

social communities und Lernern, kann die automatisierten und quantitativen Verfahren anreichern.

Aus diesen Synergieeffekten ergibt sich der Anschluss an Perspektiven zum Verhältnis von (erklärender) Annotation und quantitativen Methoden des *distant reading*. Es ist denkbar, dass Datenbanken es künftig ermöglichen, Vorgänge des erklärenden Annotierens zu automatisieren, etwa in Fällen von Ambiguität, deren Semantik erkannt und die entsprechend aufgelöst wird. In diesem Zusammenhang bedarf er der Ergänzung unserer literaturwissenschaftlichen Perspektive durch technische Expertise. Unser Vortrag stellt somit auch Fragen, die insbesondere im Zuge der Tagung diskutiert werden können. Dabei sollen auch Ideen diskutiert werden, welche Möglichkeiten es geben könnte, dass die manuelle Markierung dessen, was erklärend annotiert (also nicht im Sinne von markup) werden soll, entfällt und Erklärungskontexte definiert werden, innerhalb derer eine datenbankgestützte – und damit „automatische“ – Annotation erfolgen kann.

Der Vortrag bewegt sich an der Schnittstelle von Automatisierung und individuellen hermeneutischen Akten und damit entlang des Problems, wie im Markup eines Textes Entscheidungen getroffen werden können, welche Aspekte in einem Text relevant sind und die dem individuellen Text gerecht werden können. Wir möchten verschiedene Fallstudien aus englischsprachigen literarischen Texten vorstellen, etwa anhand von automatisierten Annotationssystemen wie x-ray, die oben geschilderte Probleme exemplarisch aufzeigen, die aber zugleich auch Synergieeffekte deutlich machen. Letztlich sollten bei dem Verhältnis von qualitativen Methoden des *close reading* und quantitativen Herangehensweisen die Nutzerfreundlichkeit sowie die Individualisierung im Vordergrund stehen. Diese Individualisierung ist dabei zweifach: zum einen bezieht sie sich auf Informationen aus Datenbanken, zum anderen auf den Text. Texte, die annotiert werden sollen, müssen mit Datenbanken korrelieren, und die erklärenden Annotationen müssen offenlegen, was im Text gemeint ist sowie was der potentielle Leser erfahren und wissen möchte. Eine solche Anreicherung von Texten ist von quantitativen Methoden bislang nicht zu leisten; umgekehrt sind qualitative Methoden momentan dadurch eingeschränkt, dass sie sich manueller Verfahren bedienen müssen. Das Zusammenspiel von Quantität und Qualität, von *close* und *distant reading*, erklärender Annotation und computergestützter Textanalyse öffnet neue Perspektiven im Bereich der Digital Humanities und kann auch einen Beitrag zu fächerübergreifenden Paradigmen leisten.

## Bibliographie

**Battestin, Martin C.** (1981): „A Rationale of Literary Annotation: The Example of Fielding’s Novels“, in: *Studies in Bibliography* 34: 1-22.

**Bauer, Matthias / Angelika Zirker** (2015): „Whipping Boys Explained: Literary Annotation and Digital Humanities“, in: Siemens, Ray / Price, Kenneth M. (eds): *Literary Studies in the Digital Age: An Evolving Anthology* <http://dlsanthology.commons.mla.org/under-review-matthias-bauer-and-angelika-zirker-whipping-boys-explained-literary-annotation-and-digital-humanities/> [letzter Zugriff 09. Januar 2016].

**Cummings, James** (2013): „The Text Encoding Initiative and the Study of Literature“, in: Siemens, Ray / Schreibman, Susan (eds.): *A Companion to Digital Literary Studies*. Oxford: Blackwell 451-76.

**Drucker, Johanna** (2012): „Humanistic Theory and Digital Scholarship“, in: Gold, Matthew K. (ed.): *Debates in the Digital Humanities*. Minneapolis: University of Minnesota Press 85-95.

