

Automatische Bild-Text-Analyse: Chancen für die Zeitschriftenforschung jenseits von reinen Textdaten

Rißler-Pipka, Nanette

nanette.rissler-pipka@ku.de

Katholische Universität Eichstätt-Ingolstadt, Deutschland

Chandna, Swati

swati.chandna@kit.edu

Karlsruhe Institute of Technology, Deutschland

Tonne, Danah

danah.tonne@kit.edu

Karlsruhe Institute of Technology, Deutschland

Zeitschriften und multimodale Wahrnehmung

Gerade die Epoche der Moderne (1850-1945) geht mit einer Veränderung der menschlichen Wahrnehmung einher, als mit den aufkommenden Avantgardebewegungen (sowie im Zuge der technischen, ökonomischen und sozialen Weiterentwicklung) weltweit die Anzahl der Kulturzeitschriften explosionsartig steigt. Nicht so sehr die vielzitierte technische Reproduzierbarkeit des Kunstwerks (Benjamin 1935-39) zeugt von dieser Veränderung als vielmehr die Fähigkeit zur multimodalen Wahrnehmung, heraus gebildet durch die dynamische Medienlandschaft zu der neben Photographie und Film vor allem die Zeitschriften gehören. Für Lateinamerika vergleicht Raquel Macciuci die Rezeption der modernen Presse und Zeitschriften mit der Nutzung digitaler Medien (Macciuci 2015: 209), während für den deutschsprachigen Raum gleich zwei Tagungen zu Zeitschriften deren Multimodalität auf der einen Seite („Illustrierte Zeitschriften um 1900: Multimodalität und Metaisierung“, 2014) und die Funktionsweise der visuellen Zeitschriftenkultur auf der anderen Seite („Deutsche illustrierte Magazine – Journalismus und visuelle Kultur in der Weimarer Republik“, 2013) beleuchten. Das Desiderat der Forschung wird in dieser Beziehung schon 2010 von Frank et al. genau bestimmt:

So wissen wir inzwischen relativ gut darüber Bescheid, welche und wie viele einschlägige Zeitschriften es gab und gibt; weniger schon, wer darin worüber aus welcher Perspektive und mit welcher Stoßrichtung geschrieben

hat; kaum mehr zuletzt aber, welche Gestaltung die Zeitschriften prägte und welcher Stellenwert dergleichen Publizistik zukam. (Frank et al.: 2010, 10)

Es geht auch hier in der Literatur- und Kulturwissenschaft vor allem um Zahlen, Metadaten und eine Möglichkeit die „Gestaltung“ der Zeitschriften quantitativ zu ermesen. Genau diese Felder sind prädestiniert für die Zusammenarbeit mit den DH.

Wenn wir für mehr als 200 Zeitschriftentitel sagen könnten, wie das quantitative Bild-Text-Verhältnis ist und Metadaten dazu vergleichen könnten, wie z.B. die programmatische Ausrichtung des Titels, die Zielgruppe, die Akteure, der Standort und Verbreitungsgrad, dann ließen sich die von Frank et al. gestellten Fragen beantworten. In einer verfeinerten Analyse der Gestaltung kann überprüft werden, ob progressives, avantgardistisches Layout (gekennzeichnet durch viel Leerraum, reduzierte Ornamente, klare Schrifttypen) mit den entsprechend programmatisch eingeordneten Titeln übereinstimmt.

Automatische Bild-Text-Erkennung: ein Versuch mit SWATI

Am konkreten Beispiel mit Daten aus dem Projekt *Revistas culturales 2.0* (Universität Augsburg) wurde vor dem Hintergrund dieses Desiderats der Versuch einer automatischen Bild-Text-Erkennung anhand von zunächst 69 Beispielseiten eines Heftes der argentinischen Kulturzeitschrift „El Hogar“ (Dec. 1919) unternommen. Während im Projekt die Metadaten der Zeitschriften bereits mit digitalen Tools zur Netzwerkvisualisierung (Ehrlicher / Herzgell 2016) oder zur Zeit/Ort-Visualisierung mit dem DARIAH-DE Geobrowser (<http://geobrowser.de.dariah.eu/storage/199501>) für Überblicksdarstellungen analysiert wurden, fehlte nach wie vor die Möglichkeit einer quantitativen Bild-Text-Analyse, die Gestaltung der Zeitschriften entsprechend.

