können, denn in den letzten dreißig Jahren, und besonders seit der Jahrtausendwende, haben viele Bibliotheken und Archive ihre Bestände digitalisiert. So wurden Hunderte Millionen Seiten durch Massendigitalisierung verfügbar gemacht, und jeden Monat kommen derzeit mehr als eine Million weitere Seiten hinzu. Vor allem Tageszeitungen, aber auch Mainstream Magazine wie die *Vogue* (siehe Paccaud im Band) und *Esquire* erfreuen sich anhaltender institutioneller Aufmerksamkeit, wie die groß angelegten Digitalisierungsbemühungen in den meisten Industrieländern zeigen. Diese digitalisierten Zeugnisse der Vergangenheit sind zu wesentlichen Ressourcen nicht nur für Forschung und Lehre, sondern auch für Politik, Journalismus und die allgemeine Öffentlichkeit geworden. Doch während sich die Verfügbarkeit digitalisierter Printmedien in der jüngsten Vergangenheit stetig ausgeweitet hat, steckt die Entwicklung digitaler Werkzeuge für eine ebenso kontinuierliche Untersuchung dieses reichhaltigen Materials noch in den Kinderschuhen. So unterliegt die Arbeit mit digitalisiertem Printmaterial – insbesondere Periodika, Fachzeitschriften, Magazinen und Zeitungen – noch immer deutlichen Einschränkungen, obgleich sich hier, gerade im Bereich der Geisteswissenschaften, viele neue Möglichkeiten für Forschungsprojekte bieten (siehe Rössler im Band).

Wie umfassend dokumentiert ist, entstanden die modernen Massenmedien, von denen Zeitungen einen signifikanten Teil ausmachten, im Laufe des 19. Jahrhunderts (siehe Cox im Band).¹ Die technischen Innovationen im Bereich der Druckerpressen und der Telegrafie, das anfängliche Fehlen von Urhebungsrechten und das wachsende Interesse

¹ Weitere Informationen zum US-amerikanischen Zeitungswesen der 1880er Jahre finden sich bei Douglas (1999: 81-94) und Kaplan (2002: 72-103). Zum internationalen Format und Design von Zeitungen, besonders am Ende des 19. Jahrhunderts siehe Barnhurst/Nerone (2001: 68-108).

der Bevölkerungen an Informationen säkularer, sensationeller und politischer Art führten zum Aufbau eines Netzwerks von Nachrichtenbörsen, das von Beginn an transnational war. Im Gegensatz dazu war die Forschung zu Zeitungen traditionell vorwiegend regional oder bestenfalls landesweit ausgerichtet. Ihr lokaler Charakter manifestierte sich als heuristisches Konzept darin, wie Zeitungen gesammelt, digitalisiert und verfügbar gemacht wurden, und zwingt Forschende bis heute zu zeit- und kostenintensiven Recherchen in Bibliotheken und Archiven vor Ort. Weiterhin war die Durchführung komparativer Forschungsprojekte zum Journalismus in verschiedenen Ländern oft von den Sprachkenntnissen der jeweiligen Forschenden und von ihrer engen Kooperation mit Forschenden in den untersuchten Ländern abhängig. Dennoch hat die Schaffung nationaler, isolierter Sammlungen und Bibliotheken die Tatsache verschleiert, dass transnationale Nachrichtenbörsen grundlegend für die Presse des 19. Jahrhunderts waren. Jane Chapman (2016: 184) fordert daher, dass die Forschung zu Periodika des 19. Jahrhunderts danach streben sollte, ihre Beschränkung durch nationale Grenzen und Einsprachigkeit zu überwinden – und digital verfügbares Material kann dabei helfen, die Forschung zu transnationalen Entwicklungen in der Vergangenheit weiter voranzutreiben (siehe Ernst im Band). Aus einer erweiterten Perspektive unterzieht Lara Putnam (2016: 394) die steigende Abhängigkeit von digitalen Recherchen in der Geschichtswissenschaft einer kritischen Betrachtung. Sie weist darauf hin, dass eine Zunahme an digitalen Primärquellen, bspw. Zeitungen und Magazine, nicht notwendigerweise zu einem umfassenderen Verständnis der Vergangenheit führt. Ihrer Argumentation nach reagieren transnationale Studien zwar auf dringende intellektuelle Belange, verlieren aufgrund der leichten Durchführbarkeit digitaler Recherchen jedoch an Wert und verwandeln die Geschichtswissenschaft damit potenziell in eine Schreibtischdisziplin, was die Gefahr oberflächlicher oder gar dilettantischer Forschung in sich birgt. Putnam zufolge sollte die Forschung deshalb weiterhin auf die Archiv- und Bibliotheksarbeit vor Ort zurückgreifen, wann immer dies möglich ist. Dabei vernachlässigt sie allerdings wichtige ökologische und finanzielle Faktoren und unterschätzt zudem, in welchem Maße digitale Sammlungen die Realisierung transnationaler Forschungsvorhaben ermöglichen, für die digitale Sondierungs-, Analyse- und Interpretationsmethoden keine Gefahr, sondern eine Bereicherung darstellen.

Eine der größten Herausforderungen in der digitalen geisteswissenschaftlichen Forschung der letzten dreißig Jahre war die Entwicklung innovativer computergestützter Modelle, Werkzeuge, Software, Datenspeichersysteme und Infrastrukturen, die eine Überwindung der Grenzen traditioneller *analoger* Archive im Hinblick auf geografische Lage und Zugänglichkeit ermöglichen und der Forschung Unterstützung bei der Untersuchung transnationaler Prozesse der Informationsweitergabe und -rezeption über die Printmedien bieten. Solche Untersuchungsmöglichkeiten sind besonders in jenen akademischen Disziplinen willkommen, die sich mit dem 19. Jahrhundert beschäftigen – einer Zeit, in der die Menschen in Europa, seinen Kolonien und Nordamerika nicht nur Zeugen einer wachsenden Anzahl von Nachrichtenquellen und lokalen Zeitungen, sondern auch von technologischen Innovationen wurden, die der Verbreitung der Zeitung als kultureller, politischer und ökonomischer Institution in westlichen Gesellschaften Vorschub leisteten. Eines der Hauptmerkmale der Zeitung des 19. Jahrhunderts war ihre Verwurzelung im »scissors-and-paste journalism« (Beals 2017: 1): In Zeitungsredak-

tionen war es damals gängige Praxis, die eigenen Seiten mit Nachrichten aus anderen Zeitungen zu füllen. Hieraus folgte eine gewisse Homogenisierung der Informationen, mit denen die Menschen weltweit in Kontakt kamen (siehe Barnhurst/Nerone 2001: 102). Diese frühe Form des Copy-Paste-Journalismus kann heute potenziell mithilfe von neu entwickelten Algorithmen aus den Bereichen der digitalen Geisteswissenschaft und der Informatik nachgewiesen werden. Wie David Smith, Ryan Cordell und Abby Mullen (2015) zeigen, können digitale Technologien und Software zu einem besseren Verständnis der Printkultur des 19. Jahrhunderts innerhalb und außerhalb der Vereinigten Staaten beitragen, die sich im Grunde als Reprint-Kultur bezeichnen lässt: Ein Text – ob Nachricht, Gedicht, Witz oder Werbung – konnte damals sozusagen viral gehen, indem er in vielen Zeitungen fast unverändert abgedruckt wurde. Dank der steigenden Verfügbarkeit digitalisierter Versionen von Printmaterialien kann die Forschung nun besser betrachten, wie die zeitungslesende Bevölkerung die Welt, die Ereignisse, über die berichtet wurde, und zugleich auch Leser:innen in anderen Erdteilen wahrnahm. Zudem haben sich die elektronischen Übersetzungstools deutlich verbessert, sodass Forschende nun Textmaterialien aus verschiedenen Sprachräumen (siehe Martin im Band) untersuchen können. Auf diese Weise lassen sich neue Einblicke in Prozesse der Verbreitung und des Austauschs von Informationen, Nachrichten und Kultur innerhalb einzelner Länder sowie zwischen verschiedenen nationalen Gemeinschaften gewinnen.