**Jannidis, Fotis** (2010): „Methoden der computergestützten Textanalyse“, in: Nünning, Vera (ed.): *Methoden der literatur- und kulturwissenschaftlichen Textanalyse*. Stuttgart: Metzler 109-132.

**Jannidis, Fotis / Flanders, Julia** (2015): „Knowledge Organization and Data Modeling in the Humanities. A whitepaper“ <http://www.wwp.northeastern.edu/outreach/conference/kodm2012/> [letzter Zugriff 09. Januar 2016].

**Jannidis, Fotis / Krug, Markus / Toepfer, Martin / Puppe, Frank / Reger, Isabella / Weimer, Lukas** (2015): „Automatische Erkennung von Figuren in deutschsprachigen Romanen“. Abstract für die DHd 2015 in Graz.

**McCarty, Willard** (2012): „Collaborative Research in the Digital Humanities“, in: Deegan, Marilyn / McCarty, Willard (eds.): *Collaborative Research in the Digital Humanities*. Farnham: Ashgate 1-10.

**Meister, Jan-Christoph** (2012): „Crowd Sourcing ‘True Meaning’: A Collaborative Approach to Textual Interpretation“, in: Deegan, Marilyn / McCarty, Willard (eds.): *Collaborative Research in the Digital Humanities*. Farnham: Ashgate 105-122.

**Moretti, Franco** (2013): *Distant Reading*. New York: Verso.

**Stroud, Matthew D.** (2006): „The Closest Reading: Creating Annotated Online Editions“, in: Bass, Laura R. / Greer, Margaret R. (eds.): *Approaches to Teaching Early Modern Spanish Drama*. New York: The MLA of America 214-129.

## Emosaic Visualisierung von Emotionen in Texten durch Farbumwandlung zur Analyse und Exploration

**von Lupin, Martin**

[martin.von.lupin@fh-potsdam.de](mailto:martin.von.lupin@fh-potsdam.de)

FH Potsdam, Deutschland

## Geuder, Philipp

philipp.geuder@fh-potsdam.de  
FH Potsdam, Deutschland

## Leidinger, Marie-Claire

marie.leidinger@fh-potsdam.de  
FH Potsdam, Deutschland

## Schröder, Tobias

schroeder@fh-potsdam.de  
FH Potsdam, Deutschland

## Dörk, Marian

doerk@fh-potsdam.de  
FH Potsdam, Deutschland

## Abstract

Das computergestützte Extrahieren und Visualisieren von Emotionen in Texten ist eine etablierte Technik des "Distant Reading". Die generelle Stimmung eines Textes kann schnell erfasst werden ohne den gesamten Text lesen zu müssen. Da Emotionen sehr komplex und die Eigenschaften zwischen verschiedenen Emotionen fließend sind, ist die visuelle Charakterisierung von Emotionen schwierig. Wir stellen *Emosaic* vor, ein Online-Tool welches Emotionen aus benutzerdefinierten Texten filtert und durch systematische und nachvollziehbare Farbumwandlung zur Exploration und Analyse innerhalb einer interaktiven Visualisierung bereitstellt. Die durch drei Dimensionen beschreibbaren Emotionen werden dabei in klar definierte Farbparameter übersetzt. Ein von uns entwickelter öffentlich zugänglicher Web- Prototyp ( vgl. Geuder et al. 2014-) zeigt anhand interaktiver Visualisierungen erste Analyse- und Explorationsmöglichkeiten dieser Methode.