Der technische Ansatz zur Durchführung einer solchen quantitativen Bild-Text-Analyse basiert auf den Entwicklungen des eCodicology Projektes, in dem die Digitalisate des „Virtuellen Skriptoriums St. Matthias“ automatisch ausgewertet und die so erfassten Merkmale in den beschreibenden Metadaten abgelegt und anschließend visualisiert wurden. Auch aus informationstechnologischer Sicht stellt sich die Frage, ob die für mittelalterliche Handschriften verwendeten Verfahren und Algorithmen generisch einsetzbar sind und Potential für alternative Anwendungsfelder, hier am Beispiel spanischsprachiger Magazine untersucht, besitzen. Die Exploration von Erweiterungsmöglichkeiten verspricht ebenfalls zusätzlichen Erkenntnisgewinn für alle beteiligten Disziplinen.

Zur Extraktion der Merkmale wird der SWATI Workflow (Software Workflow for the Automatic Tagging of

Medieval Manuscript Images) genutzt (Chandna et al. 2015). Andere bestehende Methoden zur Image Document Analysis führten nicht zum erwarteten Resultat, da sie nur auf spezielle und kleine Datensätze angewendet werden können (vgl. DIVAServices-Spotlight). Hier wurde der SWATI-Workflow speziell an die heterogene Layout-Struktur der Zeitschriften angepasst.

Als Basis der Untersuchung wurden auf den Digitalisaten der spanischsprachigen Magazine die Seitengröße sowie der Text- und Bildraum vermessen und jeweils die Fläche, Breite, Höhe, Koordinaten der linken oberen Ecke sowie der Neigungswinkel bestimmt.

Da auf einer Seite mehrere Text- und/oder Bildbereiche wie beispielsweise Überschriften, Haupttext, Notizen, Initialen, Zeichnungen oder Glossen auftreten können, werden die genannten Werte für jeden Bereich einzeln ermittelt und gespeichert. Zusätzlich werden die Werte auch als relative Angaben in Prozent unabhängig von der Einheit der Messungen aufgeführt, um die Übertragbarkeit bei unterschiedlichen Auflösungen der Digitalisate zu gewährleisten.

Im Bildbeispiel sind die Ergebnisse der jeweiligen automatischen Bild-Text-Erkennung zu erkennen: Originalbild (links) sowie segmentierte Bild- (Mitte) bzw. Textbereiche (rechts). Zusätzlich wird zur Vermessung des Seitenbereiches jeweils eine Segmentierung der Seite durchgeführt (hier nicht dargestellt)