Um die Anwendungsmöglichkeiten von Textauswertungs- und Data-Mining-Techniken für die Arbeit mit digitalisierten Printmaterialien an einem Beispiel aufzuzeigen, untersucht dieser Beitrag die Berichterstattung zum Ausbruch des Krakatau im Jahr 1883, der oft als erstes globales Medienereignis angesehen wird, aus transnationaler Perspektive. Die Basis hierfür bildet eine vergleichende Analyse der innerhalb der ersten zwei Wochen nach dem Ereignis in Österreich, Mexiko, Deutschland, den Niederlanden, Finnland, Australien, Großbritannien und den Vereinigten Staaten veröffentlichten Zeitungsartikel. Die verschiedenen Archive weisen von der Auswahl der gescannten Zeitungen bis hin zu Unterschieden in der Konvertierung, Aufbewahrung und Präsentation der Zeitungsdaten einige signifikante blinde Flecken auf; dennoch soll hier gezeigt werden, wie sich bestimmte bereits veröffentlichte Annahmen und Ergebnisse, die auf eher traditionellen Untersuchungsmethoden basieren, mithilfe digitaler Werkzeuge und Methoden der Geisteswissenschaften validieren, korrigieren und erweitern lassen. Bevor jedoch anhand der gesammelten Daten herausgearbeitet wird, wie sich die Nachricht über eine Naturkatastrophe weltweit verbreitete, soll es im Folgenden zunächst um einige grundlegende Herausforderungen gehen, die sich bei der Arbeit mit historischen Datensammlungen – und insbesondere mit Tageszeitungen – ergeben.

2. Von analog zu digital: Chancen und Herausforderungen elektronischer Zeitungsarchive

Im Allgemeinen war die Zeitung des 19. Jahrhunderts ein vergleichsweise chaotischer und sich ständig verändernder Mix, der von Fakten bis Fiktionen reichte, die in einer unendlichen Zahl amorpher Arrangements präsentiert wurden. Digitalisierte Zeitungen spiegeln dieses Chaos wider und ergänzen es um bestimmte Eigenheiten, die den heu-

tigen Untersuchungsprozess beeinflussen. Jede Datenbank enthält Bestände an Daten, Metadaten und Bildern, doch die exakte Beschaffenheit und Nuancierung dieser Daten wird häufig durch die automatischen Prozesse verdeckt, mittels derer sie kodiert und erstellt wurden. Da es noch keine universellen Standards für die datenbankübergreifende Analyse gibt, bleibt die digitale Recherche oft auf vorhandene institutionelle oder kommerzielle Bestände beschränkt, statt transnationale Untersuchungen anzuregen. Diese und andere Unregelmäßigkeiten machen die Daten anfällig für Fehlinterpretationen, sowohl in der Rezeption als auch bei der Aktualisierung der Bestände für zukünftige Forschungsprojekte.

Daher drängt James Mussell (2012: 192-202) in seinem Buch *The Nineteenth-Century Press in the Digital Age* darauf, dass Forschende bei der Arbeit mit digitalen Archivmaterialien aus dem 19. Jahrhundert die Transformation einer analogen Quelle – in diesem Fall einer Zeitung – in auf einem Computerbildschirm darstellbare und für die elektronische Weiterverarbeitung geeignete Daten nachvollziehen und kontextualisieren können sollten. Anfänglich entschieden Archive, Bibliotheken und andere Institutionen, ihre Bestände zum Zweck der Konservierung zu digitalisieren, doch meistens sind die Faktoren, die die Digitalisierung beeinflussen – bspw. Entscheidungen im Hinblick auf das Scannen, die Verarbeitung und die Präsentation der digitalen Objekte – den Nutzer:innen einer digitalen Quelle nicht bekannt. Oftmals haben diese Nutzer:innen zudem wenig oder gar keine Informationen darüber, wie groß der digitalisierte Anteil der insgesamt vorhandenen Printmaterialien wirklich ist, geschweige denn, welche Entscheidungen zur Auswahl der zu scannenden Originale geführt haben. Bspw. befinden sich nicht alle Originalzeitungen in Archiven – dies betrifft insbesondere Zeitungen mit mehr als einer Auflage pro Tag –, sodass sie auch nicht in digitalen Sammlungen erscheinen. Weiterhin werden Entscheidungen darüber, welcher Teil der Bestände digitalisiert werden soll, stark von wirtschaftlichen, rechtlichen, regionsspezifischen und praktischen Überlegungen geprägt. Da es nur wenige internationale Standards für den eigentlichen Digitalisierungsprozess gibt, können sich die Ergebnisse unterscheiden, je nachdem, wann, wo und durch wen das analoge Material in den digitalen Raum überführt wurde.² Mussell (2016: 24) liefert ein bemerkenswertes Beispiel: Was Nutzer:innen beim Lesen einer digitalen Quelle vor sich sehen, so schreibt er, sei keine Zeitung aus dem 19. Jahrhundert. Stattdessen handele es sich um eine Abbildung, die in vielen Fällen das gedruckte Medium durch ihre Präsentation und ihre spezifischen Nutzungsmöglichkeiten *simulieren* solle. Unerfahrene halten die digitale Simulation, die sie auf dem Computerbildschirm sehen, jedoch häufig für eine Eins-zu-eins-Abbildung des Originals und begreifen nicht, wie die Umwandlung von Zeitungen und Periodika in elektronische Daten die Art und den Inhalt der durch das Quelldokument vermittelten Informationen verändert.

Eines der größten Hindernisse bei der Arbeit im Kontext der digitalen Geisteswissenschaften, und ganz besonders im Zusammenhang mit älteren Printmaterialien wie Zeitungen und Zeitschriften des 19. Jahrhunderts, liegt in der Software: Die Qualität der optischen Zeichenerkennung (»optical character recognition«, OCR), mit der aus dem

² Zur Konservierung von Zeitungen siehe Bourke (1990); Silverman (2005); Walravens (2006); zur Digitalisierung von Zeitungen siehe King (2005) und zu Metadatenstandards in der Forschung zu digitalisierten Zeitungen Wilkinson (2016).

