

# Algorithmen- gestützte Analyse visuell-materieller Eigenschaften von Briefen

## Nantke, Julia

julia.nantke@uni-hamburg.de  
Universität Hamburg, Deutschland

## Reul, Christian

christian.reul@uni-wuerzburg.de  
Universität Würzburg, Deutschland

## Bläß, Sandra

sandra.blaess@uni-hamburg.de  
Universität Hamburg, Deutschland

## Flüh, Marie

marie.flueh@uni-hamburg.de  
Universität Hamburg, Deutschland

Der visuell-materiellen Gestaltung von historischen Dokumenten und insbesondere von Briefen kommt in der Forschung ein hoher Stellenwert zu. Visuell-materielle Eigenschaften wie die Papier- und Stifffarbe, die Anordnung von Schrift auf der Seite und die Schriftführung geben Auskunft über historische Bedingungen und konkrete Umstände des Schreibens und der Schreibenden sowie, im Fall von Briefen, über das Verhältnis der Schreibenden zu den jeweiligen Adressat:innen (vgl. Baasner 2008; Wiethölter/Bohnenkamp 2010; Henzel 2020; Lukas/Osthof 2016).

Obwohl die digitale Repräsentation historischer Dokumente in digitalen Editionen und Archiven entscheidende Vorteile hinsichtlich der Darstellung visuell-materieller Charakteristika bietet (vgl. Bohnenkamp-Renzen 2013; Radecke 2015), spielt deren Analyse in computer-gestützten Verfahren noch immer eine untergeordnete Rolle. Die meisten quantitativen Methoden und hierbei genutzten Tools beziehen sich auf den Inhalt und/oder den Schreibstil literarischer und historischer Texte, die bibliografischen Codes (McGann 1991: 77) bleiben jenseits ihrer Repräsentation als Bild-Digitalisate in digitalen Editionen bei der Analyse und Interpretation meistens unberücksichtigt.

Im Rahmen des geplanten Vortrags soll ein Ansatz zur quantitativen, Algorithmen-gestützten Erschließung visuell-materieller Charakteristika in einem Korpus von Briefen aus dem Zeitraum ‚Deutschland um 1900‘ präsentiert werden. Hierbei wollen wir zeigen, wie sich anhand verschiedener (teil-)automatisiert ermittelbarer Werte zur optisch erfassbaren Gestaltung der Briefe inhaltliche Aussagen sowohl über konkrete Dokumente

sowie deren Schreibende und Adressat:innen machen als auch Erkenntnisse über die historischen Dynamiken der Textsorte ‘Brief’ gewinnen lassen. Denn zum einen handelt es sich bei Briefen um einen Dokumententyp, der spätestens seit dem 18. Jahrhundert spezifischen Kodierungen unterworfen ist, bei denen die Ebenen des sprachlichen Ausdrucks und der visuell-materiellen Gestaltung komplex ineinandergreifen (vgl. Baasner 2008). Zum anderen zeichnet sich gerade der untersuchte Entstehungszeitraum der Briefe durch die Lockerung kommunikativer Etikette aus, was im Rahmen des Briefformats zunehmende Spielräume zur individuellen Ausgestaltung eröffnet (vgl. Ehlers 2004).

Bei dem analysierten Korpus handelt es sich um einen Ausschnitt aus dem ca. 35.000 Briefe umfassenden Dehmel-Archiv der Staats- und Universitätsbibliothek Hamburg (SUB), die aktuell im Projekt *Dehmel digital* wissenschaftlich erschlossen werden.<sup>1</sup> Die Briefe richten sich an den um 1900 berühmten, 1920 verstorbenen Dichter Richard Dehmel und stammen von verschiedenen anderen Künstler:innen wie Rainer Maria Rilke, Stefan Zweig, Else Lasker-Schüler, Detlev von Liliencron und Peter Behrens, die in unterschiedlich engem Kontakt zu Dehmel standen und deren Erfolg als Künstler:innen zum Zeitpunkt der Korrespondenzen ebenfalls verschieden groß war. Weiterhin enthält das Korpus auch Briefe von Vertreter:innen des Literaturbetriebs, die mit Dehmel über Publikationsvorhaben, geplante Veranstaltungen oder gemeinsame Projekte verhandelten.

Wir wollen in unserem Beitrag zeigen, wie sich anhand der visuell-materiellen Gestaltung der Briefe Aussagen über die Lebensumstände der Briefschreibenden, deren Beziehung zu Dehmel sowie die Charakteristika epistolarer Kommunikation um 1900 und die Veränderung der Normen der Briefschreibung ableiten lassen, welche sich in dieser Zeit vollziehen.

