

# GOEDOC – Dokumenten- und Publikationsserver der Georg-August-Universität Göttingen

---

---

2018

## Tagungsbericht über die Digital Humanities 2018 – Konferenz vom 25. –29. Juni 2018 in Mexiko-Stadt

Benjamin Krautter<sup>1</sup> Sandra Schell<sup>2</sup>

<sup>1</sup>Universität Stuttgart / QuaDramA

<sup>2</sup>Universität Heidelberg

DARIAH-DE Working Papers

Nr.28

Krautter, Benjamin; Schell, Sandra: Tagungsbericht über die Digital Humanities 2018 : Konferenz vom 25. –29. Juni 2018 in Mexiko-Stadt  
Göttingen : GOEDOC, Dokumenten- und Publikationsserver der Georg-August-Universität, 2018  
(DARIAH-DE working papers 28)

Verfügbar:

PURL: <http://resolver.sub.uni-goettingen.de/purl/?dariah-2018-3>

URN: <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2018-3-5>

## Bibliographische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

*Erschienen in der Reihe*  
DARIAH-DE working papers

ISSN: 2198-4670

*Herausgeber der Reihe*  
DARIAH-DE, Niedersächsische Staats- und Universitätsbibliothek

Mirjam Blümm, Thomas Kollatz, Stefan Schmunk und Christof Schöch

---

---

**Abstract:** Die diesjährige internationale Digital Humanities Konferenz gastierte vom 25. bis zum 29. Juni 2018 in Mexiko-Stadt. Der interdisziplinäre Anspruch der Tagung fokussierte in diesem Jahr insbesondere Wissensbestände ethnischer Minderheiten und multilinguale Präsentationsformen der Beiträge. Ziel war es, eine plurale Diskussionsplattform für möglichst viele WissenschaftlerInnen zu bieten, die im weiten Feld der digitalen Geisteswissenschaften forschen. Der folgende Tagungsbericht möchte einerseits einen Einblick in diese besondere Rahmung der DH2018 bieten und andererseits eine Auswahl literaturwissenschaftsaffiner Beiträge der Konferenz vorstellen. Wir diskutieren die Beiträge, sie entstammen unterschiedlichen Präsentationsformen (Poster, Long Paper, Short Paper, Panel Session), unter drei Forschungsperspektiven: Methodologie, Aufbereitung digitaler Korpora und Bestände sowie Anwendung digitaler Methoden bei der Textanalyse.

**Keywords:** Digital Humanities, DH2018, Literaturwissenschaft, Tagungsbericht, Methodologie, digitale Korpora, Textanalyse

Digital Humanities, DH2018, literary studies, conference report, methodology, digital corpora, text analysis

# Tagungsbericht über die Digital Humanities 2018

Konferenz vom 25.–29. Juni 2018 in Mexiko-Stadt

Benjamin Krautter<sup>1</sup>      Sandra Schell<sup>2</sup>

<sup>1</sup>Universität Stuttgart / QuaDramA

<sup>2</sup>Universität Heidelberg



Benjamin Krautter, Sandra Schell: „Tagungsbericht über die Digital Humanities 2018“. [DARIAH-DE Working Papers](#) Nr. 28. Göttingen: DARIAH-DE, 2018. URN: [urn:nbn:de:gbv:7-dariah-2018-3-5](#).

Dieser Beitrag erscheint unter der  
Lizenz [Creative-Commons Attribution 4.0](#) (CC-BY).

Die *DARIAH-DE Working Papers* werden von Mirjam Blümm,  
Thomas Kollatz, Stefan Schmunk und Christof Schöch  
herausgegeben.



## Zusammenfassung

Die diesjährige internationale Digital Humanities Konferenz gastierte vom 25. bis zum 29. Juni 2018 in Mexiko-Stadt. Der interdisziplinäre Anspruch der Tagung fokussierte in diesem Jahr insbesondere Wissensbestände ethnischer Minderheiten und multilinguale Präsentationsformen der Beiträge. Ziel war es, eine plurale Diskussionsplattform für möglichst viele WissenschaftlerInnen zu bieten, die im weiten Feld der digitalen Geisteswissenschaften forschen. Der folgende Tagungsbericht möchte einerseits einen Einblick in diese besondere Rahmung der DH2018 bieten und andererseits eine Auswahl literaturwissenschaftsaffiner Beiträge der Konferenz vorstellen. Wir diskutieren die Beiträge, sie entstammen unterschiedlichen Präsentationsformen (Poster, Long Paper, Short Paper, Panel Session), unter drei Forschungsperspektiven: Methodologie, Aufbereitung digitaler Korpora und Bestände sowie Anwendung digitaler Methoden bei der Textanalyse.

## Schlagwörter

Digital Humanities, DH2018, Literaturwissenschaft, Tagungsbericht, Methodologie, digitale Korpora, Textanalyse

## Keywords

Digital Humanities, DH2018, literary studies, conference report, methodology, digital corpora, text analysis

## Inhaltsverzeichnis

1	Einführung	4
2	Methodologie	6
3	Aufbereitung digitaler Korpora und Bestände	9
4	Anwendung digitaler Methoden	11
5	Schlussbemerkungen	13
6	Literaturverzeichnis	16

## 1 Einführung

Inmitten der Hochphase von Fußball-Weltmeisterschaft und mexikanischem Präsidentschaftswahlkampf diskutierten vom 25. bis 29. Juni 2018 über 600 internationale WissenschaftlerInnen die neuesten Entwicklungen im breiten Feld der Digital Humanities. Das für die DH2018 in Mexiko-Stadt gewählte Motto „Puentes–Bridges“ akzentuierte nicht nur den interdisziplinären Anspruch dieser Tagung, sondern implizierte bereits politische Signalwirkung.<sup>1</sup> Erstmals im ‚globalen Süden‘ zu Gast war die diesjährige DH ein „momentous achievement for an organization that has always strived to be truly global, diverse, and inclusive“, wie die OrganisatorInnen Élika Ortega, Glen Worthey, Isabel Galina und Ernesto Priani betonten.<sup>2</sup>

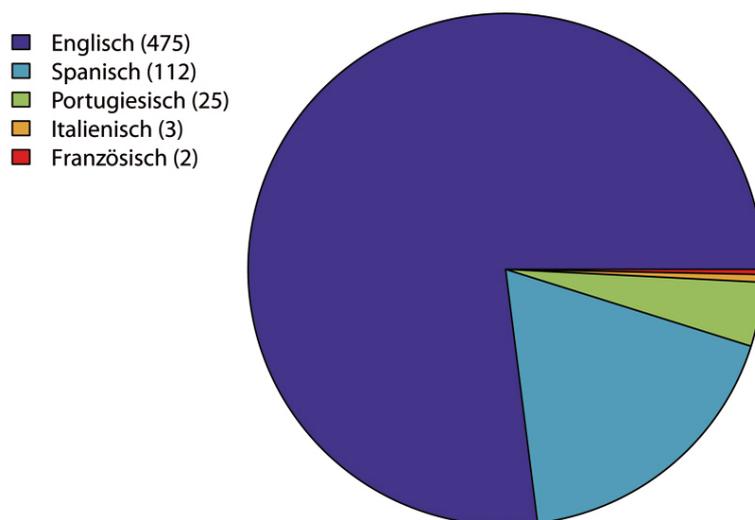


Abbildung 1: Einreichungen für die DH2018 nach Sprachen.