Originaldokument ein digitales Surrogat erzeugt wird und die eine exakte Umwandlung des gedruckten Zeichens in ein maschinenlesbares Objekt gewährleisten soll, ist häufig unzureichend. Damit der Text einer gedruckten Quelle digital gelesen, durchsucht und bearbeitet werden kann, muss das Bild des gescannten Dokuments in ein Textformat umgewandelt werden. Dieser automatisierte Prozess führt bei den Daten häufig zu Verrauschungen, die wiederum ins Gewicht fallen können, wenn digitalisierte Dokumente weiterverarbeitet und analysiert werden sollen. Die automatische Texterkennung in Dokumentencans ist wegen des komplexen Spaltenlayouts der Zeitungen und aufgrund von typografischen Elementen, bspw. ungewöhnlichen Schriftarten oder Tintenflecken, oft nur teilweise erfolgreich (siehe Podewski im Band). Somit führen Fehler aus dem Scanprozess durch eine fehlerhafte OCR-Auslesung des Originals manchmal zu signifikanten Unterschieden zwischen analogem Original und digitaler Kopie und behindern die Maschinenlesbarkeit von Objekten im digitalen Archiv. Johan Jarlbrink und Pelle Snickars (2017: 1231) wiesen im Rahmen einer Qualitätsprüfung schwedischer digitalisierter Zeitungsbestände durch OCR produzierte Fehler nach. Es wurde eine Genauigkeit von 50 bis 70 Prozent auf Wortebene und von 71 bis 98 Prozent auf Zeichenebene festgestellt. Für die geisteswissenschaftliche Forschung ist ein Verständnis dieser und anderer technischer Prozesse wichtig, da diese an sich oft schon eine Interpretationsinstanz darstellen: Sie beeinflussen die angezeigten Ergebnisse, wenn ein Wort oder eine Reihe von Wörtern in die Suchmaske eines Archivs eingegeben wird, das möglicherweise Millionen fehlerhaft gescannter und konvertierter Objekte enthält.

Vor diesem Hintergrund ist es wichtig, bei der Entwicklung spezifischer methodischer Prozesse für die Analyse der transnationalen Informationsflüsse im 19. Jahrhundert über Staatsgrenzen hinweg die Tatsache zu berücksichtigen, dass Zeitungsarchive früher von lokalen Bibliotheken oder bestenfalls von Nationalbibliotheken digitalisiert wurden. Des Weiteren ist es für die Arbeit mit digitalen Archiven nahezu unerlässlich, dass Forschende sich entweder entsprechendes Computerwissen aneignen oder mit Informatiker:innen zusammenarbeiten, damit computergestützte Prozesse zur Optimierung verschiedener Schritte in der Verarbeitung historischer Zeitungsdaten entwickelt werden können. Dabei kann es um eine sichere Erkennung von Eigennamen, die Artikel-Segmentierung der Zeitungsseiten, die Trennung von Text und Bild, die Ordnungskriterien für Kategorien (bspw. Nachrichten, Werbung, Fortsetzungsromane), die Darstellung und Vervollständigung der Artikelmetadaten, die Modellierung von Daten für bestimmte Zeitspannen oder um die Unterstützung von manuellen Suchen nach relevanten Artikeln in großen digitalen Archiven bzw. sogar archivübergreifenden Suchen gehen.

3. Recherchen und Data-Mining in digitalisierten Tageszeitungen des 19. Jahrhunderts: Der Ausbruch des Krakatau (1883) in der internationalen Presse

Um zu illustrieren, wie diese neuen Möglichkeiten und Herausforderungen die Forschung zum transnationalen Austausch von Informationen durch das Medium Zeitung beeinflussen, soll im Folgenden die Berichterstattung zum Ausbruch der indonesischen

Vulkaninsel Krakatau am 27. August 1883 näher betrachtet werden. Hierzu wird eine vergleichende Analyse von Nachrichtenzyklen aus mehreren Ländern zu dem Vulkanausbruch und seinen Folgen in sieben digitalisierten Zeitungsarchiven vorgenommen. Auf diese Weise lässt sich besser verstehen, wie Menschen in verschiedenen Ländern Nachrichten über ein Ereignis erhielten, das für viele von ihnen Tausende Kilometer entfernt stattfand.³ Zur Untersuchung der geotemporalen Nachrichtenverbreitung und zum Vergleich bestimmter Elemente der Berichterstattung wurden automatische Reprint-Erkennungsalgorithmen und maßgeschneiderte Visualisierungssoftware verwendet. Die Einbeziehung solcher computergestützter Untersuchungsmethoden kann, wie in diesem Beitrag gezeigt werden soll, zu einem genaueren Verständnis für verschiedene politische, kulturelle, ideologische und ökonomische Hintergründe der Nachrichtenberichterstattung im späten 19. Jahrhundert beitragen.

Die ersten Anzeichen für einen bevorstehenden großen Ausbruch des Krakatau gab es – auch wenn sie nicht als solche erkannt wurden – am 10. Mai 1883, als eine vergleichsweise geringe vulkanische Aktivität festgestellt wurde. Da der Vulkan Teil der niederländischen Kolonie Java war, berichteten niederländische Zeitungen über diese ersten Eruptionen, während sich in den digitalen Zeitungsarchiven anderer Länder nichts darüber findet. Der Hauptausbruch begann am 26. August 1883 um 13:06 Uhr Ortszeit. Er dauerte mehrere Stunden an und begann mit einer Serie verheerender Explosionen, die Aschewolken erzeugten. Nach 20:00 Uhr Ortszeit erreichten Tsunamiwellen Höhen von 15 bis 40 Metern und vernichteten zahlreiche Küstendorfer im Westen Indonesiens. Die Explosionen waren an Orten in bis zu 3.000 Kilometern Entfernung zu hören und damit die lautesten jemals gemessenen Geräusche (siehe Winchester 2003: 235, 262). In den Morgenstunden des 27. August führten drei finale Explosionen, größer und lauter als die vorherigen, zum partiellen Einsturz des Krakatau und verursachten weitere zerstörerische Tsunamis. Seinen Höhepunkt erreichte der Ausbruch um 10:02 Uhr. Um diese Zeit hatte die Wolke aus Asche, Feuer und Bimsstein eine Höhe von etwa 40 km erreicht, und die ganze Region war für mehrere Stunden in Dunkelheit getaucht. Dem finalen Ausbruch am Montag, dem 27. August, folgten noch kleinere Explosionen, bis der Krakatau schließlich wieder schwieg und das Tageslicht in den Morgenstunden des 28. August zurückkehrte. Nach offiziellen Angaben kamen 36.417 Menschen ums Leben.⁴

Betrachtet man lediglich die Anzahl der Todesopfer, war der Ausbruch von Krakatau sicher nicht die verheerendste Katastrophe dieser Zeit. Doch im Gegensatz zu allen anderen bis dahin aufgezeichneten Vulkanausbrüchen wurden die Auswirkungen in nahezu allen Erdregionen gespürt, gesehen oder gehört – sowohl unmittelbar als auch noch Monate später. Mit kurz zuvor entwickelten wissenschaftlichen Geräten konnte man damals Druckwellen messen, die mehrfach um die Erde gegangen waren; Auswirkungen

3 Aufgrund verschiedener rechtlicher, eigentumsrechtlicher und praktischer Restriktionen ist eine Untersuchung aller digitalisierten Zeitungen noch nicht möglich. Das hier vorgestellte Forschungsprojekt nutzte vorrangig Daten aus den folgenden Quellen: Australiens Trove, Zeitungsarchiv der British Library, Chronicling America, Europeana, Hemeroteca Nacional Digital de México, finnische Nationalbibliothek, niederländische Nationalbibliothek.