## Einleitung

Wörtern wurden bereits in verschiedenen Studien Emotionen anhand von Emotionsdimensionen wie Valenz (V, engl. valence), Erregung (A, engl. arousal) und Dominanz (D, engl. dominance) zugewiesen (Osgood 1962). Diese Dimensionen sind eine Möglichkeit, Emotionen anhand numerischer Werte zu beschreiben und nach der Affect Control Theory Grundlage aller sozialen Interaktionen (Rogers et al. 2013). Mithilfe der in diesem Forschungskontext empirisch entstandenen Emotionswörterbücher können einem Text somit verschiedene Emotionen numerisch zugeordnet werden. Die dreidimensionale numerische Beschreibung von

Emotionen ist hingegen dem intuitiven Verständnis nicht leicht zugänglich. Eine visuelle Entsprechung für sämtliche Emotionszustände würde helfen, die mehrdimensionale Emotionalität eines Textes schnell erfassbar und zudem explorierbar zu machen.

## Verwandte Arbeiten

Die Analyse und Visualisierung von benutzergenerierten Inhalten ist ein spannendes und aktuelles Forschungsgebiet. Sowohl wegen des Entstehens großer Textmengen durch Trends wie Online-Weblogs, durch maschinell auslesbare Schnittstellen zu solchen Diensten und durch die Leistungssteigerung von Hardware und Software, hat das Interesse daran zugenommen.

**We feel fine** (Kamvar et al. 2011) ist ein interaktives Tool welches die Exploration von Emotionen in Weblogs ermöglicht mit dem Ziel den Menschen dabei behilflich zu sein sich selbst bzw. andere Personen besser zu verstehen. Die Ausgabe erfolgt anhand einer experimentellen Visualisierung. Es kann als Suchmaschine verstanden werden, welche das Web nach Emotionen durchsucht. Jede gefundene Emotion wird als Kreis dargestellt. Die Füllfarbe ist abhängig von der zu repräsentierenden Emotion. Glückliche Emotionen werden durch ein helles Gelb vertreten, während Ärger durch ein Rot dargestellt wird. Diese Farbkodierung hilft dem Benutzer bei der groben Unterscheidung von Emotionen, es können diese jedoch nicht weitergehend charakterisiert werden.

**We feel** (Milne et al. 2015) ist ein webbasiertes Tool welches Emotionen in sozialen Medien verfolgt und aufzeichnet. Das Ziel ist die Aufzeichnung der Weltstimmung bzw. der Stimmung eines Landes um die Verbreitung von physischen Problemen zu erforschen. Die Ausgabe erfolgt anhand interaktiver Graphen und Diagramme. Die Emotionen werden durch deren Namen in Textform dargestellt. Eine Farbkodierung innerhalb der Diagramme hilft bei der Orientierung wobei auch hier keine differenzierte Identifikation von Emotionen bzw. Nuancen zwischen ähnlichen Emotionen möglich ist.

Während beide Tools als Datengrundlage Emotionen aus Webeinträgen nutzen, ist es hingegen unser Ziel die benutzerdefinierte Eingabe eines Textes zu ermöglichen. Zusätzlich sollen sowohl differenzierte als auch nachvollziehbare farbliche Repräsentationen für Emotionen verwendet werden. Die Farbkodierung von Emotionen zur besseren Orientierung innerhalb visueller Darstellungen scheint eine etablierte Methode zu sein, jedoch lassen die bisher vorliegenden Kodierungssysteme keine fundierte Emotionsbeschreibung zu. *Emosaic* hingegen nutzt eine direkte Umwandlung der drei Emotionsebenen in klar definierte Farbparameter, sodass Worte nicht notwendig sind, um Emotionen in Texten unterscheiden und klassifizieren zu können.

## Methodisches Vorgehen

### Farbzuweisung

Da die Emotionalität eines Wortes nach der Affect Control Theory durch genau einen Punkt im dreidimensionalen Emotionsraum beschrieben werden kann, ist eine Farbzuweisung anhand eines dreiparametrischen Farbraums plausibel. Nach mehreren Iterationen in verschiedenen Farbräumen, wählten wir schließlich den HSV-Farbraum. Die drei Parameter des HSV-Farbraums weisen einen hohen eigenständigen Einfluss auf die resultierende Farbe auf (Farbwert, Farbsättigung und Hellwert). Da der Einfluss eines Parameters auf das Farbergebnis klar sichtbar ist, eignet sich dieser Farbraum gut für die Farbumwandlung der drei Emotionsebenen. Von Probanden als plausibel wahrgenommene Ergebnisse erzielten wir, wenn Valenz den Farbwert bestimmt (H, engl. hue), Erregung die Farbsättigung (S, engl. saturation) und Dominanz den Hellwert (V, engl. value).