Der multilinguale Charakter der Konferenz stand exemplarisch für die angestrebte Diversität. Zum wiederholten Mal wurde auch ForscherInnen eine Stimme gegeben, die sich nicht auf Englisch vorstellen konnten oder wollten: Etwa 20% der Poster, Vorträge und Panels wurden in einer anderen Sprache präsentiert. In der eröffnenden (spanischen) und abschließenden (englischen) Keynote verdeutlichte sich nicht nur die Multilingualität – beide Vorträge wurden souverän simultanübersetzt –, sondern auch einer weiteren Ausdeutung des diesjährigen Mottos wurde Aufmerksamkeit verliehen: indigenen Völkern, ihren Sprachen und Dialekten.

<sup>1</sup>Exemplarisch für die meisten *Twitter*-Reaktionen auf die Konferenzeröffnung steht Kyle K. Courtney, der in Anlehnung an Isabel Galinas Begrüßung den folgenden Tweet verfasste: „Theme of building bridges, NOT walls [...] No disciplinary walls, nor academic walls, nor global walls. We need bridges & never walls, never cages. Families belong together“ <https://twitter.com/KyleKCourtney/status/1011733975143010304> (letzter Zugriff: 17.07.2018).

<sup>2</sup>Ortega, Élika, Glen Worthey, Isabel Galina und Ernesto Priani: *Welcome to DH2018*. In: Book of Abstracts Digital Humanities 2018. URL: <https://dh2018.adho.org/en/abstracts/> (letzter Zugriff: 27.07.2018).

**Janet Chávez Santiago** verknüpfte in ihrer eröffnenden Keynote *Tramando la palabra (Weaving the World)* die in ihrem Heimatdorf weitverbreitete Handarbeit des Webens mit einer Metapher für indigene Sprachen und Kulturen. Diese werden häufig als ausgestorben wahrgenommen – der Wikipedia-Eintrag zu Santiagos Muttersprache ist bspw. in Vergangenheitsform verfasst –, die allgemeine Abwesenheit indigener Kulturen in den Medien negiere zugleich ihre Existenz, so Santiagos Beobachtung. Ihr Anliegen sei es, der digitalen Welt die Dokumentation und die didaktische Aufbereitung ihres Heimatdialekts Zapotec zugänglich zu machen: „We as speakers of indigenous languages must truly appropriate these spaces, to weave our word well, in order to liberate ourselves from the denial of the present.“<sup>3</sup> **Schuyler Esprit** berichtete in der abschließenden Keynote *Digital Experimentation, Courageous Citizenship and Caribbean Futurism*, in welchem Maße Hurrikan Maria (Sep.-Okt. 2017) ihre Vorstellung von Heimat und Heimatlosigkeit veränderte. Mit *Carisealand* stellte sie ein Projekt vor, das sich nachhaltig den traumatischen Nachwirkungen der Verwüstung widmen möchte: „Rebuilding is not limited to material things, it is about rebuilding an intangible spirit that necessitates the input of people all around the world.“<sup>4</sup> Esprits Projekt möchte auf die Schlüsselfunktion von Sprache und (kulturellen) Narrativen im Prozess des Wiederaufbaus hinweisen. Digitale Methoden haben dabei v.a. dokumentarische und didaktische Funktionen: Wie man indigene und afrikanische Wissensbestände und Bewusstsein auf das Digitale überträgt, um Veränderungsprozesse zu mobilisieren, wird hierbei für Esprit zur Leitfrage.

Santiagos und Esprits Aktivismus in unterschiedlichen Gebieten des ‚globalen Südens‘ – in Oaxaca (Mexiko) und der Dominikanischen Republik –, die zudem von Kolonial- und Neokolonialmächten beeinflusst sind, rahmten die Konferenz, eröffneten dem Publikum eine spezifische Perspektive auf die Welt und zeigten, wie digitale Projekte diese Gebiete im Kleinen gestalten können. Das Plädoyer, politische, ethnische und gesellschaftliche Minderheiten anzuerkennen und ihre Wissensbestände für neue Denkansätze in einem globalen Kontext zu nutzen, wurde in sechs Sektionen<sup>5</sup> und etwa 10% der Poster diskutiert. Insgesamt präsentierten WissenschaftlerInnen aus über 20 Nationen Forschungsbeiträge mit fast 900 beteiligten AutorInnen,<sup>6</sup> die in einem Peer Review Verfahren von 633 GutachterInnen ausgewählt wurden. Hierfür boten 16 Workshops, 22 Long-Paper und 22 Short-Paper Sessions, 33 Panels und 116 Poster in bis zu sieben parallelaufenden Sektionen eine Plattform. Mehr als 13.750 Tweets unter dem Hashtag DH2018<sup>7</sup> begleiteten den Konferenzzeitraum und wurden weit über den Tagungs-Mikrokosmos des Sheraton Hotels hinausgetragen und rezipierbar gemacht. Unser Tagungsbericht möchte diese Rezeption ergänzen und zudem einen fokussierten Blick auf eine Auswahl literaturwissenschaftsaffiner Beiträge der DH2018 werfen. Unter den drei Forschungsperspektiven *Methodologie, Aufbereitung digitaler Korpora und Bestände* sowie *Anwendung digitaler Methoden* werden wir im Folgenden Beiträge unterschiedlicher Präsentationsformen vorstellen und diskutieren.<sup>8</sup>

---

<sup>3</sup>Santiago, Janet Chávez: *Weaving the World*. In: Book of Abstracts Digital Humanities 2018, S. 30.

<sup>4</sup>Vortragszitat.

<sup>5</sup>Hierzu zählen wir u.a. die Sektionen *Indigenous languages* und *DH in the Global South*, und einzelne Beiträge wie, *Bridging Cultures Through Mapping Practices: Space and Power in Asia and America*, *Las Humanidades Digitales en la Mixteca de Oaxaca: reflexiones y proyecciones sobre la Herencia Viva o Patrimonio*, *Digital Decolonizations: Remediating the Popol Wuj*.

<sup>6</sup>Im Folgenden beschränken wir uns auf die Nennung der Vortragenden und verweisen mit u.a. auf Co-AutorInnen, die sich im *Book of Abstract* oder Konferenzprogramm vollständig nachvollziehen lassen.

<sup>7</sup>Vgl. [https://epriego.blog/2018/06/30/tweets-per-user\\_lang-in-a-dh018-archive/](https://epriego.blog/2018/06/30/tweets-per-user_lang-in-a-dh018-archive/) (letzter Zugriff: 27.07.2008).

<sup>8</sup>Die nachfolgend diskutierten Beiträge sind als kleine Auswahl des sehr dichten und vielseitigen Konferenzprogramms zu verstehen. Sie wurden aus einer größtenteils literaturwissenschaftlichen Perspektive gewählt.

## 2 Methodologie

Ein großer Teil der literaturwissenschaftlich perspektivierten Vorträge konzentrierte sich auf methodologische Implikationen bei der computergestützten Textanalyse. Im Vordergrund stand dabei zumeist das zu vertiefende Verständnis und weniger das eigentliche Anwendungspotential der vorgestellten Methoden, die oftmals, etwa bei der stilometrischen Autorschaftsattributions, bereits in einer konkreten Forschungspraxis etabliert sind. So waren es häufig Klassifikationsaufgaben – welcher Gattung ist ein Text zuzuordnen, welche Nationalität hat ein Autor/eine Autorin, welche literarische Figuren sind die ProtagonistInnen? –, die den Antrieb methodologischer Innovation bildeten. Doch während die Attribution von Autorschaft oder die Klassifikation von Nationalität auf gesicherten Daten getestet werden kann, ist schon die Annotation literarischer Kategorien, etwa der Gattung eines Textes, nicht trivial – man denke hier bspw. an die Vielzahl verschiedener Novellendefinitionen. Solche Kategorien sind häufig an besonders exemplarischen Prototypen ausgerichtet und können verschiedene Textmerkmale ins Zentrum rücken – etwa Thema, Struktur, Stil, Personal, Länge.