4 Für weitere Berichte über das Wie und Warum des Krakatau-Ausbruchs vgl. Simkin/Fiske (1983); Thornton (1996).

der Tsunamis, die auf den Ausbruch um 10:02 Uhr folgten, wurden im britischen Ärmelkanal festgestellt, und in den folgenden Monaten wurden durch die Ausbreitung vulkanischer Partikel in der Atmosphäre überall auf der Welt ungewöhnliche Wetterphänomene und zahlreiche optische Effekte bemerkt, die Forschung, Literatur und Malerei inspirierten.⁵ Der globale Informationsfluss über den Vulkanausbruch ermöglichte dem Zeitungspublikum zu erkennen, dass lokale Ereignisse tatsächlich Auswirkungen auf der ganzen Welt haben können. Bspw. wurden vielen Menschen globale Verbindungen bewusst, als sie darüber lasen, wie die Wissenschaft Zusammenhänge zwischen den ungewöhnlichen Wetterphänomenen zu erkennen begann. Des Weiteren lernte die Leser:innenschaft nicht nur neue und ›exotische‹ Namen wie ›Sumatra‹, ›Java‹ oder ›Krakatau‹ kennen und erfuhr, welche Orte sich dahinter verbargen, sondern erlebte auch aus erster Hand, dass Ereignisse ›dort drüben‹ beeinflussen können, was ›hier bei uns‹ geschieht.

Bei der Untersuchung der transnationalen Nachrichtenströme zum Ausbruch des Krakatau hat es sich als hilfreich herausgestellt, Bob Nicholsons (2012: 243–245) drei Schritte für die quantitative Forschung mithilfe digitalisierter Printarchive zu folgen: (i) passende Zeitungen herausgreifen, (ii) Schlüsselwörter für die Suche auswählen und verfeinern und (iii) Daten sammeln, analysieren und darstellen, um die Interpretation weiter voranzubringen. Unser Untersuchungsmaterial waren digitalisierte Zeitungsseiten aus sieben verschiedenen Nationalbibliotheken, die erst kürzlich ihre ›echten‹ Zeitungen in (mehr oder weniger) maschinenlesbare Formate konvertiert hatten.⁶ Das

5 Im Jahr 2007 veröffentlichte eine Gruppe von Astronom:innen einen Artikel mit interessanten Implikationen für die Kunstgeschichte: Ihre Forschungsergebnisse ließen darauf schließen, dass der ikonische Sonnenuntergang in Edvard Munchs berühmtem Gemälde *Der Schrei* mit seinen intensiven Farben weniger ein frühes Werk des Expressionismus war als vielmehr eine realistische Wiedergabe des Dämmerlichts nach der Eruption des Krakatau. Bei dem Ausbruch wurden u.a. gewaltige Mengen an Partikeln und Gasen in die Atmosphäre geschleudert, was zu ungewöhnlich farbigen und intensiven Dämmerlichtsituationen, grünen und blauen Sonnen- und Mondeffekten, Koronen, Wolkenschleieren und Bishopschen Ringen führte. Diese Phänomene ließen sich nach dem Ausbruch 1883 bis in die ersten Monate des Jahres 1884 hinein beobachten: Von der südlichen Halbkugel wurden die Vulkanemissionen in die nördlichen Breitengrade transportiert und lösten atmosphärische Anomalien in verschiedenen Erdregionen aus, auch in Norwegen, wo der berühmte Maler damals lebte (Olson/Olson/Doescher 2007). Für weitere Studien zu dem Ausbruch und seinen Folgen in Kunst und Literatur vgl. Altick (1960); Daly (2011); Dobosiewicz (2020).

6 Im Folgenden beziehen sich ›wir‹ und ›unser‹ auf eine Gruppe von Forscher:innen, die an einem transnationalen Data-Mining-Projekt zu Zeitungen teilnahmen, insbesondere Max Franke, Jana Keck, Moritz Knabben, Steffen Koch, Ryan Muller, Mila Oiva, Sebastian Padó, James Parker, Ernesto Priani, Martin Riedl, Hannu Salmi, Melissa Terras und Jaap Verheul. Das internationale Forschungsprojekt *Oceanic Exchanges: Tracing Global Information Networks in Historic Newspaper Repositories, 1840–1914* sammelte mehr als 100 Millionen digitalisierte Zeitungsseiten aus sieben Ländern. Wissenschaftler:innen und Forscher:innen aus jedem der untersuchten Länder mit Erfahrungen in computergestützter Zeitungsforschung arbeiteten disziplinübergreifend zusammen, um Muster im Informationsfluss über Landes- und Sprachgrenzen hinweg zu untersuchen. Um den transnationalen Charakter der Presse des 19. Jahrhunderts weiter zu untersuchen, entwickelte die Projektgruppe digitale Fallstudien, die Einblick in die erste Phase der Globalisierung von Informationen geben und die Entwicklung einer modularen und erweiterungsfähigen Ontologie für die

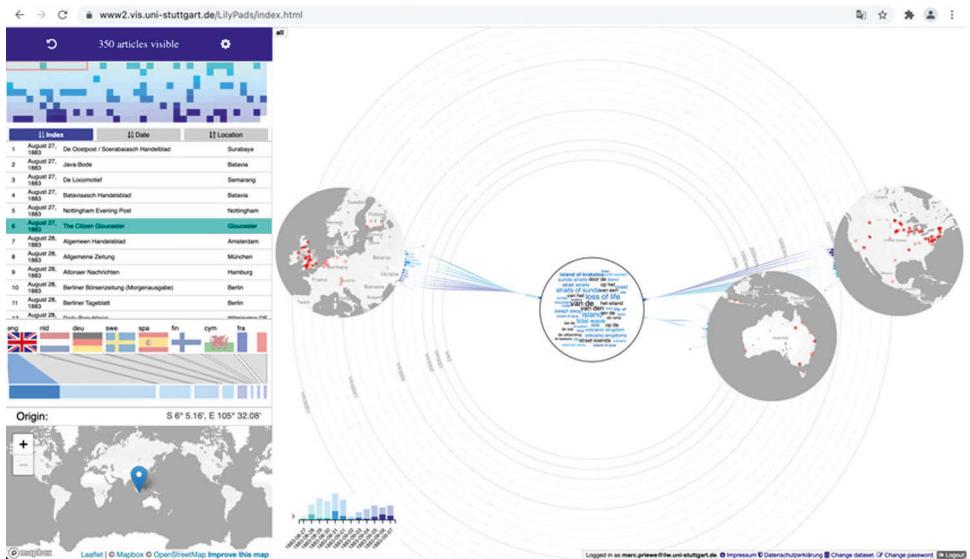
hieraus erstellte Arbeitskorpus umfasste 350 Artikel über den Vulkanusbruch, die zwischen dem 27. August, dem Tag der finalen kataklysmischen Eruption der Vulkaninsel, und dem 7. September 1883 veröffentlicht wurden. Wir wählten einen Zeitraum von zwei Wochen, um die Anzahl der Artikel überschaubar zu halten und gleichzeitig die Möglichkeit zu haben, Trends und Muster aus dem anfänglichen Nachrichtenzyklus ableiten zu können. Anschließend kam eine Kombination aus Close Reading und Distant Reading zum Einsatz: Wir lasen zunächst alle Artikel sorgfältig und nutzten dann quantitative Text- und Data-Mining-Techniken aus der Computerlinguistik sowie Datenvisualisierungen, um nach linguistischen, stilistischen und geografischen Mustern zu suchen, die bei einem einfachen Close Reading möglicherweise übersehen worden wären. Daraufhin wurden bestehende Algorithmen zur Reprint-Erkennung so weiterentwickelt, dass sich mit ihrer Hilfe die Wiederverwendung des ersten Telegramms, das das Ausmaß der Explosion zusammenfasste, innerhalb unseres Korpus nachvollziehen ließ.