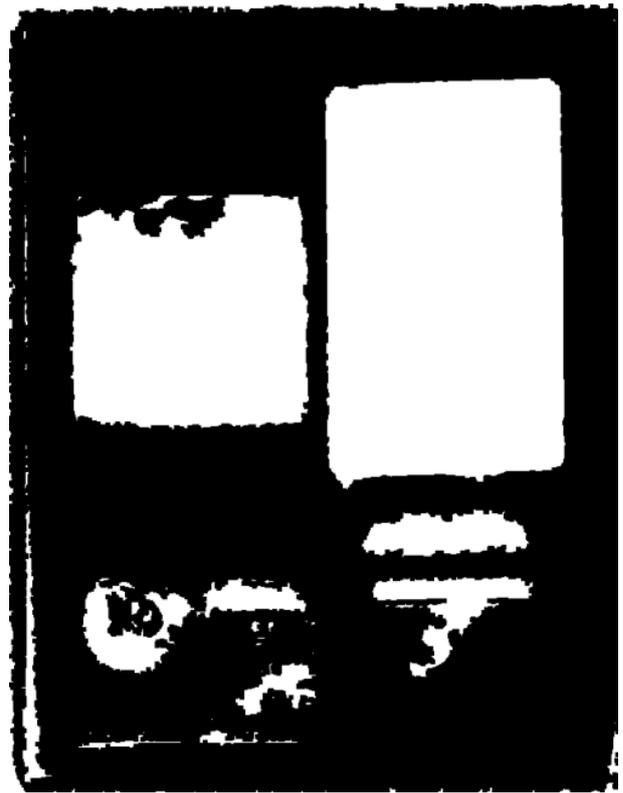


Abbildung 2: Bildererkennung



Abbildung 1: Originalseite „El Hogar“

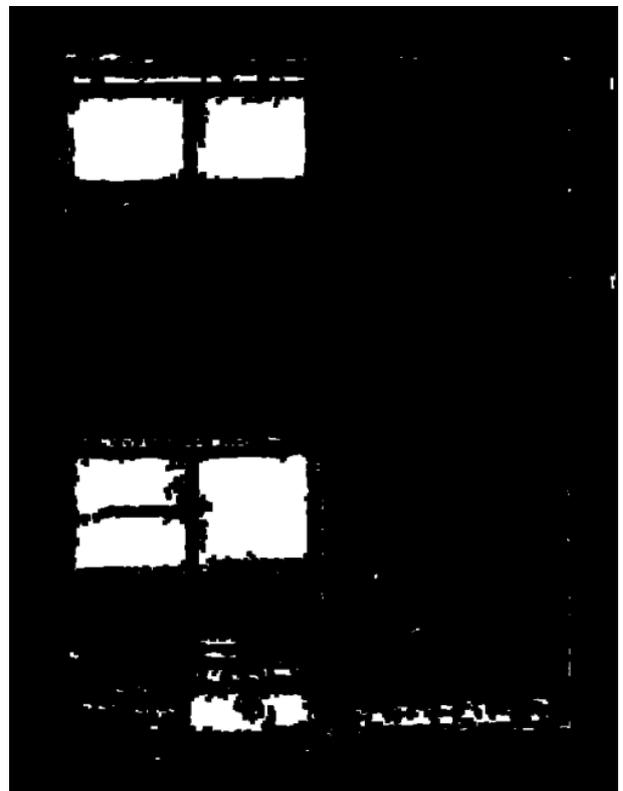


Abbildung 3: Texterkennung



Abbildung 4: Originalseite "El Hogar"