**David L. Hoovers** (New York) Vortrag fokussierte auf eine Gegenüberstellung verschiedener Methoden der Autorschaftsattributions viktorianischer Dramen. Ausgangspunkt seiner Analysen war die von Brian Vickers forcierte These, dass Wort-Ngramme besser geeignet seien als die Frequenz einzelner Wörter, um die Autorschaft frühneuzeitlicher Dramen zu bestimmen. Grund dafür sei gemäß Vickers insb. die idiomatische Natur der Sprache. Hoover gelang es nicht, diese These auf seinem Korpus von 125 viktorianischen Dramen plausibel zu machen. Denn sowohl Wortfrequenzen als auch Zeichen-Ngramme erzielen bedeutend bessere empirische Testergebnisse als Wort-Ngramme. Hoover versuchte dabei die Studie von Vickers möglichst akkurat zu reproduzieren, nutzte also bspw. ebenfalls Dramensegmente mit einer Länge von 2.600 Wörtern als Testkorpus. Die Frequenz seltener 3-6-Wort-Ngramme, die Vickers aufgrund der Sequentialität von Sprache favorisiert, führe zu einem „spectacular failure“ bei der tatsächlichen Attribution. Hoover bilanzierte, dass die eigentlich überzeugenden a priori Argumente für Wort-Ngramme, keine empirische Bestätigung fänden.

In Übereinstimmung mit forensischer Praxis schlug **Patrick Juola** (Pittsburgh, PA) bereits 2015 vor, in Fragen der Autorschaftszuschreibung einem formalen Protokoll zu folgen. Eine leicht modifizierte Version dieses Protokolls evaluierte Juola in seinem Vortrag anhand von 4.000 AutorInnen des „Blog Authorship Corpus“, das 140 Millionen Wörter und Texte von 20.000 BloggerInnen enthält. Dem zu bestimmenden Textsegment (100 Sätze) werden neben einem weiteren Segment desselben Autors/derselben Autorin zehn zufällig gewählte AutorInnen als Vergleichskorpus zur Seite gestellt. Fünf verschiedene und idealiter voneinander unabhängige Analysen (Wortschatzübereinstimmung, Wortlängen, Zeichen-4-Gramme, Zeichensetzung, 50 mfw) werden einzeln betrachtet und in eine Rangfolge gebracht. Aus diesen Rangfolgen ergibt sich durch Aufsummieren ein Ähnlichkeitsmaß, das wiederum in p-Werte umgewandelt wird. Juola konstatierte, dass das Protokoll sehr gute Ergebnisse für die Identifikation von Textsegmenten gleicher AutorInnen liefere. Die p-Werte für unterschiedliche AutorInnen signalisierten jedoch, dass die fünf durchgeführten Analysen in einer gegenseitigen Abhängigkeit stehen könnten, durchschnittliche Wortlängen also bspw. mit den meistfrequenten Wörtern zusammenhängen. Denn die p-Werte von unterschiedlichen AutorInnen sind nicht gleichmäßig verteilt, wie es bei zufälligen Relationen der Einzelanalysen anzunehmen wäre, sondern favorisieren die beiden Extrempunkte der Werteskala.

Die Netzwerkanalyse dramatischer Werke hat eine hohe Ausgangsplausibilität, die sich v.a. aus der starken Strukturiertheit der Texte und der zentralen Stellung der Figurenrede ergibt. Um den von Manfred Pfister geprägten Begriff der ‚quantitativen Dominanzrelation‘ neu zu perspektivieren, gemeint sind damit quantitativ erfassbare Relationen des Dramenpersonals, nutzten **Frank Fischer** (Moskau) u.a. für ihren Vortrag einen mehrdimensionalen Ansatz. Dieser kombinierte fünf Netzwerkmaße (Degree; Weighted Degree; Betweenness Centrality; Closeness Centrality; Eigenvector Centrality) mit drei zählbasierten Werten (Anzahl der gesprochenen Worte einer Figur; Anzahl der Äußerungen und Anzahl der Szenen, in denen eine Figur präsent ist) und fand Anwendung auf einem Korpus von 465 deutschsprachigen Texten. Einzelne Fallbeispiele konnten zeigen, dass die methodische Mehrdimensionalität v.a. bei weniger tragenden Figuren („middle characters“) zu interessanten Verschiebungen zwischen Netzwerkmaßen und zählbasierten Werten führt. Die nach dieser Metrik als ProtagonistInnen identifizierten Figuren – ein Goldstandard existiert hierfür nicht – rangieren dagegen in 50% der Fälle zumindest in sieben der acht Rankings an erster Stelle. Für den historischen Verlauf von 1730 bis 1930 konstatieren Fischer u.a. eine Abnahme quantitativ dominanter Figuren. Fischer u.a. nutzten die acht quantitativen Messwerte auch, um eine Typologisierung der Figurenverteilung in Dramen mithilfe einer Regressionsanalyse vorzunehmen. Entgegen ihrer ursprünglichen Erwartung scheint eher eine quadratische Verteilung der Werte denn eine *Power Law-Verteilung* vorzuliegen.

**Lawrence Evalyn** (Toronto) u.a. testeten die Möglichkeiten sozialer Netzwerkanalysen für die Gattungsbestimmung von William Shakespeares Dramen. Personalwechsel auf der Bühne, die über maschinenlesbare Auf- und Abtrittsmarker automatisch erkannt werden, grenzen dabei die bemessenen Einheiten voneinander ab. Die für die Netzwerkanalysen zentrale Figurenpräsenz kann so potentiell noch feingranularer ausgewertet werden. Evalyn u.a. ordneten die 37 Stücke Shakespeares zu drei Untergruppen: Tragödien, Komödien und Historien. Tragikomödien wie *The Comedy of Errors* werden dabei den Komödien zugeschrieben. Wiederum scheint ein multidimensionaler Ansatz – bis zu 17 Netzwerkmaße finden Anwendung – am vielversprechendsten für die Auflösung. Verschiedene Kombinationen dieser Maße erreichen bei der Gattungsklassifikation eine Accuracy von 100%. Netzwerkgrafiken, so Evalyns Fazit, könnten als Proxy für Handlung aufgefasst werden. Die gezeigten Netzwerke waren jedoch immer eine Zustandsbeschreibung des kompletten Dramentexts, dessen Handlung zu einem statischen Netzwerk aufsummiert wird.