Unser Forschungsteam unternahm verschiedene Schritte, um die Analyse der ausgewählten internationalen Zeitungsarchive praktikabler und produktiver zu gestalten und damit Lösungen für einige der typischen Schwierigkeiten bei der Arbeit mit digitalen Archiven zu schaffen. So waren etwa bei der Digitalisierung entstandene Verrauschungen vorrangig auf Artikelsegmentierung und Zeichenerkennung zurückzuführen. Durch Segmentierung und Clustering der historischen Zeitungsdaten sowie die Einführung und Verbesserung von Nomenerkennungsmethoden (»named-entity recognition«, NER) konnten in diesen beiden Problembereichen Verbesserungen erzielt werden. Aufgrund ihres komplexen Textsatzes sind historische Zeitungsarchive oft nicht nach Artikeln, sondern nach Seiten sortiert. Dabei befinden sich lange Artikel nicht immer auf einer Seite, sondern werden über mehrere Seiten, manchmal sogar über mehrere Ausgaben hinweg fortgesetzt. Zur Reduzierung der Verrauschungen bei den verfügbaren Zeitungsdaten wurden manuelle Annotationen vorgenommen und im Anschluss Textsegmentierungsmethoden zur Identifizierung der Segmente getestet. Anschließend nutzten wir computergestütztes Clustering, um fortgesetzte Artikel zusammenzuführen und praktikablere Segmentierungen und Trennungen der digitalisierten Zeitungsartikel zu erreichen (vgl. Riedl/Betz/Padó 2019; Riedl/Padó 2018).

Als nächstes entwickelten wir interaktive Visualisierungstools, mit denen wir den einzelnen Zeitungsartikeln geografische Koordinaten zuwiesen. Weil die Koordinaten nicht Bestandteil der zugrundeliegenden Datensätze waren, wurden sie für alle Zeitungen mithilfe von automatischen und manuellen Methoden festgesetzt. Hierfür wurden vorhandene Metadaten verwendet, die meistens von den Bibliotheken oder Archiven bereitgestellt wurden, die den Originalbestand, d.h. das archivierte Material der jeweiligen Zeitung, digitalisiert oder kuratiert hatten. Insbesondere die Titel von Zeitungen lassen häufig auf den Ort ihrer Herausgabe schließen, doch für die geografische Zuordnung ist im Allgemeinen eine Disambiguierung erforderlich (*Berliner Morgenpost* wird bspw. zu »Berlin, Deutschland« und mit den entsprechenden Geokoordinaten verknüpft). Eine eindeutige Verortung war allerdings nicht in allen Fällen möglich, und oftmals traten

Mehrdeutigkeiten oder Widersprüche hervor, die manuell aufgelöst werden mussten. Außerdem wurde die Sprache der Zeitungsaufsätze mithilfe eines Tools zur Sprachbestimmung ermittelt und falsche Zuordnungen wurden korrigiert. Da zuverlässige Sprachinformationen die Basis für komparative Analysen und dementsprechend auch für die Implementierung verschiedener Visualisierungstools und -prototypen bilden, erwies es sich als hilfreich, die im digitalen Korpus zusammengestellten Artikel mithilfe von Google Translate ins Englische zu übersetzen (vgl. Franke et al. 2020; siehe Abb. 1).

Abbildung 1: Nachrichtenberichterstattung zu Krakatau vom 27.08.1883 bis zum 07.09.1883.



Visualisierung von Max Franke.

Bezüglich der internationalen Berichterstattung zum Ausbruch des Krakatau im Jahr 1883 argumentiert Simon Winchester (2003: 185, 198), dass diese Naturkatastrophe den Beginn des von Marshall McLuhan begrifflich eingeführten »Global Village« verkörpere: Es sei das erste Medienereignis gewesen, an dem Menschen in aller Welt teilhatten und habe zu einem noch nie dagewesenen Gefühl von globaler Verbundenheit und einem neuen Bewusstsein für den Planeten geführt. Dennoch sollte der Beginn des Global Village von McLuhan nicht ausschließlich an diesem Ereignis festgemacht werden, denn die anfänglichen Nachrichten von dieser Katastrophe und ihren Auswirkungen erreichten nicht alle Menschen auf der Welt gleichermaßen. Bei Betrachtung der Nachrichtenflüsse am Ende des 19. Jahrhunderts lassen sich sogar deutliche Asymmetrien feststellen. Einige Zeitungen, und somit auch ihr jeweiliges Publikum, befanden sich 1883 näher an den Nachrichtenkanälen als andere. Dies lässt sich bspw. anhand einiger US-amerikanischen Kleinstädte zeigen, deren Zeitungen beeindruckend schnell und genau über die Ereignisse in Indonesien berichteten. Andere Lesergruppen, bspw. in Mexico-Stadt, wurden

dagegen erst knapp eine Woche nach dem Ausbruch von den Meldungen erreicht. Daher waren der Ausbruch und der teilweise Einsturz des Krakatau kein globales Ereignis nach den Maßstäben des 21. Jahrhunderts – einer Ära, in der die Menschen weltweit quasi zeitgleich Zugang zu neuen Informationen erhalten. Durch Textauswertungsmethoden, Visualisierungssoftware und die wachsenden Datenbestände der digitalen Zeitungsarchive lässt sich heute detaillierter nachvollziehen, wie die Nachrichten über diese Katastrophe in verschiedene Länder auf der ganzen Welt getragen wurden und sich dort verbreiteten.

Um die Frage nach der globalen Nachrichtenverfügbarkeit und damit auch die angenommene Entstehung eines Bewusstseins für globale Zusammenhänge durch die Nachrichtenmedien des 19. Jahrhunderts genauer zu betrachten, sollen im Folgenden einige Beobachtungen aus der komparativen Untersuchung der derzeit verfügbaren Berichterstattung zum Ausbruch des Krakatau in den oben genannten digitalen Archiven vorgestellt werden. Der Nachrichtenfluss zu der Naturkatastrophe in Niederländisch-Ostindien begann mit einem Telegramm, das am 27. August 1883 aus Batavia an die Nachrichtenagentur Reuters in London geschickt wurde⁷:

Terrific detonations from the volcanic island Krakatoa were heard last night, and were audible as far as Soerakarta, showers of ashes falling as far as Cheribon. The flashes from the volcano are plainly visible from here. Serang is now in total darkness. Stones have fallen at that place. Batavia is also nearly in darkness. All the gaslights were extinguished during the night. It is impossible to communicate with Anjer, and it is feared that some calamity has happened there. Several bridges between Anjer and Serang have been destroyed and a village near the former place has been washed away, the rivers having overflowed through the rush of the sea inland. (Reuters zit.n. Eliot 2006: 167)