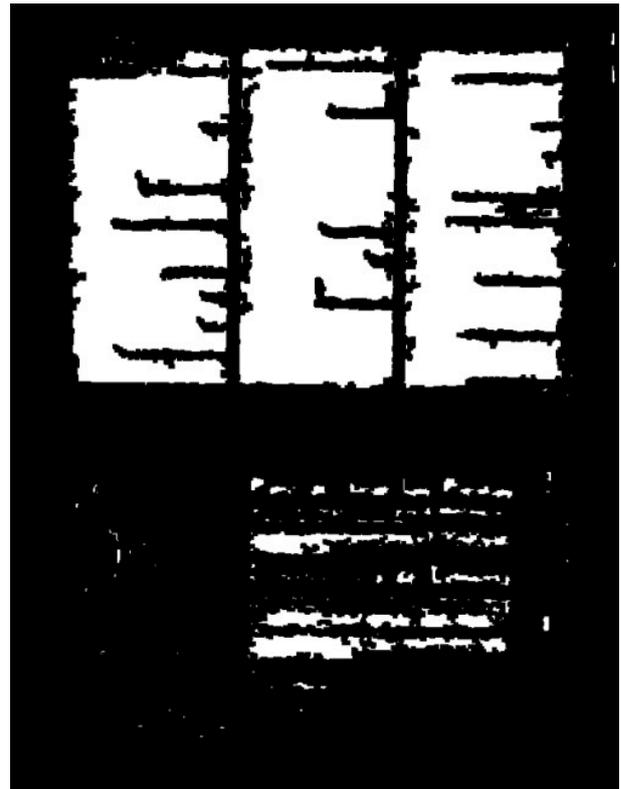


Abbildung 6: Texterkennung

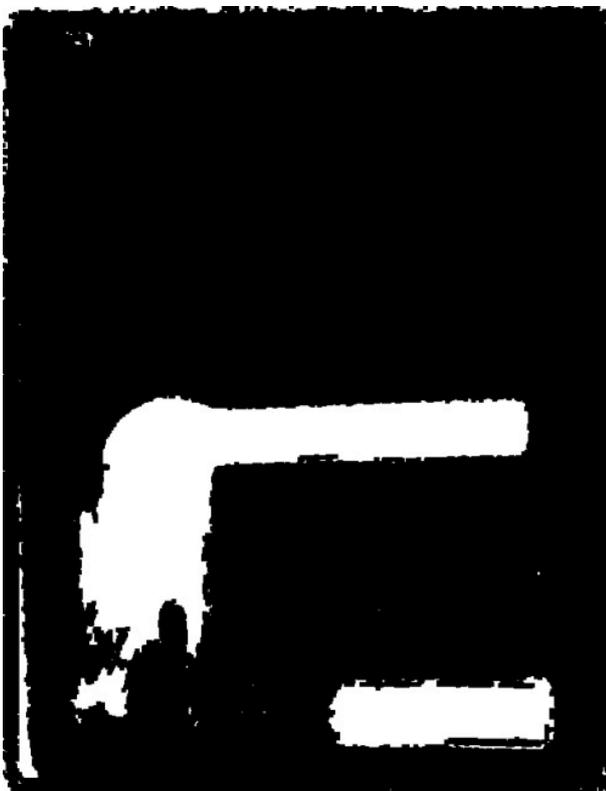


Abbildung 5: Bilderkennung

Wie man an diesen beiden Beispielseiten der Zeitschrift „El Hogar“ (Dec. 1919) erkennt, funktioniert die Bild- und Texterkennung sehr zuverlässig auch bei komplexeren Bild-Text-Gefügen, wie z.B. in Werbeanzeigen. Größere ins Bild integrierte Schriften, z.B. die Werbeüberschrift in Abb. 1 oder 4 wird als Teil des Bildes erkannt. Das ist aus gestaltungsanalytischer Sicht auch nicht falsch, hat die Werbeüberschrift doch gleichzeitig Bild- und Textfunktion.

Ein anderes Problem ist die Zuordnung kleinerer Bild- oder Textflächen zu einer Einheit, z.B. gehört in Abb. 4-5 der Schriftblock am unteren rechten Bildrand zum Werbebild und gibt dem zentral stehenden Text eine Rahmung. Als Messergebnis erscheinen aber zwei verschieden große, unabhängige Text- bzw. Bildbereiche.

Die eigentliche Analyse der gewonnenen Metadaten (Messdaten) ist dann für alle Beteiligten eine erneute Herausforderung, kann aber genau das oben beschriebene Problem der Zuordnung semantisch zusammenhängender Text- und Bildbereiche lösen. Durch eine Filtereinstellung werden bei mehr als 10 erkannten Text/Bildbereichen nur die 10 größten ausgewählt. Diese Selektion vermindert den Einfluss von Messartefakten und erleichtert die statistische Auswertung der Daten sowie deren Visualisierung. Auch dieser Bereich wurde im Zusammenhang des eCodicology Projekts bereits für das Korpus mittelalterlicher Handschriften erprobt (Chandna et al. 2016) und konnte auf das vorliegende Fallbeispiel übertragen werden.