Die Grundlage für die maschinenlesbare Erschließung visuell-materieller Eigenschaften der Dokumente bilden die Strukturinformationen auf der Briefseite. Diese werden im Rahmen der Layoutanalyse, einem Teilschritt der HTR (Handwritten Text Recognition), semi-automatisch mittels OCR4all<sup>2</sup> (vgl. Reul et al. 2019) erfasst. Hierbei können Strukturen wie der Seitenspiegel, der Haupttext, Grußformeln, Briefköpfe etc. entweder manuell ausgezeichnet und typisiert oder ein automatisch generierter Vorschlag gezielt korrigiert werden, wobei manuell geprüfte Daten als Trainingsbeispiele für die fortlaufende Verbesserung der algorithmischen Methoden herangezogen werden können. Aus den so erzeugten Daten können anschließend die Informationen über die visuell erfassbaren Merkmale der Briefe automatisiert extrahiert werden, um im Rahmen statistischer Analysen Zusammenhänge zwischen Layout und Briefinhalt aufzudecken (vgl. Busch/Hegel 2017; Hurlbut 2013). So lassen sich z.B. statistische Mittelwerte für den Weißraum der Briefe einzelner Korrespondenzpartner:innen berechnen und miteinander vergleichen oder die Papierfarbe eines Briefs ins Verhältnis zu den innerhalb des Korpus üblichen Färbungen setzen.

Die von uns untersuchten Merkmale lassen sich grob in zwei Kategorien unterteilen: Erstens gibt es grundsätzlich vorhandene visuell-materielle Eigenschaften wie das Format der Briefbögen, das Verhältnis von Weißraum und Textraum sowie der Abstand zwischen Grußformeln

und Textblock, die für die epistolare „Respektsemiotik“ (Ehlers 2004: 21) von großer Aussagekraft sind. Diese können direkt aus den oben erwähnten Auszeichnungen abgeleitet bzw. berechnet werden. Ebenfalls in den Bereich der grundsätzlichen Eigenschaften gehören die Papier- und Stifffarbe, deren „Auswahl und Wirkungsweise [...] in enger Abhängigkeit – historisch wandelbarer – ökonomischer, kultureller, sozialer, ästhetischer u.a. Faktoren sowie des jeweiligen Inhalts und der Funktion des Briefs“ stehen (Henzel 2020: 222). Die computergestützte Identifikation dieser Merkmale erfordert allerdings im Vergleich einen etwas größeren Aufwand: Zunächst wird das originale Farbbild in ein Binärbild umgewandelt, sodass in den zuvor ausgezeichneten Textregionen die Vordergrundpixel weitestgehend der Schrift und die Hintergrundpixel weitestgehend dem unbeschriebenen Papier entsprechen. Nach dem Ausschließen von Übergangspixeln, um Störeffekte zu minimieren, werden die Entsprechungen im Farbbild gesammelt und der jeweiligen Klasse zugewiesen. Abschließend wird separat die durchschnittliche Schrift- und Papierfarbe berechnet, indem zunächst je ein Mittelwert für die drei Farbkanaäle Rot, Grün und Blau gebildet wird und diese zu je einer Farbe kombiniert werden.

Neben diesen generellen visuell-materiellen Eigenschaften beziehen wir zweitens spezifische Charakteristika wie Briefköpfe, Abbildungen und Zeichnungen in unsere Analysen mit ein, die nur in einem Teil der Dokumente enthalten sind und bei denen vorerst lediglich erfasst wird, ob sie vorhanden sind oder nicht. Diese Erfassung zielt darauf, die Dokumente anschließend nach den entsprechenden Charakteristika filtern zu können, um auf breiter Basis Aussagen über die Verbreitung und die konkreten Eigenschaften der Gestaltungsmittel machen zu können.

Zwei Beispiele sollen im Folgenden die dargestellte Vorgehensweise sowie die Aussagekraft der Ergebnisse der visuell-materiellen Analysen exemplarisch illustrieren.

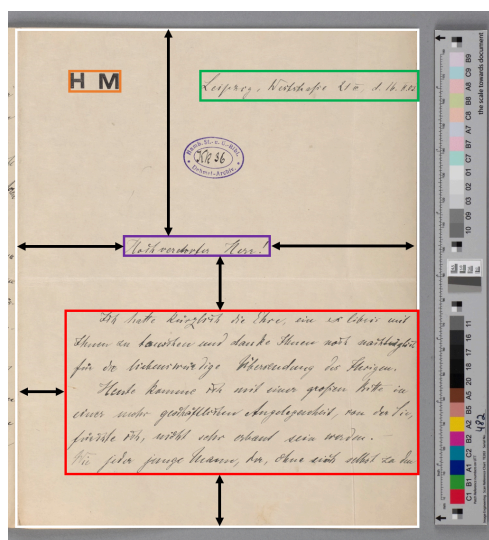


Abb. 1: Brief von Heinz Möller an Richard Dehmel vom 16.02.1903, Dehmel-Archiv der SUB, HANSb18474, S. 1