1883 gab es zwei Möglichkeiten, ein Telegramm von Jakarta (Batavia) nach London zu senden: entweder auf dem Landweg – diese Verbindung war langsamer und weniger sicher, aber vergleichsweise kostengünstig – oder über die neuen Unterseeleitungen, was schneller, aber auch teurer war. Im Durchschnitt benötigte die Übertragung von Nachrichten oder anderen Informationen von Indonesien nach England über Land etwa eine Woche. Nach der Fertigstellung der Unterwasserverbindung Mitte der 1870er Jahre benötigten die Nachrichten aus den Überseegebieten des britischen Empires bis in sein Zentrum nur noch drei Stunden (siehe Winchester 2003: 193). Berücksichtigt man die Zeitdifferenz von sieben Stunden zwischen Batavia/Jakarta und London, dann haben einige der Morgenzüge mit guter Verbindung an Reuters es geschafft, die Nachricht von der Naturkatastrophe am anderen Ende der Welt beinahe in Echtzeit an ihr Publikum weiterzugeben. Nach unseren Daten brachten zwei Zeitungen in Großbritannien die Meldung noch am selben Tag, also am 27. August, wogegen die meisten Zeitungen in den großen Städten Europas, Nordamerikas und Australiens erst am nächsten Tag, dem 28. August, über den Ausbruch berichteten. Generelle Aussagen über die Darstellung der Krakatau-Meldungen in den für die vorliegende Studie ausgewählten Zeitungen

7 Für weitere Informationen zur Rolle der Nachrichtenagenturen bei der internationalen Verbreitung von Nachrichten siehe Silberstein-Loeb (2014: 163-225).

aus verschiedenen Ländern lassen sich schwer treffen, weil sie sich in Form und Layout stark unterschieden. In den meisten Zeitungen erschienen die ersten Berichte über den Vulkanausbruch unter den Auslandsnachrichten oder in der Telegrafenkolumne, meist auf den Innenseiten (Seiten 2-3 in vierseitigen Ausgaben; Seiten 4-5 in Ausgaben mit acht Seiten), und nicht, wie man es heute erwarten würde, fettgedruckt auf der Titelseite (siehe Barnhurst/Nerone 2001: 99-105).

Mithilfe der Reprint-Erkennungsalgorithmen konnten wir in Bezug auf alle sieben staatlichen Zeitungsarchive herausarbeiten, dass das oben zitierte Reuters-Telegramm aus Jakarta vom 27. August in der ersten Woche nach der finalen Explosion häufig wiederverwendet wurde. Einige Zeitungen gaben das Telegramm leicht gekürzt wieder, andere fügten zusätzliche Informationen über ›Java‹ hinzu (ein Begriff, der im Westen damals nur wenigen bekannt war), meist wurde der Text jedoch wortgetreu abgedruckt. In den ersten zwei Wochen nach dem Ausbruch ähnelte sich der Nachrichtenzyklus in den Metropolen: Die meisten Zeitungen in London, Berlin, Helsinki, New York und vielen anderen Großstädten der westlichen Welt druckten in den ersten Tagen die Telegramme aus der Krisenregion ab. Erst ab der zweiten Woche finden sich stärker ausgearbeitete Berichte und Leitartikel über die Auswirkungen des Vulkanausbruchs.

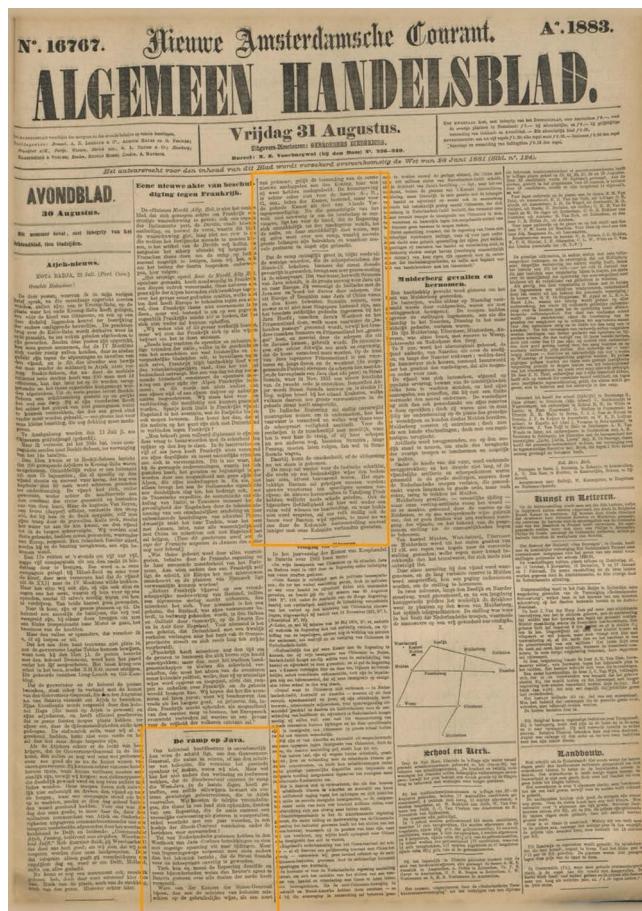
Im Folgenden soll der Fokus weniger auf die Details der Berichterstattung, sondern vielmehr auf einige thematische und stilistische Muster gelenkt werden, die sich aus den Daten ergeben. Erstens sind bei der Berichterstattung über Krakatau Zeichen einer Bloch'schen ›Gleichzeitigkeit des Ungleichzeitigen‹ festzustellen: Es gab eine unmittelbare und umfangreiche Berichterstattung über Detonationen und Zerstörungen in der lokalen indonesischen Presse – diese Informationen wurden jedoch nicht in den Zeitungen anderer Metropolen abgedruckt. Deren Nachrichtenzyklus zum Ausbruch des Krakatau begann mit dem Reuters-Telegramm, das die anfängliche Berichterstattung in weiten Teilen Europas und den Vereinigten Staaten dominierte, jedoch nicht von den in Java herausgegebenen Zeitungen der niederländischen Kolonien abgedruckt wurde. Zweitens gab es signifikante stilistische und thematische Unterschiede bei der Berichterstattung außerhalb Indonesiens – bspw. fokussierte sich die australische Presse hauptsächlich auf die Einschränkungen in der Schifffahrt und auf den Handelsrouten in der Region; die mexikanische Presse befasste sich weitgehend mit den menschlichen Tragödien in Java – der Name des Vulkans wurde in den ersten 10 Tagen nach der Eruption nicht erwähnt; das Korpus der Vereinigten Staaten zeigt eine größere Anzahl von eigenen, eher sensationslüsternen Berichten als alle anderen untersuchten Datenkorpora; das finnische Archiv illustriert, wie gut die Leserschaft aus den Randregionen Europas über Menschen und Ereignisse in Tausenden Kilometern Entfernung informiert war; deutsche Zeitungen verwendeten im Vergleich zu den Zeitungen anderer Länder eine bildhafte Sprache, um das tragische Ausmaß des Ereignisses zu unterstreichen. Drittens wurde die Berichterstattung über die Katastrophe Bestandteil der kolonialen Rivalitäten zwischen Großbritannien und den Niederlanden, wobei es hauptsächlich um die Frage ging, wer die Berichterstattung und Informationsflüsse kontrollierte.