Ein ähnliches Ziel verfolgte **Ulrike Henny-Krahmer** (Würzburg). Sie nutzte Sentiment-Analysen – mit der Hilfe von Emotions- oder Valenzwörterbüchern werden Wörter, Sätze oder größere Textabschnitte als graduell positiv oder negativ eingestuft und Grundemotionen zugeordnet –, für die Genreklassifikation spanisch-amerikanischer Romane. Zwei Hypothesen markierten ihren Ausgangspunkt: (1) Grad und Art der Emotionen würden sich in verschiedenen Genres unterscheiden und (2) es mache einen Unterschied, ob Emotionen in direkter Figurenrede oder vom Erzähler wiedergegeben werden. Ihr vorgestelltes Textkorpus umfasste 16 AutorInnen und 30 Romane, die zwischen 1840 und 1910 geschrieben wurden. Henny-Krahmer ordnete die Romane in vier Subgenres: rührende (sentimental), historische, soziopolitische und Sittenromane („Costumbrista“). Für die Klassifikation testete sie verschiedene Emotionswörterbücher und Parameter. Ihr bestes Klassifikationsresultat mit einem Decision Tree erreichte einen F1-Wert von 0,66 und nutzte das NRC Emotion Lexicon. Tests mit Support Vector Machines, die für Klassifikationsaufgaben häufig dem Decision Tree vorgezogen werden, erzielten erstaunlicherweise nur unwesentlich bessere Ergebnisse. Eine mögliche Erklärung steckt in den Genrezuschreibungen selbst, die möglicherweise nicht trennscharf genug sind.

Eine erste Version ihres *Ontological Model for Inferring Psychological Profiles and Narrative Roles of Characters* (OLC) zeigte das Poster von **Mattia Egloff** (Lausanne) u.a. Ausgehend von linguistischen Beschreibungen der Charakterzüge von Figuren in literarischen Texten (z.B. durch Adjektive) untersucht das Tool deren Rolle für die Handlung. OLC ist als integratives Modell angelegt, das ontologische Modelle (*LEMON, OE, Plutchik, Hourglass*) mit dem psychologischen Modell der Big Five verbindet, um literarische Charaktere verschiedenen Figurentypen zuordnen zu können – vorgestellt wurden *Hero, Anti-Hero* und *Villain*. Als Beispiel diente Shakespeares Figur *Shylock*, der über die Charakterisierung im Nebentext automatisch einer der Big Five und daraus folgend dem Figurentypen *Anti-Hero* zugeordnet werden konnte.

Was ist Distinktivität? Mit dieser Frage eröffneten **Christof Schöch** (Trier), **Daniel Schlör** (Würzburg) u.a. ihren Vortrag, der sich zur Aufgabe machte, das Distinktivitätsmaß Zeta und einige Varianten zu evaluieren. Zeta wurde 2007 von John Burrows eingeführt, um die Unterscheidungskraft textueller Merkmale, etwa Wortformen oder Lemmata, in zwei Textgruppen zu bestimmen. Da Zeta auf einem Dispersionsmaß beruht, favorisiere es weniger frequente Inhaltswörter, wodurch die Ergebnisse eine hohe Interpretierbarkeit versprechen. Um das Distinktivitätsmaß besser zu verstehen, untersuchten Schöch, Schlör u.a. verschiedene Zeta-Varianten und den Einfluss wichtiger Parameter, etwa der herangezogenen Segmentlänge. Die vielversprechendsten Resultate erzielt eine log<sub>2</sub>-transformierte Zeta-Variante.<sup>9</sup> Die Klassifikationsergebnisse – spanischsprachige Romane sollten nach ihrem Herkunftskontinent (Amerika oder Europa) klassifiziert werden – versprechen eine verbesserte Leistungsfähigkeit und größere Robustheit bei verschiedenen Segmentlängen. Unklar bleibt bislang, ob die besseren Klassifikationsergebnisse zu Lasten der Interpretierbarkeit gehen. Denn durch die log<sub>2</sub>-Transformation können auch Textmerkmale mit geringen Segmentanteilen hohe Zeta Werte erzielen. Das habe zur Folge, dass zum Teil Orts- oder Personennamen eines Texts als distinktiv für eine ganze Gruppe von Texten erachtet werden.

**Maciej Eder** (Krakau) verfolgte in seinem Vortrag das Ziel, die Dynamiken sprachlicher Veränderungen zu modellieren. Grundlage seiner Untersuchung war das „Corpus of Historical American English“ (Texte zwischen 1810–2009, insg. 400 Millionen Wörter). Eder ging es dabei nicht um die Bestätigung bereits selektierter a priori Annahmen, sondern um eine Trendsuche ohne vorherige Auswahl. Um das zu erreichen, nutzte er die 1.000 häufigsten Wortformen (Lemmata) für eine iterative Klassifikation. Ausgangspunkt ist ein Jahr *x*, das hypothetisch einen linguistischen Wandel im Korpus markiere. Daran anschließend wird – mit einer Unterbrechung von zehn Jahren – ein Zeitraum von 20 Jahren vor und nach dem gewählten Jahr abgedeckt. Beiden Teilmengen werden 500 zufällig gewählte Textbeispiele mit einer Länge von 1000 Tokens zugeteilt. Die Teilmengen werden wiederum halbiert, wodurch Trainings- und Testdaten entstehen, auf Basis derer eine überwachte Klassifikation (Nearest Shrunken Centroids) durchgeführt wird. Sprachlicher Wandel lasse sich dann an der Genauigkeit der Klassifikation bemessen. Die Klassifikationsergebnisse bekräftigten, dass sprachlicher Wandel nicht linear, sondern in Perioden mit beschleunigter und verlangsamer Veränderung verlaufe. Verantwortlich für den Wandel zeigen sich dabei eher eine Vielzahl kleiner Merkmale als wenige, aber dafür starke Prädikatoren. Aus Eders Sicht eher kontraintuitiv sind unter den relevantesten Merkmalen viele Funktionswörter (*the, and, that* usw.) zu beobachten. Die Nutzung von Funktionswörtern scheint also stärkeren Schwankungen zu unterliegen als die von thematischen Wörtern.

---

<sup>9</sup>Die Segment Proportions werden, wie von Burrows vorgeschlagen, aber weiterhin subtrahiert.

Insgesamt betrachtet ließ sich ein größtenteils reflektierter Umgang mit Methoden der Textanalyse beobachten, auch wenn das angedeutete Erklärungspotential in manchen Fällen überpointiert wurde. Limitierungen und Schwierigkeiten einzelner Methoden wurden ebenso diskutiert wie deren Vorzüge und Perspektiven. Problematischer erscheint die Operationalisierung literaturwissenschaftlicher Konzepte, die sich oftmals nur indirekt oder gar nicht in messbare Verhältnisse übertragen lassen. Gleiches gilt auch für deren Annotation. Ein klarer methodologischer Schwerpunkt der DH2018 lag auf der Netzwerkanalyse, die vielfach Anwendung fand. Verglichen mit der DH2017 erlebte die Stilometrie eine kleine Renaissance. Topic Modeling und Sentiment Analysen wurden dagegen seltener genutzt, auch im Verhältnis zu früheren Konferenzen. Das mag auch am eher methodologischen Erkenntnisinteresse der Beiträge gelegen haben, da Topic Modeling und Sentiment Analysen oftmals zu einer unüberwachten explorativen Erstsichtung eines Textkorpus herangezogen werden.

### 3 Aufbereitung digitaler Korpora und Bestände

Die digitale Bild-, Text- oder Korpusaufbereitung ist eine wichtige Grundlage, um computergestützte Analysen überhaupt erst zu ermöglichen. Dazu zählt einerseits der Digitalisierungsprozess, also die Erstellung maschinenlesbarer Dokumente, andererseits aber auch die Ausstaffierung dieser Dokumente mit Metadaten. Im Folgenden werden einige der vorgestellten Projekte diskutiert, die sich just diesen grundsätzlichen Fragen widmeten.