Der oben abgebildete Brief des Leipziger Herausgebers Heinz Möller vom 16. Februar 1903 ist offensichtlich auf

eigenem Briefpapier verfasst. Wie anhand des Briefkopfs mit den Initialen des Verfassers zu erkennen, tritt Möller im eigenen Namen auf und handelt nicht im Auftrag einer Organisation. Des Weiteren orientiert sich die Gestaltung des Briefs geradezu prototypisch an den Regeln der Respektsemiotik: Das Ausmaß der textfreien Weißräume an den Briefrändern und zwischen den einzelnen Briefteilen (oberer Briefrand und Anrede, Anrede und Textblock) und das äußerst sorgfältige Schriftbild bringen auf visuell-materieller Ebene die Hochachtung des Schreibenden gegenüber dem angeschriebenen Richard Dehmel zum Ausdruck und korrespondieren dabei mit verbalsprachlichen Formeln wie „Hochverehrter Herr“ und „In aufrichtiger Verehrung und Dankbarkeit“. Insbesondere im Vergleich mit dem zweiten Beispiel wird deutlich, dass gerade im Rahmen geschäftlicher Korrespondenz auch zu Beginn des 20. Jahrhunderts die seit der Antike in stetig aktualisierten „Briefstellern“ verbreiteten formalen Regeln zur Abfassung von Briefen noch Gültigkeit besaßen (vgl. Schiegg 2020). Zugleich lässt sich aus der Gestaltung eine Aussage über das Hierarchieverhältnis der beiden Korrespondenzpartner zueinander ableiten: Möller wendet sich hier, wie gesagt in eigener Sache, an Dehmel als „Obmann“ des Kartells lyrischer Autoren, um einen Nachlass für die Honorare einer von Möller geplanten Lyrik-Anthologie zu erbitten. Der Umstand, dass es sich hier um ein Bittschreiben handelt, spiegelt sich eindrücklich in der visuell-materiellen Gestaltung des Dokuments.

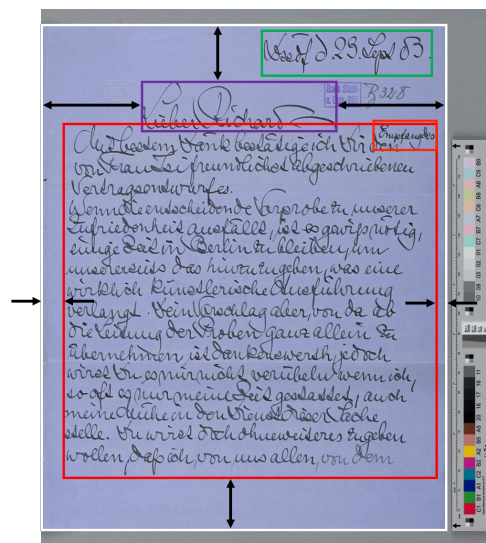


Abb. 2: Brief von Peter Behrens an Richard Dehmel vom 23.09.1903, Dehmel-Archiv der SUB, HANSb313015, S. 1

Das zweite Beispiel, der Brief des Kunsthandwerkers und Designers Peter Behrens an Richard Dehmel vom 23. September 1903, zeigt eine andere Form der signifikanten Individualisierung von Briefpapier: Behrens' Briefbogen enthält zwar keinen aufgedruckten Briefkopf, die farbliche Gestaltung, das lilafarbene Briefpapier hebt ihn aber sehr stark aus dem Feld der breiteren Masse von Briefen aus dem untersuchten Zeitraum hervor und macht ihn, etwa in einem Stapel von Briefen, sofort als einen Brief des Verfassers Behrens erkennbar. Darüber hinaus weist die Gestaltung des Dokuments durch-

aus grundlegende visuell-materielle Charakteristika der Briefschreibung auf, die im Rahmen einer quantitativen Analyse als Eigenschaften des etablierten Brieflayouts der Zeit um 1900 ermittelt werden können: Es besteht, wie für den Brief von Möller festgestellt, ebenfalls ein gewisser Abstand zwischen dem oberen Blattrand und der Anrede sowie zwischen Blattrand und Brieftext, die Anrede ist, wiederum analog zum Brief Möllers, zentriert und der eigentliche Brieftext als Block gesetzt. Im Vergleich mit dem ersten Beispiel zeigen sich allerdings auch deutlich messbare Differenzen, die Rückschlüsse auf das Verhältnis der Korrespondenzpartner erlauben. So besteht ein deutlich geringerer Abstand zwischen oberem Blattrand und Anrede und kein gegenüber dem sonstigen Zeilenabstand vergrößerter Abstand zwischen Anrede und Brieftext. Hinzu kommt eine Wort-Einfügung über der Zeile, die ebenfalls einen Bruch mit der klassischen epistolaren Etikette darstellt und von einer größeren Vertrautheit der Korrespondenzpartner zeugt. Wiederum spiegelt hier die visuell-materielle die verbalsprachliche Gestaltung. Peter Behrens war nicht nur künstlerisch mit Richard Dehmel verbunden, sondern auch ein guter Freund, was sich in der Anrede als „Lieber Richard“ niederschlägt.