Es wäre zu erwarten gewesen, dass das niederländische Zeitungspublikum am besten über die Vorgänge in der Kolonie Java unterrichtet war. Aufmerksam Lesende hätten die ersten Berichte im Mai 1883 zum ersten Ausbruch des Vulkans nicht übersehen,

der nur wenige Tage anhielt und kaum Schäden verursachte. In dem von uns untersuchten digitalisierten Zeitungskorpus berichtete einzig die niederländische Presse durchgehend seit dem ersten Ausbruch im Mai über Krakatau, was verständlich ist, wenn man berücksichtigt, dass die Niederlande ein ureigenes Interesse am Erhalt der Schifffahrtsverbindung in ihre Überseekolonie und an der Kontrolle über deren Regierungsapparat hatten. Angesichts der zahlreichen Kommunikationskanäle zwischen dem Zentrum des niederländischen Kolonialreichs und seiner Peripherie überrascht es, dass der erste Bericht in der niederländischen Presse nach der finalen Eruption im August 1883 eine Übersetzung des Reuters-Telegramms von Jakarta nach London war, auf das sich auch die meisten anderen großen Zeitungen verließen. Nachdem die niederländischen Zeitungen sich in den ersten Tagen des Nachrichtenzyklus auf die Auswirkungen des Vulkanausbruchs konzentriert hatten, wies am 31. August das *Algemeen Handelsblad* in Amsterdam darauf hin, dass die Niederlande unter einem Informationsnachteil litten und klagte damit indirekt die Regierung an, dass sie im imperialen Streit um die Informationskontrolle und damit auch im Wettbewerb mit Großbritannien unterlänge. Unter dem Titel »De ramp op Java« (»Die Katastrophe in Java«) auf der Titelseite der Abendausgabe kritisierte der Artikel die Regierung als unfähig, angemessene Informationen über den Verbleib niederländischer Familien in Indonesien zu liefern, weil sie auf den Nachrichtenfluss über London angewiesen sei (siehe Abb. 2).

Dass die Nachrichten aus Niederländisch-Ostindien zuerst in London statt in Amsterdam eingingen, lag hauptsächlich in der Tatsache begründet, dass Großbritannien zu dieser Zeit führend in der Produktion und Verlegung von Land- und Unterseekabeln für die Telegrafie war. Im Wettstreit um die Vorherrschaft zwischen den Niederlanden und Großbritannien gewann letzteres tatsächlich die Oberhand bei der Errichtung transnationaler, kolonialer Kommunikationsnetze, die Tom Standage (1998) als das »viktorianische Internet« bezeichnet hat. Durch seinen direkten Zugang zu der Kommunikationsstruktur und deren Kontrolle war Großbritannien beim Aufbau einer Informationshegemonie klar im Vorteil. Dies war wiederum ausschlaggebend für den Ausbau des Handels und bildete die Grundlage für die Regierung und Expansion des britischen Empires (siehe Wobring 2005: 319-336; vgl. Ahvenainen 1995; Boyce 1995).

Abbildung 2: »De ramp op Java«, *Algemeen Handelsblad*, 31. August 1883, S. 1. Hervorh. d. Verf.



Quelle: www.delpher.nl

4. Schlussausführungen

Die immer weiter wachsenden Archive digitalisierter Zeitungen aus einer zunehmenden Anzahl von Ländern ermöglichen stärker differenzierte und transnationale Vergleiche der Nachrichtenberichterstattung. Unsere Daten zeigen, dass der Fluss von Nachrichten zum Ausbruch des Krakatau im August 1883 nicht unmittelbar global verlief, sondern letztendlich eher diskontinuierlich und heterogen war. Während die Bevölkerung der großen Städte in Europa, Nordamerika und Australien tatsächlich innerhalb von Stunden und Tagen von der Katastrophe in Indonesien erfuhr, erreichten die Nachrichten die Menschen in abgeschiedeneren Regionen Lateinamerikas oder in ländlichen Gegendern im Allgemeinen sehr viel langsamer und in einigen Fällen überhaupt nicht. Es ist somit vielleicht etwas übertrieben, den Ausbruch als die erste Naturkatastrophe zu

bezeichnen, von der die gesamte Menschheit wusste, und die die Welt entsprechend der von Winchester übernommenen McLuhan'schen Metapher zu einem globalen Dorf ›schrumpfen‹ ließ. Die Behauptung, dass der Krakatau ein frühes Beispiel für die Globalisierung der Medien war, übersieht bei den Nachrichtenströmen wichtige Details, die sich heute mittels Informationstechnologie und digitaler Archive genauer nachverfolgen lassen.

Im Rahmen des hier vorgestellten Forschungsprojekts konnte eine Anzahl von Annahmen und Ergebnissen aus früheren Untersuchungen zur Herausbildung zunehmend globaler Informationsflüsse durch die Zeitungen des 19. Jahrhunderts bestätigt und erweitert werden. Unsere Datenanalyse baute auf Sybille Rechbergers (2011) ›analogen‹ Untersuchungsmethoden zur Betrachtung der Nachrichtenberichterstattung zum Ausbruch des Krakatau im Jahr 1883 auf und erweiterte diese. So konnte etwa gezeigt werden, dass die internationalen Nachrichtensysteme zwar wichtige Themen weltweit bekannt machten, ein allgemeines Bewusstsein dafür schufen und breite Debatten generierten – die jeweiligen lokalen, kulturellen und politischen Umstände hatten jedoch stets einen deutlichen Einfluss auf die Perspektive und die Rezeption der Berichterstattung zu diesen global relevanten Themen. Der Zugang zur technischen Infrastruktur (bspw. verbesserte Druckerpressen, Telegrafie, Unterseekabel) beeinflusste sowohl die Geschwindigkeit, mit der die Nachrichten verschiedene Orte erreichten, als auch die vorherrschenden Perspektiven der Berichterstattung in verschiedenen Ländern.

Wie sich ebenfalls zeigte, bleiben möglicherweise im Digitalisierungsprozess von analogen Archiven auftretende Fehler und Verzerrungen oft unentdeckt und erschweren die Umsetzung transnationaler Untersuchungen mithilfe digitaler Archive. Es erweist sich daher als schwierig, dem Aufruf Mussells (2012) zu folgen und sich die Schritte, Entscheidungen und Abläufe im Rahmen des Digitalisierungsprozesses vollständig zu vergegenwärtigen: Für Archivnutzer:innen sind die Algorithmen und Mechanismen, die historische Daten in etwas verwandeln, das auf einer grafischen Oberfläche lesbar ist, äußerst selten einsehbar und zudem ohne IT-Kenntnisse kaum verständlich. Hinzu kommt noch, dass die meisten digitalen Archive wenig Informationen darüber herausgeben, wie die präsentierten Materialien entstanden sind, und im Allgemeinen auch nicht auf mögliche bei der Digitalisierung entstandene Fehler hinweisen. Da es nur wenig erschwingliche, zuverlässige und effektive Software für die Arbeit mit digitalisiertem Printmaterial gibt, sind Forschende neben der Anstellung von oder Zusammenarbeit mit IT-Expert:innen auf den Ausbau ihrer eigenen Computerkenntnisse angewiesen.

Im Rahmen der Untersuchungen zum Ausbruch des Krakatau hat unsere interdisziplinäre Forschungsgruppe neue Methoden und Werkzeuge für transnationale Untersuchungen getestet und entwickelt, die eine Kombination aus digitaler Suche und Mustererkennung in historischen Zeitungssammlungen ermöglichen. Eine solche computergestützte Herangehensweise kann auch für die Betrachtung von Texten und Berichten zu anderen Themen – historischen Ereignissen, Persönlichkeiten, Ideen und Konzepten – eingesetzt und angepasst werden, die in der Vergangenheit im Umlauf waren und in digitalisierten Sammlungen erfasst wurden. Das Ermitteln linguistischer oder visueller Muster in einzelnen Datensammlungen oder über mehrere digitale Archive hinweg kann als eine Form des Distant Reading einen methodischen Beitrag zur Erforschung und Analyse von historischen Dokumenten leisten. Unabhängig davon, ob ein Forschungs-

projekt digitale Zeitungen, Zeitschriften, Periodika oder andere online verfügbare Printmaterialien nutzt, ist zu erwarten, dass die kontinuierliche Verbesserung von Methoden zur Extraktion, Verarbeitung und Analyse von Texten als Daten künftig zu weiteren Forschungsfortschritten sowohl in der Geisteswissenschaft als auch der Informatik führt.