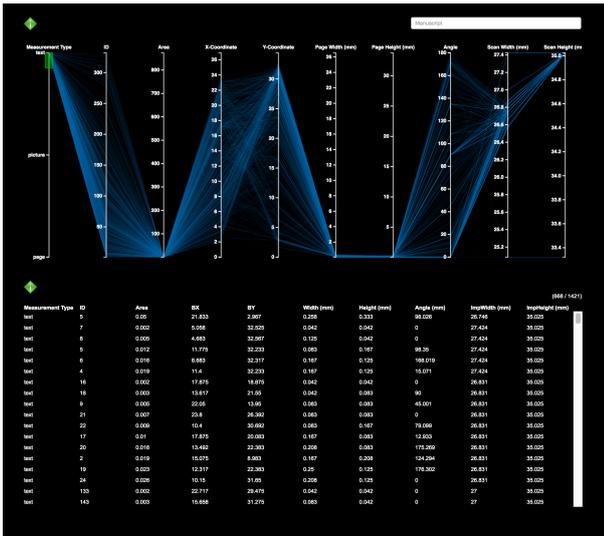


Abbildung 7: Visualisierung der 69 Beispielseiten, hier nur Textbereiche

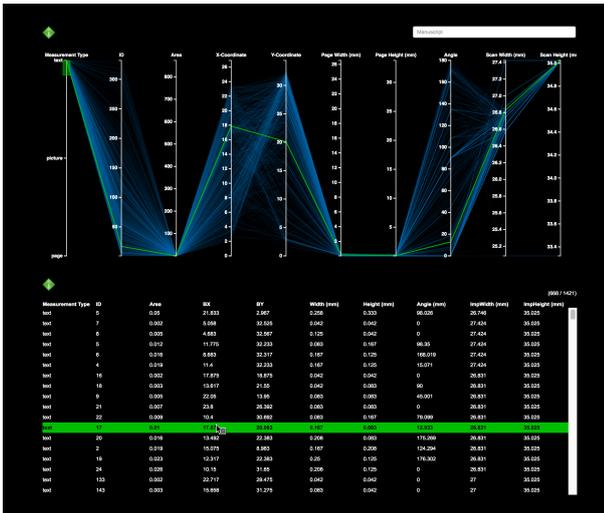


Abbildung 8: Visualisierung der 69 Beispielseiten, nur Textbereiche, Auswahl eines bestimmten Textbereichs (grüne Markierung)

Auf einen Blick sehen wir die Verteilung des Textes, wahlweise im Vergleich mit derjenigen der Bilder auf den Zeitschriftenseiten. Durch die interaktive Funktionalität des Visualisierungstools (CodiVis) können bestimmte Bereiche, einzelne Seiten oder ganze Text- oder Bildblöcke, die bestimmte Gemeinsamkeiten aufweisen, ausgewählt werden.

Wir können auf diese Weise schon beantworten, wie viel Text- und Bildbereiche es im Vergleich auf den gesamten 69 Seiten gibt. Wenn in Zahlen 668 Textbereiche von insgesamt 1421 Bereichen angezeigt werden, ist dabei zwar schon eine Selektion der größeren erfolgt, aber wir können noch nicht sagen, ob 668 einzelne Texte, Spalten, Textteile von Werbeanzeigen oder Überschriften gemeint

sind. Umgekehrt sind es 753 Bildbereiche, aber das sagt uns noch nicht, dass es in Summe ebenso viele einzelne Illustrationen oder Werbebilder sind.

Nichtsdestotrotz ist das Ergebnis für einen ersten Versuch mit den Zeitschriftendokumenten erstaunlich. Die automatische Erkennung von Bild und Text funktionierte und es wurden zu jeder einzelnen Seite genaue Metadaten erhoben, die unabhängig von der Annotation einzelner Forscher oder Nutzer sind und insbesondere reproduzierbar erzeugt werden können. Ein solches Verfahren bietet die Möglichkeit, große Korpora automatisiert zu erschließen und so spezifische Fragestellungen an das Material zu ermöglichen.