Digitalisiertes Archivmaterial durch Textsuche zugänglich zu machen, war das Ziel von **Sofia Ares Oliveira** (Lausanne) u.a. Sie konzentrierten sich dabei auf venezianische Manuskripte des 18. Jahrhunderts, die über zwei Millionen Seiten umfassen. Ein Trainings- und Test-Set von 23.000 Bildsegmenten wurde in den vergangenen zwei Jahren von ArchivarInnen manuell transkribiert. Mittels einer Deep Learning Architektur (Convolutional und Recurrent Neural Networks) entwickelten Oliveira u.a. ein Transkriptionssystem für handgeschriebene Texte. Die Zeichenfehlerrate dieser Architektur liegt bei etwa 8%. Das sei ein guter Wert, wie Oliveira bemerkte, mit dem bspw. die Suche nach Entitäten (mit Fuzzy Matching) gelingen müsste. Ein Crowdsourcing Experiment sollte dieses Ergebnis mit dem durchschnittlichen menschlichen Leser vergleichen. 36 ausgewählte italienischsprechende Freiwillige transkribierten hierfür 8.674 Segmente. Die Zeichenfehlerrate lag dabei, je nach Versuchsaufbau, zwischen 10 und 13%, immer aber über dem entwickelten Transkriptionssystem.

Mit Hilfe von maschinellem Lernen (CRF Modell) näherten sich **Chao-Lin Liu** (Taipeh) u.a. der Satzsegmentierungen in Grabbigraphien der Tang Dynastie an. Klassisches Chinesisch besitzt keine Markierungszeichen für die Trennung von Sätzen, deren Verständnis jedoch maßgeblich von der Satztrennung abhängt. Selbst für Experten sei die Segmentierung keine triviale Aufgabe und immer schon Textinterpretation. Für die automatische Auflösung arbeiteten Liu u.a. mit einem Korpus von 5.119 annotierten Grabbigraphien, das 5.505 einzigartige Schriftzeichen, 1.461.000 Tokens und 424.000 Satzzeichen enthält. Die Satzsegmentierung wurde dabei als Klassifikationsproblem aufgefasst: Folgt auf ein Schriftzeichen ein Satzzeichen oder nicht? Liu u.a. testeten hierfür verschiedene Parameter: die zu betrachtende Kontextgröße, Bigramme im Kontext, Informationen auf Wortebene und Informationen der Aussprache. Die beiden letzteren konnten das Klassifikationsergebnis jedoch nicht verbessern. Der beste F1-Wert lag bei 0,83. Erste Ansätze mit Deep Learning Methoden (BiLSTM) versprechen noch bessere Resultate. Längerfristig soll diese automatische Satzsegmentierung die Zugänglichkeit von klassischen chinesischen Texten in der praktischen Arbeit verbessern.

Unter dem Titel *The Scholar Index* präsentierten **Matteo Romanello** (Lausanne) u.a. eine Plattform, um Literatur und Primärquellen der Geisteswissenschaften über ihre Zitation zu erschließen. Romanello u.a. argumentierten, dass eine umfassende Indexierung von Referenzen und Verweisen ein großes Potential für Suchkapazitäten anbiete und so Quellensammlungen verbinden könne, die momentan getrennt seien. Um die Referenzen automatisch erschließen zu können, müssten diese in zwei Schritten als solche erkannt und anschließend richtig zugeordnet werden (Deep Learning Ansatz mit 40.000 annotierten Referenzen als Trainingsdaten). Die Referenz wird dann als Relation zwischen einer zitierenden Publikation und einer zitierten Quelle aufgefasst. Romanello u.a. zeigten einen ersten Prototyp zur Geschichtsschreibung Venedigs, der über 3.000 Bände listet. Fast vier Millionen extrahierte Referenzen sollen nach und nach eingepflegt werden. Die Metadaten der gelisteten Publikationen sind durchsuchbar, können von den NutzerInnen ergänzt werden und verweisen direkt auf die zitierten Quellen. Auch Publikationsprofile einzelner AutorInnen lassen sich abrufen. Im Idealfall – abhängig von rechtlichen Fragen – soll der *Scholar Index* mit einer digitalen Bibliothek verbunden werden, die den direkten Textzugriff erlaubt.

An der Schnittstelle von Korpusaufbereitung und Anwendungsbeispielen digitaler Methoden präsentierten sich die drei nachfolgend vorgestellten Projekte: **Ethan Reed** (Charlottesville, VA) näherte sich mit Hilfe von Sentiment-Analysen der Bedeutung von ‚Unruhe‘ in Gedichten des *Black Arts Movements*. Die dezidiert politische Ausrichtung der sich in den 1960er und 70er Jahren um Malcom X formierenden Bewegung richtete sich vornehmlich an die ‚Black Community‘ und kommuniziere, so Reeds Hypothese, v.a. über Provokationen und negative Gefühlsäußerungen. Unter der Prämisse, dass Gefühle eine kulturelle Struktur formieren, untersuchte er ein manuell erstelltes Korpus (26 Gedichtbände) mit verschiedenen Sentiment-Klassifikatoren. Ziel war die Annäherung an die Relation zwischen ‚feelings‘ und ‚systemic injustice‘ sowie deren Kodierung von ‚race‘ und ‚gender‘. Reed nutzte die digitalen Methoden größtenteils explorativ, um seinen Blick auf das Korpus neu zu perspektivieren. Seine Analysen führten hierbei jedoch zu einem „interpretative disjoint between the denotative affective import of words [...] and their more nuanced affective import as shaped by poetic, literary, social, and political contexts.“<sup>10</sup> Das Adjektiv ‚black‘ illustriert Reeds Problematik mit Emotionswörterbüchern. Mehrfach wird ‚black‘ von diesen mit teils extrem negativen Emotionswerten belegt, obwohl es in seinem Korpus zumeist positiv als Ausdruck von Gruppenzugehörigkeit gesetzt sei und identitätsstabilisierend wirke.

Hinter ihrem Datenbank-Projekt steht für **Gary Stringer** (Exeter) u.a. ein literatursoziologisches Interesse: *The Poetry of The Lancashire Cotton Famine (1861-65): Tracing Poetic Responses to Economic Disaster*. Für eine frei zugängliche und durchsuchbare Webapp (eXist-db) sammeln Stringer u.a. Gedichte und deren Audioaufnahmen, die als literarische, politisch konnotierte Reaktion auf die ökonomische Krise entstanden sind. Die v.a. in Lokalzeitungen publizierten, meist in Dialekt verfassten Werke geben der Arbeiterklasse eine Stimme, die bis dato in der Forschung kaum berücksichtigt werden konnte. Von der lokalen Katastrophe ausgehend, bietet das Korpus ferner Material – Reflexionen über Arbeit, Armut, Krieg und Sklaverei – für literatursoziologische Fragestellungen, um sowohl das Engagement der Arbeiterklasse als auch die Dichtung im amerikanischen Konflikt der Südstaaten angemessener beschreiben zu können.

---

<sup>10</sup>Reed, Ethan: *Measured Unrest in the Poetry of The Black Arts Movement*. In: Book of Abstracts Digital Humanities 2018, S. 477f., hier S. 478.