Die beiden Beispiele zeigen andeutungsweise, wie die automatisierte, quantitative Auswertung der visuell-materiellen Gestaltungsformen in einem Briefkorpus dazu beitragen kann, bereits vor der genauen inhaltlichen Sichtung der Briefe Hypothesen über das Verhältnis der Korrespondenzpartner:innen und den Zweck der Kommunikation anzustellen sowie darüber hinaus grundsätzliche Charakteristika epistolarer Kommunikation in einem bestimmten Zeitraum zu erschließen.

## Fußnoten

1. Vgl. <https://dehmel-digital.de>.
2. <https://www.ocr4all.org>

## Bibliographie

- Baasner, Rainer.** Stimme oder Schrift? Materialität und Medialität des Briefs. Adressat: Nachwelt. Briefkultur und Ruhmbildung. Hg. v. Detlev Schöttker. München: Wilhelm Fink, 2008, S. 53–69.
- Bohnenkamp-Renken, Anne, Hg.** Medienwandel / Medienwechsel in der Editionswissenschaft. De Gruyter, 2013, <https://doi.org/10.1515/9783110300437>.
- Busch, Hannah, und Philipp Hegel:** Automatic Layout Analysis and Storage of Digitized Medieval Books. In: Digital Philology. A Journal of Medieval Cultures, 6, 2 (2017), S. 196–212, <https://doi.org/10.1353/dph.2017.0010>.
- Dehmel digital.** Hg. v. Julia Nantke unter Mitarbeit von Sandra Bläß und Marie Flüh, seit 2021, <https://dehmel-digital.de>.
- Ehlers, Klaas-Hinrich:** Raumverhalten auf dem Papier. Der Untergang eines komplexen Zeichensystems dargestellt an Briefstellern des 19. und 20. Jahrhunderts. In: Zeitschrift für germanistische Linguistik 32, 1 (2004), S. 1–31.
- Henzel, Katrin:** Materialität des Briefs. In: Handbuch Brief. Von der Frühen Neuzeit bis zur Gegenwart. Hg. v. Marie Isabel Matthews-Schlinzig, Jörg Schuster, Gesa Steinbrink u. Jochen Strobel. Berlin/Boston: De Gruyter 2020, S. 222–231.
- Hurlbut, Jesse:** The Manuscript Average, Part 1. Dezember 2013, [https://jessehurlbut.net/wp/mssart/?page\\_id=2097](https://jessehurlbut.net/wp/mssart/?page_id=2097).
- Lukas, Wolfgang und Matthias Osthof.** „Physische vs. gedutete Räumlichkeit. Zur Auszeichnung spatialer Informationen in der historisch-kritischen Ausgabe C.F. Meyers Briefwechsel“. Jahrbuch für Computerphilologie online, 2016, <http://computerphilologie.digital-humanities.de/jg09/lukasosthof.pdf>.
- McGann, Jerome:** The Textual Condition. Princeton: Princeton University Press 1991.
- Radecke, Gabriele:** Materialautopsie. Überlegungen zu einer notwendigen Methode bei der Herstellung von digitalen Editionen am Beispiel der Genetisch-kritischen und kommentierten Hybrid-Edition von Theodor Fontanes Notizbüchern. TextGrid. Von der Community – für die Community. Eine Virtuelle Forschungsumgebung für die Geisteswissenschaften. Hg. v. Heike Neuroth u. a., 2015, S. 39–56.
- Reul, Christian; Christ, Dennis; Hartelt, Alexander; Balbach, Nico; Wehner, Maximilian; Springmann, Uwe; Wick, Christoph; Grundig, Christine; Büttner, Andreas; Puppe, Frank:** OCR4all – An open-source tool providing a (semi-) automatic OCR workflow for historical printings. In: Applied Sciences, 9(22), 2019.
- Schiegg, Markus: Briefsteller.** In: Handbuch Brief Von der Frühen Neuzeit bis zur Gegenwart. Hg. v. Marie Isabel Matthews-Schlinzig, Jörg Schuster, Gesa Steinbrink u. Jochen Strobel. Berlin/Boston: De Gruyter 2020, S. 276–290.
- Wiethölter, Waltraud u. Anne Bohnenkamp (Hg.):** Der Brief, Ereignis & Objekt. Berlin: Stroemfeld, 2010.