Fügt man diesen Metadaten die vorhandenen Metadaten (primär Erscheinungszeitraum, Titel, Ort) jeder Zeitschrift hinzu und lässt den beschriebenen Workflow über das gesamte Korpus des Revistas culturales 2.0 Projekts laufen, kann man bereits historische Vergleiche die Gestaltung der Kulturzeitschriften entsprechend anstellen. Konkret erhoffen wir uns, aktuelle Annahmen, wie die Korrelation von bildlicher vs. textueller Gestaltung mit programmatischer Ausrichtung einer Zeitschrift, beantworten zu können. Lassen sich davon Gesetzmäßigkeiten ablesen, z.B. dass avantgardistische Zeitschriften generell mehr nicht-bedruckten Blattanteil haben als andere? Gibt es da Unterschiede, die die Herkunft und das kulturelle Umfeld der Zeitschriften betreffen? Lassen sich bestimmte Muster in der Gestaltung von Zeitschriften erkennen, die regional, personal oder programmatisch zugeordnet werden können? Anhand dieser Fragen können die Parameter, die für die Gestaltungsentscheidung historisch wichtig waren, definiert werden.

Ausblick

Eine automatische Bild-Text-Erkennung und folgende Analyse des gesamten Zeitschriftenkorpus des Revistas culturales 2.0 Projekts ist nach diesem Testlauf der nächste Schritt. Durch Kombination der bereits vorhandenen mit den automatisch erzeugten Metadaten kann die Visualisierungskomponente erweitert und vervollständigt werden. Danach soll geklärt werden, welche Informationen zusätzlich zum reinen Bild-Text-Verhältnis für jede Zeitschrift, Ausgabe, Seite aus den Messdaten gewonnen werden können. Die Position von Bild und Text auf jeder Seite ist neben der Quantität eine ebenso wichtige gestalterische Information. Durch Kontextwissen zu jedem Zeitschriftentitel und den Metadaten können daraufhin Gestaltung und intellektuelles Konzept (also auch Layout und Inhalt) miteinander verglichen werden.

Sowohl aus informationstechnischer als auch aus geisteswissenschaftlicher Sicht ist die Ausweitung der automatisch erkannten Merkmale eine wichtige Herausforderung. Es stellt sich die Frage, welche zusätzlichen Merkmale für eine vollständige Analyse des Bestandes notwendig sind und durch welche angepasst

oder zu entwickelnden Algorithmen diese bestimmbar sind. Auch zurzeit fehlerhaft erkannte Elemente könnten durch weiterentwickelte Funktionalitäten verbessert zugeordnet werden.

Langfristig bietet sich eine Analyse aller verfügbaren digitalisierten Zeitschriftentitel (deren Bildqualität ausreicht) an. Somit könnten nicht nur transatlantisch spanischsprachige Kulturzeitschriften verglichen werden, sondern auch die internationale Szene modernistischer Zeitschriften (vgl. Blue Mountain Project) oder deutschsprachige illustrierte Magazine (vgl. Projekt: "illustrierte Magazine").

Bibliographie

AsymEnc: <http://asymenc.wp.hum.uu.nl> [letzter Zugriff 20. November 2016].

Benjamin, Walter (1935-39): „Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit“ in: Tiedemann, Rolf (ed.): *Walter Benjamin. Gesammelte Schriften* 1 (2). Frankfurt am Main: Suhrkamp, 1980, 471–508.

DIVAServices-Spotlight: <http://wuersch.pillo-srv.ch/#/> [letzter Zugriff 22. November 2016].

Chandna, Swati / Tonne, Danah / Stotzka, Rainer / Busch, Hannah / Vanscheidt, Philipp / Krause, Celia (2016): „An Effective Visualization Technique for Determining Co-Relations in High-Dimensional Medieval Manuscripts Data“, in: *Visualization and Data Analysis 2016*, San Francisco, California, USA, 14.–18. Februar, 1–6 <http://ist.publisher.ingentaconnect.com/contentone/ist/ei/2016/00002016/00000001/art00013>

Chandna Swati / Tonne, Danah / Jejkal, Thomas / Stotzka, Rainer / Krause, Celia / Vanscheidt, Philipp / Busch, Hannah / Prabhune, Ajinkya (2015): „Software workflow for the automatic tagging of medieval manuscript images (SWATI)“, in: Ringger, Eric K. / Lamiroy, Bart (eds.): *Proceedings SPIE9492, Document Recognition and Retrieval XXII*, 940201 (8. Februar 2015), San Francisco.