Während die Sättigung und die Helligkeit Minima und Maxima analog zu Dominanz und Erregung beschreiben, stellt der Farbwert einen kontinuierlichen Farbverlauf dar. Um eine Farbwertannäherung an den Rändern zu vermeiden, haben wir bei der Farbzuweisung einen Farbwertbereich bewusst ausgespart, sodass eine Grenze zwischen Minima und Maxima deutlich hervortritt. Blau entspricht dem Minimum, rot dem Maximum und grün einem mittleren neutralen Valenzwert.

Für die Beurteilung der Zuordnung orientierten wir uns an den sechs Basisemotionen (Liebe, Überraschung, Freude, Wut, Trauer und Angst), wobei wir die Zuweisung der Emotionsdimensionen auf die Farbdimensionen so wählten, dass Liebe einem Rot- / Pinkton entspricht, um der tradierten Farb-Emotions-Zuweisung in der westlichen Kultur zu entsprechen (Abbildung 1a). Eine informelle Studie zeigte, dass diese Form der Zuordnung als intuitiv bewertet wurde. Daneben zeigte sich in Übereinstimmung mit unserer Farbzuweisung, dass negative Gefühle eher dunkel sind und kühlen Farbtönen wie blau oder grün zugeordnet werden (Abbildung 1b), dagegen positive Gefühle eher hell sind und mit warmen Farbtönen wie gelb oder orange in Verbindung gebracht werden (Abbildung 1c).

### Datengrundlage für die Charakterisierung der Emotionen

Wir verwenden für die Farbumwandlung das ANEW-Wörterbuch (Warriner et. al 2013) mit über 13.000 englischen Wörtern. Das Wörterbuch wurde ebenfalls vom Projekt "We feel" verwendet.

## Möglichkeiten der Farbübersetzung

Durch die Farbübersetzung können emotionale Stimmungen in Texten visuell miteinander verglichen werden. Zusätzlich können Emotionen eines Textes in verschiedenen Ebenen analysiert werden. Sowohl die Gesamtstimmung eines Textes als auch einzelne Sätze können in den Fokus gerückt werden. Stimmungsänderungen im Text können so visuell dargestellt werden. Zudem ist das Filtern von Emotionen möglich. Die farbliche Kodierung unterstützt dabei die Regulierung der Filter.

### Aufbau und Funktionsweise des Tools

Grundlegend für die Funktionsweise des Tools ist die Eingabe eines Textes. Der Nutzer kann aus vorgegebenen Texten verschiedenster Länge wählen oder einen eigenen Text innerhalb eines Textfeldes platzieren. Nach der serverseitigen Textanalyse sind verschiedene statische und interaktive Darstellungen verfügbar. Die Darstellung der Emotionsanalyse teilt sich in drei Bereiche auf: Makroansicht, Textansicht und Mikroansicht (Abbildung 2). Die drei Bereiche sind miteinander verlinkt. Grundlegend für die dynamische Änderung einer Ansicht ist die Auswahl von einzelnen Emotionen bzw. Wörtern oder einem Bereich innerhalb einer Emotionsdimension.