Literatur

Ahvenainen, Jorma (1995): The Role of Telegraphs in the 19th Century Revolution of Communications. In: North, Michael (Hg.): Kommunikationsrevolutionen. Die Neuen Medien des 16. und 19. Jahrhunderts. Köln: Böhlau. S. 73-80.

Altick, Richard D. (1960): Four Victorian Poets and an Exploding Island. In: Victorian Studies 3.3 (1960). S. 249-260.

Barnhurst, Kevin G./Nerone, John (2001): The Form of News. A History. New York: Guildford Press.

Beals, Melodee H. (2017): Scissors and Paste. The Georgian Reprints, 1800-1837. In: Journal of Open Humanities Data 3.1 (2017). S. 1-5.

Bloch, Ernst (1985 [1962]): Erbschaft dieser Zeit. Frankfurt a.M.: Suhrkamp.

Bourke, Thomas A. (1990): Scholarly Micropublishing, Preservation Microfilming, and the National Preservation Effort in the Last Two Decades of the Twentieth Century: History and Prognosis. In: Microform Review 19.1 (1990). S. 4-16.

Boyce, Robert (1995): Submarine Cables as a Factor in Britain's Ascendancy as a World Power. In: North, Michael (Hg.): Kommunikationsrevolutionen. Die Neuen Medien des 16. und 19. Jahrhunderts. Köln: Böhlau. S. 81-100.

Chapman, Jane (2016): Transnational Connections. In: King, Andrew/Easley, Alexis/Morton, John (Hg.): The Routledge Handbook to Nineteenth-Century British Periodicals and Newspapers. New York: Routledge. S. 175-184.

Daly, Nicholas (2011): The Volcanic Disaster Narrative. From Pleasure Garden to Canvas, Page, and Stage. In: Victorian Studies 53.2 (2011). S. 255-285.

Dobosiewicz, Ilona (2020): ›The wrathful sunset glared...‹: The Krakatoa Sunsets in Victorian Science and Art. In: Anglica Wratislaviensia LVIII (2020). S. 11-20.

Douglas, George H. (1999): The Golden Age of the Newspaper. Westport: Greenwood Press.

Eliot, Nils Lindah (2006): Mediating Nature. New York: Routledge.

Franke, Max/John, Markus/Knabben, Moritz/Keck, Jana/Blascheck, Tanja/Koch, Steffen (2020): LilyPads. Exploring the Spatiotemporal Dissemination of Historical Newspaper Articles. In: Proceedings of the 11th International Conference on Information Visualization Theory and Applications 3 (2020). S. 17-28.

Jarlbrink, Johan/Snickars, Pelle (2017): Cultural Heritage as Digital Noise. Nineteenth Century Newspapers in the Digital Archive. In: Journal of Documentation 73.6 (2017). S. 1228-1243.

Kaplan, Richard L. (2002): Politics and the American Press. The Rise of Objectivity, 1865-1920. Cambridge: Cambridge University Press.

King, Edmund (2005): Digitisation of Newspapers at the British Library. In: The Serials Librarian. From the Printed Page to the Digital Age 49.1-2 (2005). S. 165-181.

McLuhan, Marshall/Powers, Bruce R. (1989): *The Global Village. Transformations in World Life and Media in the 21st Century*. Oxford: Oxford University Press.

Mussel, James (2016): Digitization. In: King, Andrew/Easley, Alexis/Morton, John (Hg.): *The Routledge Handbook to Nineteenth-Century British Periodicals and Newspapers*. New York: Routledge. S. 17-28.

Mussell, James (2012): *The Nineteenth-Century Press in the Digital Age*. Basingstoke: Palgrave Macmillan.

Nicholson, Bob (2012): Counting Culture; or, How to Read Victorian Newspapers from a Distance. In: *Journal of Victorian Culture* 17.2 (2012). S. 238-246.

Olson, Marilynn S./Olson, Donald W./Doescher, Russell L. (2007): The Blood-Red Sky of Munch's 'The Scream'. In: *Environmental History* 12.1 (2007). S. 131-135.

Putnam, Lara (2016): The Transnational and the Text-Searchable. Digitized Sources and the Shadows They Cast. In: *American Historical Review* 121.2 (2016). S. 377-402.

Rechberger, Sybille Christine (2011): Der Ausbruch des Krakataus 1883 und seine Bedeutung in der zeitgenössischen Presse. Diplomarbeit, Universität Wien.

Riedl, Martin/Padó, Sebastian (2018): A Named Entity Recognition Shootout for German. In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Short Papers)*. Melbourne Australia, July 15-20, 2018. S. 120-125; <https://www.aclweb.org/anthology/P18-2020.pdf> (26.04.2022).

Riedl, Martin/Betz, Daniela/Padó, Sebastian (2019): Clustering-Based Article Identification in Historical Newspapers. In: *Proceedings of the 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*. Minneapolis, MN, June 7, 2019. S. 12-17; <https://www.aclweb.org/anthology/W19-2502.pdf> (26.04.2022).

Silberstein-Loeb, Jonathan (2014): *The International Distribution of News: The Associated Press, Press Association, and Reuters, 1848-1947*. Cambridge: Cambridge University Press.

Silverman, Randy (2005): Retaining Hardcopy Papers Still Important in Digital Age. In: *Newspaper Research Journal* 36.3 (2005). S. 363-372.

Simkin, Tom/Fiske, Richard S. (1983): *Krakatau 1883. The Volcanic Eruption and Its Effects*. Washington: Smithsonian Institution Press.

Smith, David A./Cordell, Ryan/Mullen, Abby (2015): Computational Methods for Uncovering Reprinted Texts in Antebellum Newspapers. In: *American Literary History* 27.3 (2015). S. 1-15.

Standage, Tom (1998): *The Victorian Internet: The Remarkable Story of the Telegraph and the Nineteenth Century's Online Pioneers*. London: Weidenfeld & Nicolson.

Thornton, Ian (1996): *Krakatau. The Destruction and Reassembly of an Island Ecosystem*. Cambridge: Harvard University Press.

Walravens, Hartmut (Hg.) (2006): *International Newspaper Librarianship for the 21st Century*. München: K. G. Saur.

Wilkinson, M. D. et al. (2016): The FAIR Guiding Principles for Scientific Data Management and Stewardship. In: *Scientific Data* 3 (2016). Article Number: 160018; <https://www.nature.com/articles/sdata201618> (26.04.2022).

Winchester, Simon (2003): *Krakatoa: The Day the World Exploded*. New York: Viking.

Wobring, Michael (2005): Die Globalisierung der Telekommunikation im 19. Jahrhundert. Pläne, Projekte und Kapazitätsausbauten zwischen Wirtschaft und Politik. Frankfurt a.M.: Lang.