**Cole Daniel Crawford** (Harvard) diskutierte in seinem Beitrag *Exploring Big and Boutique Data through Laboring-Class Poets Online* die Bedeutung der Auswertung einzelner Aufzeichnungen und Daten. In Rekurs auf Cheryl Ball u.a.<sup>11</sup> führte er den Begriff der ‚boutique data‘ ein, um „the ongoing importance of small, localized, partial, and qualitative datasets“ hervorzuheben.<sup>12</sup> Ähnlich wie Stringer zielt Crawfords Fallstudie zur Arbeiterdichtung darauf ab, nicht-kanonische AutorInnen in ihren Biographien und Werken der Forschung zugänglich zu machen, um an dieses erstmals erschlossene Korpus literatursoziologische und genreästhetische Fragen zu stellen. Diesen drei Projekten gemeinsam ist das relativ kleine, manuell erstellte Korpus. Die von Stringer und Crawford genutzten Datenbanken verstehen sich zudem als ein kontinuierlich erweiterbares Datenset, das einen „boutique approach to historical and literary information“<sup>13</sup> verfolge, der – im Gegensatz zu ‚big data‘ – Ambiguitäten und Lücken abzubilden vermag und hierdurch Einsichten vermittelt, die GeisteswissenschaftlerInnen das komplexe Verständnis von Raum und Zeit als sich kontinuierlich entwickelnde Ereignisse ermöglichen.

Die Aufbereitung digitaler Bestände wurde in den vorgestellten Beiträgen mit diversen und teilweise innovativen Methoden realisiert. Nur eine unter vielen Visionen für die Zukunft könnten multimediale Datenbanken sein, die in Korpora verschiedene Medienformate zusammenführen. Damit ließen sich Themengebiete, die medial unterschiedlich repräsentiert werden, auch auf größeren Skalen vergleichbar machen. Methodologisch erscheint die Zusammenführung verschiedener Techniken ebenfalls vielversprechend – zu denken wäre an die Kombination von Distant Reading- und Distant Viewing-Techniken, die bspw. ein Filmskript und die visuelle Gestaltung des Films analysieren.

## 4 Anwendung digitaler Methoden

Der folgende Abschnitt rückt Anwendungsbeispiele in den Vordergrund, die sich literaturwissenschaftlichen oder literatur- und sprachgeschichtlichen Fragestellungen annäherten oder an diese anknüpften. Die Rückbindung neuer Erkenntnisse an Forschungsfragen der Fachdisziplinen bleibt eines der zentralen Anliegen der Digital Humanities, auch wenn die disziplinäre Verortung und institutionelle Verankerung im Wissenschaftsbetrieb noch vielfach unklar ist.

Die Vorträge von **Luis Meneses** (Victoria, BC) u.a. zur bilingualen Lyrik von Pablo Picasso und **Pablo Ruiz Fabo** (Madrid) u.a. zu einem über 4.000 Gedichte umfassenden spanischen Sonettkorpus diskutierten die interpretatorischen Rückschlüsse, die Korpusanalysen für Einzeltexte bieten. Meneses u.a. untersuchten sprachspezifische semantische Konzepte in Picassos spanischer und französischer Lyrik und verfolgten zunächst die Hypothese, dass diese Konzepte sprachgebunden seien. Ein taxonomischer computergestützter Ansatz (Latent Dirichlet Allocation und Term Frequency sowie Inverse Document Frequency) konnte diese Hypothese jedoch widerlegen. Der interpretatorische Mehrwert liege folglich darin, repräsentative Themen und Zusammenhänge zwischen verschiedenen Konzepten und Sprachen

---

<sup>11</sup>Ball, C., T. S. Graban und M. Sidler: *The Boutique is Open: Data for Writing Studies*. In: J. Rice und B. McNely (Hrsg.): *Networked Humanities: Within and Without the University* [in Vorbereitung].

<sup>12</sup>Crawford, Cole Daniel: *Exploring Big and Boutique Data through Laboring-Class Poets Online*. In: *Book of Abstracts Digital Humanities 2018*, S. 353f., hier S. 353.

<sup>13</sup>Ebd., S. 354.

zu identifizieren und lokalisieren.<sup>14</sup> Die Frage nach wiederkehrenden poetischen Phänomenen stellten auch Fabo u.a. an ihr ‚Spanish Golden-Age‘-Sonettkorpus (4.000 Einzeltexte, über 1.000 spanische AutorInnen). Allerdings konzentrierten sie ihre Untersuchung auf Kriterien der Genredefinition, der Metrik und des Reimschemas. Erste Ergebnisse bestätigen laut Fabo u.a. die in der Forschung etablierte metrische Beschreibung spanischer Sonette, deuten jedoch auch auf eine zunehmende metrische Varianz im 19. Jahrhundert hin.

**Chelsea Miya** (Edmonton, AB) präsentierte in ihrem Poster eine stilometrische Einzeltextanalyse zu Herman Melvilles *Moby Dick*, die an eine breite literaturwissenschaftliche Debatte zur Erzählstruktur des Romans anschließt. Melville integriert in die mittleren Abschnitte seines Romans Kapitel zur Walkunde, die nahezu frei von Figuren- oder Handlungsentwicklung sind. Miyas stilometrische Analyse – sie zeigte in Cluster- und Netzwerkdarstellungen, dass den Walkunde-Kapiteln eigene stilometrische Signale zugrunde liegen – scheint die These zu stützen, dass Melville zwei separate Texte möglicherweise sogar zu verschiedenen Zeitpunkten geschrieben und diese letztlich zu einem Roman verbunden habe. Die Ergebnisse könnten aber auch in der Anlage der Textstruktur liegen, also durch den unterschiedlichen Einsatz von homo- und heterodiegetischem Erzähler, Dialogen oder Tempusvariationen erklärt werden.

Ebenfalls in einem Poster veranschaulichte **Gabrielle Kirilloffs** (Lincoln, NE) eine durch Pattern-Matching und Sentiment-Analysen gestützte Untersuchung der rhetorischen Setzung von Leser-Apostrophen in 2.000 modernen, englischsprachigen Romanen. Ihre Korpusuntersuchung bildet eine konsistente Setzung der Apostrophen über eineinhalb Jahrhunderte ab. Apostrophen sind in 90% der Romane präsent, durchschnittlich enthält jeder Roman 49 Apostrophen. Ihre Analyse offenbarte zudem eine Korrelation zwischen Geschlecht des Autors und der Apostrophenfunktion: Autoren nutzen Apostrophen häufiger, sie sprechen den Adressaten dabei eher als „reader“ denn als „you“ an. Apostrophen von Autorinnen weisen dagegen deutlich positivere Emotionswerte auf.

In ihrem digitalen Visualisierungsprojekt untersuchte **Randa El Khatib** (Victoria, BC) *The Moral Geography of Milton's ‚Paradise Lost‘*. Sie versteht das epische Gedicht als eine Art „geospatial, text-to-map“<sup>15</sup> Projekt angelegt, das sie auf historischen und biblischen Karten rekonstruiert und abzubilden versucht. Die Komplexität des Projekts entstehe durch Miltons Raumsemantik, die sich multitemporaler (biblischer, historischer und zeitgenössischer) Referenzen bediene, so Khatib. Die Visualisierung erfolgt daher über verschiedene Ebenen der Karten. Hinweise zu intertextuellen Referenzen, Stellenkommentare und Zitate ergänzen die eingezeichneten Handlungsorte auf den Karten und machen Khatibs Projekt zu einem interpretationsstarken Analysetool.