Die **Textansicht** ist der zentrale Bereich des Tools (Abbildung 3b). Oberhalb befindet sich ein Histogramm mit permanenter Positionierung und darunter der zu Beginn eingegebene Text, welcher durch Scrollen in voller Länge gelesen werden kann. Ist noch keine Auswahl getroffen, werden alle emotionsrelevanten Wörter innerhalb des Textes durch Hinterlegung mit der korrespondierenden Farbe hervorgehoben. Bei einer Auswahl tritt die Farbhinterlegung von Wörtern, welche außerhalb der Auswahl liegen, in den Hintergrund. Im Histogramm werden alle im Text vorkommenden Farben angeordnet. Die x-Achse kann mit einer der Emotionsdimensionen belegt werden. Durch diesen Ansicht wird die Verteilung innerhalb der Dimensionen sichtbar. Das Histogramm dient zum einen als emotionaler Fingerabdruck und zum anderen als Auswahlwerkzeug. Der Nutzer hat die Möglichkeit durch Klicken und Ziehen einen Bereich im Histogramm auszuwählen. Hierdurch verändert sich dynamisch die Auswahl an Emotionsworten. Durch dieses Brushing ändern sich entsprechende Elemente in der Makro- und Mikroansicht. Alternativ zur Mehrfachauswahl kann eine Einzelauswahl durch einen Mausklick auf das entsprechende Wort vorgenommen werden.

Die **Makroansicht** bietet einen ersten Überblick über allgemeine Emotionstendenzen und -entwicklungen im

Text (Abbildung 3a). Mit dem links platzierten Diagramm kann untersucht werden, inwiefern sich die einzelnen Werte (V, A, D) innerhalb des Textes verändern und wie sie in Beziehung zueinander stehen. Hat der Nutzer eine Auswahl getroffen bietet der Index daneben eine Übersicht darüber, wo sich die entsprechenden Textstellen befinden. Klicken auf den Index ermöglicht schnelles Springen zur entsprechenden Textpassage.

Die **Mikroansicht** ermöglicht das Explorieren des Textes im Detail (Abbildung 3c). Die Wörter innerhalb der getroffenen Auswahl werden hier aufgelistet. Zu der Emotionsfarbe und der Häufigkeit des Vorkommens erfährt der Nutzer hier auch die Zusammensetzung aus den vad-Werten. Hat der Nutzer nur ein Emotionswort gewählt, werden zusätzlich emotionsverwandte Worte aufgelistet, um weiteres Explorieren des Textes zu ermöglichen.

## Zukünftige Arbeiten

Die Analyse und Exploration von Emotionen in Texten anhand von Farben, basierend auf deren Emotionsdimensionen ist ein innovativer Ansatz zur Textanalyse. Der Umgang durch Nutzer mit unserem Tool ist dabei Bestand weiterer Untersuchungen. Hierzu ist bereits eine langfristig angesetzte Evaluation gestartet worden, die neben automatisch generierten Parametern der Texte wie Länge und Emotionalität auch von Nutzern angegebene Feedback wie zum Beispiel dem Verwendungszweck des Tool beinhaltet. Auf Basis der Evaluationsergebnisse planen wir Erkenntnisse zu der Wirkung von Farbumwandlung von Emotionen und dem generellen emotionsbezogenen Interessensfokus von Nutzern zu gewinnen. Limitiert durch unser verwendetes Wörterbuch ist die Analyse momentan ausschließlich auf englische Texte beschränkt. Wir planen weitere Sprachen zu integrieren, insofern ähnliche Wörterbücher für diese Sprachen zur Verfügung stehen. Zudem muss ein möglicher Mehrwert neuer bzw. alternativer Textanalysemethoden untersucht werden. Da Farben und deren emotionale Empfindung kulturabhängig sind, wäre die Untersuchung von Sprachen aus verschiedenen Kulturräumen ebenfalls besonders interessant. Des Weiteren stellt sich die Frage, wie mit der Kombination von Adjektiven und bedeutungsverändernden Partikeln („a little bit“, „very“, „not“, „only“) und der damit einhergehenden Veränderung des emotionalen Gehalts des Wortes umgegangen wird.