Chinese Women's Magazines: <http://kjc-sv013.kjc.uni-heidelberg.de/frauenzeitschriften/index.php> [letzter Zugriff 20. November 2016].

Chinesische Unterhaltungspresse: <http://projects.zo.uni-heidelberg.de/xiaobao/index.php?p=start> [letzter Zugriff 20. November 2016].

Die Fackel, Austrian Academy Corpus: <http://corpus1.aac.ac.at/fackel> [letzter Zugriff 20. November 2016].

eCodicology Projekt: <http://www.ecodicology.org> [letzter Zugriff 20. November 2016].

Ehrlicher, Hanno / Herzog, Teresa (2016): „Zeitschriften Als Netzwerke Und Ihre Digitale Visualisierung: Grundlegende Methodologische Überlegungen Und Erste Anwendungsbeispiele“, in: *Revistas Culturales 2.0*. <http://www.revistas-culturales.de/de/buchseite/hanno->

[ehrlischer-teresa-herzog-zeitschriften-als-netzwerke-und-ihre-digitale](http://www.revistas-culturales.de/de/buchseite/hanno-ehrlischer-teresa-herzog-zeitschriften-als-netzwerke-und-ihre-digitale) [letzter Zugriff 20. November 2016].

ESPrIt: <http://www.espr-it.eu> [letzter Zugriff 20. November 2016].

Europeana Newspapers: <http://www.europeana-newspapers.eu> [letzter Zugriff 20. November 2016].

illustrierte magazine: <http://magazine.illustrierte-presse.de> [letzter Zugriff 20. November 2016].

Frank, Gustav / Podewski, Madleen / Scherer, Stefan (2010): „Kultur – Zeit – Schrift. Literatur- und Kulturzeitschriften als ‚kleine Archive‘“, in: *Internationales Archiv für Sozialgeschichte der deutschen Literatur (IASL)* 34 (2): 1–45.

Macciuci, Raquel (2015): „Técnicas, soporte, ámbitos de sociabilidad y mecanismos de legitimación: sobre la construcción de espacios de literatura en la prensa periódica“, in: Schlünder, Susanne / Macciuci, Raquel (eds.): *Literatura y técnica: derivas ficcionales y materiales: Libros, escritores, textos, frente a la máquina y la ciencia*. Actas del VIII Congreso Orbis Tertius. La Plata: Ediciones del lado de acá 205–231.

Revistas culturales 2.0: Virtuelle Forschungsumgebung zur Erforschung spanischsprachiger Kulturzeitschriften der Moderne (2014–2016): Universität Augsburg. 2014–2016. <http://www.revistas-culturales.de> [letzter Zugriff 20. November 2016].

Rißler-Pipka, Nanette (2014): „Sobre Los Problemas de Investigación Con Revistas Culturales Digitalizadas Del Mundo Hispanohablante“, in: Rißler-Pipka, Nanette / Ehrlicher, Hanno (eds.): *Almacenes de un tiempo en fuga: Revistas culturales en la modernidad hispánica*. Aachen: Shaker 59–80.

Virtuelles Skriptorium St. Matthias: <http://stmatthias.uni-trier.de> [letzter Zugriff 20. November 2016].

WeChangEd: <http://www.wechanged.ugent.be> [letzter Zugriff 20. November 2016].

Yang, Tze-I / Torget, Andrew / Mihalcea, Rada (2011): „Topic Modeling on Historical Newspapers“, in: *Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities LaTeCH*, 24. Juni 2011 Portland, Oregon, USA, Proceedings of the Workshop, Association for Computational Linguistics, Stroudsburg, USA, 96–104. <https://www.aclweb.org/anthology/W/W11/W11-15.pdf> [letzter Zugriff 20. November 2016].

ZEFYS: <http://zefys.staatsbibliothek-berlin.de> [letzter Zugriff 20. November 2016].