Ausgehend von Edward W. Saids Monographie *On Late Style* warf **Jonathan Pearce Reeve** (New York) die Frage auf: *Does „Late Style“ Exist?* Mit Hilfe von Stilometrie und Document Embeddings ging Reeve der Hypothese nach, dass Künstler gegen Ende ihrer Karriere ‚neue Wege‘ einzuschlagen scheinen. Seine zwölf Korpora umfassen AutorInnen wie Marcel Proust, Thomas Mann, Charles Dickens und Joseph Conrad. Reeve problematisierte dabei nicht nur Saids mehrdeutige Auffassung von ‚late‘, sondern auch seinen Stilbegriff, der stark an Themen gebunden sei.

---

<sup>14</sup>Vgl. Mallen, Enrique und Luis Meneses: *Distinctions between Conceptual Domains in the Bilingual Poetry of Pablo Picasso*. In: Book of Abstracts Digital Humanities 2018, S. 434–436, hier S. 436.

<sup>15</sup>El Khatib, Randa: *The Moral Geography of Milton's ‚Paradise Lost‘*. In: Book of Abstracts Digital Humanities 2018, S. 365f., hier S. 365.

Ob ‚late‘ die Nähe zum Tod, die Vorkenntnis des Todes, ein bestimmtes Alter des Künstlers oder den Umfang des Œuvres betrifft, ließ Reeve aber selbst offen. Aus seinen computergestützten Analysen schloss er indes, dass nicht der späte Stil, sondern eher die frühen Texte eines Autors/einer Autorin besonders unterscheidungstragend seien.

Lässt sich die Handlung eines Erzähltextes über Formen visualisieren? **David William McClure** (Cambridge, MA) und **Scott Enderle** (Philadelphia, PA) verfolgten die Idee, Wortverteilungen in großen Korpora im Verlauf der Erzählzeit zu modellieren. Um sich dieser Fragestellung anzunähern, untersuchten McClure und Enderle ein Korpus von ungefähr 50.000 Romanen. Umso größer das Korpus, so die Vermutung, umso deutlicher zeichnen sich Trends in der Verteilung einzelner Wörter ab. Sie versuchten dabei systematisch die Wörter zu identifizieren, die am unregelmäßigsten über die Erzählzeit verteilt sind, die also bspw. eine besondere Rolle für Anfang, Mitte oder Ende eines Erzähltextes einnehmen. Die Verteilung von Inhaltswörtern folge größtenteils unserer Intuition, folgerten McClure und Enderle. Das Ende sei etwa gekennzeichnet durch Heirat, Tod und den Ausdruck von Emotionen. Zu interessanteren Erkenntnissen könnten Funktionswörter führen: etwa die Verteilung von unbestimmtem und bestimmtem Artikel. McClure und Enderle konnten zeigen, dass beide zu Beginn der Erzählzeit überrepräsentiert sind. Während die Frequenz von ‚the‘ sehr viel schneller auf ein konstantes Niveau abfällt und zum Ende nochmals merkbar steigt, sinkt ‚a‘ zuerst langsamer, aber dafür kontinuierlich, ehe die Frequenz im letzten Fünftel der Erzählzeit deutlich einbricht. Während der unbestimmte Artikel also die Unterschiede zwischen Anfang und Ende von Erzähltexten betone, zeige der bestimmte Artikel deren Gemeinsamkeiten.

## 5 Schlussbemerkungen

Die Digital Humanities zeigen sich als diverses Forschungsfeld, das unterschiedlichste Fachdisziplinen unter dem Dach der Digitalität zusammenführt und deren Austausch anregt. Wie vielfältig die Forschungsbeiträge auf der DH2018 tatsächlich waren, verrät bereits ein Blick in das Programm.<sup>16</sup> Sessions zu *Digital Archaeology and Ethnography* stehen neben Themenfeldern wie *Digital Approaches to Poetry, Performing Arts, Sound, Media* oder *3D, Augmented Reality*. Die Auswertung aller Beitragstitel zeigt, wie viele verschiedene Fachdisziplinen, Objektgegenstandsbereiche und methodologische Ansätze tatsächlich aufeinandertrafen. Auch unser schmaler Auszug literatur- bzw. textwissenschaftlich geprägter Beiträge konnte zeigen, wie divers Zielsetzung und Anwendung von digitalen Methoden im Bereich der Analyse und Aufbereitung von (literarischen) Texten ausfällt. Diese Interessenspluralität birgt jedoch auch Implikationen für die disziplinäre Verortung der Digital Humanities. Wer ist letztlich der anvisierte Adressatenkreis digital forschender GeisteswissenschaftlerInnen und was ist das Objekt des angestrebten Erkenntnisgewinns? Hat die Rückbindung an hermeneutische Fragestellungen und Terminologien der Fachdisziplinen Priorität oder reicht methodologische Innovation als Konsens für eine eigenständige Disziplin vielfältigster Gegenstandsbereiche?

---

<sup>16</sup>Die nachfolgend aufgeführten Zahlen sind dem offiziellen Programm der DH2018 entnommen.



Abbildung 2: Wordcloud aller Beitragstitel, 80 mfw, die Stoppwortliste beinhaltet auch die Begriffe „Digital“, „Humanities“ und einige ihrer Varianten.

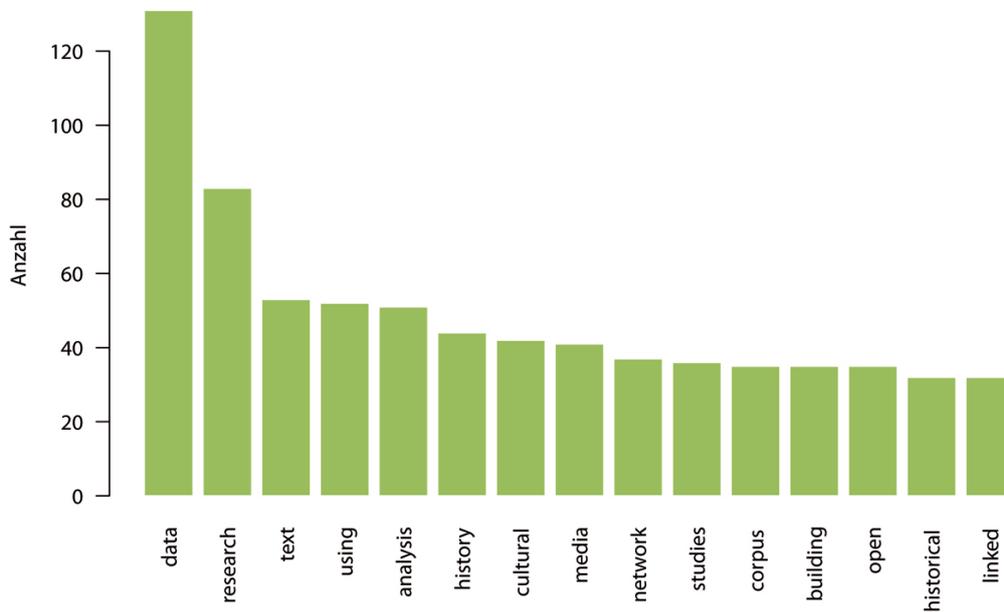


Abbildung 3: Die 15 häufigsten Wörter der Wordcloud.