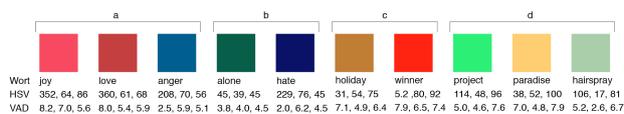


Abbildung 1: Farbumwandlung verschiedener Wörter inklusive der Farbwerte (HSV) und Emotionsdimensionen (VAD): Ausschnitt der Basemotionen (a), negative Wörter (b), positive Wörter (c), auffällige Farben (d).

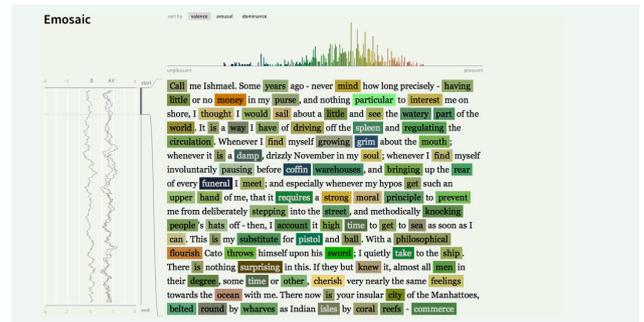


Abbildung 2: Ansicht nach Eingabe des zu analysierenden Textes.

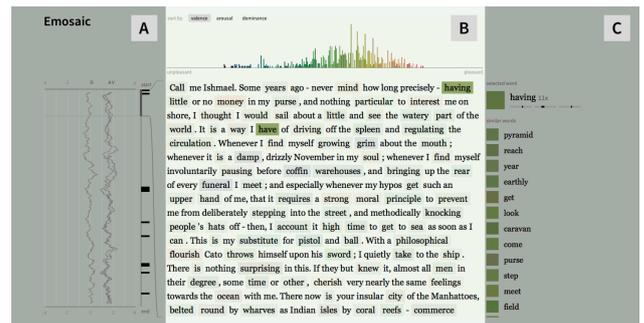


Abbildung 3: Darstellung des Tools nach Auswahl eines einzelnen Wortes. Die grauen Farbübersichtungen dienen hier zur Erklärung der verschiedenen Bereiche: (A) Makroansicht mit Werteverlauf innerhalb des Textes von valence, arousal und dominance und Index zur Lokalisierung der Textpassagen, (B) Histogramm und eingabebezogener Text mit farblich hervorgehobenen emotionsrelevanten Wörtern, (C) Mikroansicht mit ausgewähltem Wort und verwandten Wörtern innerhalb des Textes.

## Bibliographie

- Geuder, Philipp / Leidinger, Marie-Claire / von Lupin, Martin** (2014-): *emosaic* <http://emosaic.de/> [letzter Zugriff 10. Februar 2016].
- Kamvar, Sepandar D. / Harris, Jonathan** (2011): "We feel fine and searching the emotional web", in: *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*. ACM 117–126.
- Milne, David / Paris, Cecile / Christensen, Helen / Batterham, Philip / O'Dea, Bridianne** (2015): "We Feel: Taking the emotional pulse of the world", in: *Proceedings of the 19th Triennial Congress of the International Ergonomics Association (IEA 2015)*.
- Moretti, Franco** (2013): *Distant Reading*. London / New York: Verso.
- Osgood, Charles E.** (1962): "Studies on the generality of affective meaning systems", in: *American Psychologist* 17, 1: 10-28.
- Rogers, Kimberly B. / Schröder, Tobias / von Scheve, Christian** (2014): "Dissecting the sociality of emotion: A multilevel approach", in: *Emotion Review* 6: 24-33.
- Warriner, Amy Beth / Kuperman, Victor / Brysbaert, Marc** (2013): "Norms of valence, arousal, and dominance for 13,915 English lemmas", in: *Behavior Research Methods* 45: 1191-1207.