Dieser von den OrganisatorInnen postulierten und von den TeilnehmerInnen beobachteten Vielfalt steht die institutionelle Herkunft der Beitrags-AutorInnen selbst entgegen, die zu einem sehr großen Teil (38,1%) in den USA beschäftigt sind. Insgesamt waren zwar Institutionen aus 39 Ländern an den 330 Beiträgen beteiligt, die Kluft zwischen den USA und den übrigen Nationen bleibt dennoch bemerkenswert groß. Der Austragungsort in Mexiko-Stadt ermöglichte, dass TeilnehmerInnen aus Zentral- und Südamerika – insbes. aus Mexiko und Kolumbien – sehr viel präsenter als auf früheren DH-Konferenzen waren.<sup>17</sup> Dadurch gewann auch der multilinguale Charakter der Konferenz, der bereits in der Tagungseröffnung eine zentrale Stellung einnahm, an Signifikanz.

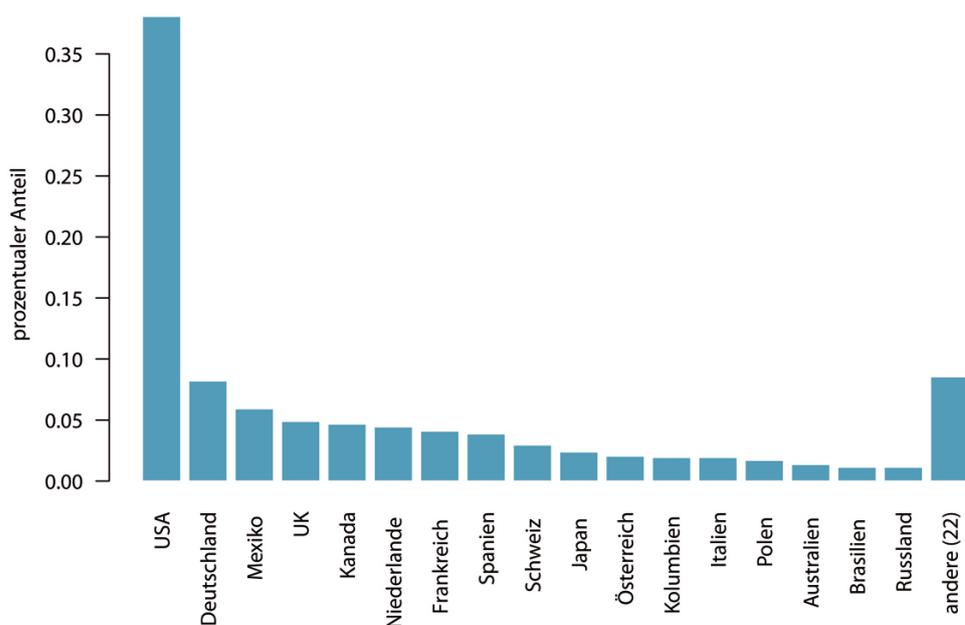


Abbildung 4: Länder, in denen die Institutionen der WissenschaftlerInnen ansässig sind.

Erstmals nutzte eine DH-Konferenz Double-Blind Peer Reviews – sowohl die Identität der AutorInnen als auch der GutachterInnen wird hierbei geheim gehalten –, um die Einreichungen zu begutachten. Ca. 54% aller Einreichungen wurden für die Konferenz angenommen, verglichen mit den Jahren 2013 bis 2015 ist das eine niedrige Akzeptanzquote.<sup>18</sup> Der Anteil beteiligter Frauen an angenommenen Einreichungen stieg jedoch zugleich auf 43,6%.

Ausgehend von diesen ‚hard facts‘ konnten die OrganisatorInnen also mit einigem Recht behaupten: „Mexico’s sociocultural diversity makes it an ideal location for converging DH from distinct cultures, contexts, and socio-political realities. We believe that our steps towards bridging cultural, technological, political, and ideological borders will lead to the creation of a DH community that is truly global, diverse, and inclusive.“<sup>19</sup>

<sup>17</sup>Vgl. Scott B. Weingarts Blog: <http://scottbot.net/submissions-to-digital-humanities-2014-pt-3/> (letzter Zugriff: 26.7.2018).

<sup>18</sup>Von 2013 bis 2015 lag die Akzeptanzquote zwischen 59 und 72%. Vgl. Scott B. Weingarts Blog: <http://scottbot.net/acceptances-to-digital-humanities-2015-part-1/> (letzter Zugriff: 27.07.2018).

<sup>19</sup>Ortega, Élika, Glen Worthey, Isabel Galina und Ernesto Priani: *Welcome to DH2018*. In: *Book of Abstracts Digital Humanities 2018*.

## 6 Literaturverzeichnis

Ball, C., T. S. Graban und M. Sidler: *The Boutique is Open: Data for Writing Studies*. In: J. Rice und B. McNely (Hrsg.): *Networked Humanities: Within and Without the University* [in Vorbereitung].

Courtney, Kyle K.: Tweet vom 26.06.2018. URL: <https://twitter.com/KyleKCourtney/status/1011733975143010304> (letzter Zugriff: 17.07.2018).

Crawford, Cole Daniel: *Exploring Big and Boutique Data through Laboring-Class Poets Online*. In: *Book of Abstracts Digital Humanities 2018*, S. 353f. URL: <https://dh2018.adho.org/en/abstracts/> (letzter Zugriff: 27.07.2018).

El Khatib, Randa: *The Moral Geography of Milton's ‚Paradise Lost‘*. In: *Book of Abstracts Digital Humanities 2018*, S. 365f. URL: <https://dh2018.adho.org/en/abstracts/> (letzter Zugriff: 27.07.2018).

Mallen, Enrique und Luis Meneses: *Distinctions between Conceptual Domains in the Bilingual Poetry of Pablo Picasso*. In: *Book of Abstracts Digital Humanities 2018*, S. 434–436. URL: <https://dh2018.adho.org/en/abstracts/> (letzter Zugriff: 27.07.2018).

Ortega, Élika, Glen Worthey, Isabel Galina und Ernesto Priani: *Welcome to DH2018*. In: *Book of Abstracts Digital Humanities 2018*, o. S. URL: <https://dh2018.adho.org/en/abstracts/> (letzter Zugriff: 27.07.2018).

Priego, Ernesto: *Tweets per User\_Lang in a #DH2018 Archive*. Blogeintrag vom 30.06.2018. URL: [https://epriego.blog/2018/06/30/tweets-per-user\\_lang-in-a-dh018-archive/](https://epriego.blog/2018/06/30/tweets-per-user_lang-in-a-dh018-archive/) (letzter Zugriff: 27.07.2018).

Reed, Ethan: *Measured Unrest in The Poetry of the Black Arts Movement*. In: *Book of Abstracts Digital Humanities 2018*, S. 477f. URL: <https://dh2018.adho.org/en/abstracts/> (letzter Zugriff: 27.07.2018).

Santiago, Janet Chávez: *Weaving the World*. In: *Book of Abstracts Digital Humanities 2018*, S. 30. URL: <https://dh2018.adho.org/en/abstracts/> (letzter Zugriff: 27.07.2018).

Weingart, Scott B.: *Acceptances to Digital Humanities 2015 (part 1)*. Blogeintrag vom 24.06.2015. URL: <http://scottbot.net/acceptances-to-digital-humanities-2015-part-1/> (letzter Zugriff: 27.07.2018).

Weingart, Scott B.: *Submissions to Digital Humanities 2015 (part 3)*. Blogeintrag vom 07.11.2014. URL: <http://scottbot.net/submissions-to-digital-humanities-2014-pt-3/> (letzter Zugriff: 26.7.2018).