

NEUE SCHRIFTEN ZUR ONLINE-FORSCHUNG

Martin Welker / Carsten Wunsch (Hrsg.)

Die Online-Inhaltsanalyse

Forschungsobjekt Internet

Herbert von Halem Verlag

Bibliografische Information der Deutschen Bibliothek

Die deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

Martin Welker / Carsten Wunsch (Hrsg.):

Die Online-Inhaltsanalyse.

Forschungsobjekt Internet

Neue Schriften zur Online-Forschung, 8

Köln: Halem, 2010

Alle Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung, vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung des Verlages reproduziert oder unter Verwendung elektronischer Systeme (inkl. Online-Netzwerken) gespeichert, verarbeitet, vervielfältigt oder verbreitet werden.

© 2010 by Herbert von Halem Verlag, Köln

ISBN 978-3-938258-53-8

<http://www.halem-verlag.de>

E-Mail: info@halem-verlag.de

SATZ: Herbert von Halem Verlag

DRUCK: FINDR, s.r.o., Tschechische Republik

GESTALTUNG: Claudia Ott, Grafischer Entwurf

Copyright Lexicon ©1992 by The Enschedé Font Foundry.

Lexicon® is a Registered Trademark of The Enschedé Font Foundry.

Neue Schriften zur Online-Forschung

hrsg. von der Deutschen Gesellschaft für Online-Forschung e.V. (DGOeF)

Geschäftsführender Reihenerausgeber

Prof. Dr. Martin Welker, MHMK München

Beratendes Herausbergremium der Gesamtreihe

Dr. Wolfgang Bandilla, GESIS – Leibniz-Institut für Sozialwissenschaften

Prof. Dr. Bernad Batinic, Universität Linz

Prof. Dr. Gary Bente, Universität zu Köln

Prof. Dr. Nicola Döring, Technische Universität Ilmenau

Holger Geißler, psychonomics AG

Johannes Hercher, Rogator Software AG

Olaf Hofmann, Skopos Institut für Markt- und

Kommunikationsforschung GmbH

Dr. Lars Kaczmirek, GESIS – Leibniz-Institut für Sozialwissenschaften

Dr. Uwe Matzat, Eindhoven University of Technology

Dr. Wolfgang Neubarth, TNS Infratest

Marc Smaluhn, Research NOW!

Ronald Zwartkruis, NetQuestionnaires AG

Inhalt

I. EINLEITUNG

MARTIN WELKER / CARSTEN WÜNSCH / SASKIA BÖCKING / ANNEKATRIN BOCK / ANNE FRIEDEMANN / MARTIN HERBERS / HOLGER ISERMANN / THOMAS KNIEPER / STEFAN MEIER / CHRISTIAN PENTZOLD / EVA JOHANNA SCHWEITZER Die Online-Inhaltsanalyse: methodische Herausforderung, aber ohne Alternative	9
---	---

II. THEORETISCHE BASIS

1. Forschungsfragen und Gegenstände der Online-Inhaltsanalyse

PATRICK RÖSSLER Das Medium ist nicht die Botschaft	31
---	----

EVA JOHANNA SCHWEITZER Politische Websites als Gegenstand der Online-Inhaltsanalyse	44
---	----

2. Grundgesamtheit und Stichprobenziehung

STEFAN MEIER / CARSTEN WÜNSCH / CHRISTIAN PENTZOLD / MARTIN WELKER	103
---	-----

Auswahlverfahren für Online-Inhalte

STEFAN MEIER / CHRISTIAN PENTZOLD Theoretical Sampling als Auswahlstrategie für Online-Inhaltsanalysen	124
--	-----

SEBASTIAN ERLHOFER 144
Datenerhebung in der Blogosphäre:
Herausforderungen und Lösungswege

3. Analyseeinheit und Kategorienbildung

MONIKA TADDICKEN / KERSTIN BUND 167
Ich kommentiere, also bin ich.
Community Research am Beispiel des
Diskussionsforums der *ZEIT Online*

LARS KACZMIREK / CHRISTIAN BAIER / CORNELIA ZÜLL 191
Wie empfinden Teilnehmer die Fragen
in Online-Befragungen? Entwicklung eines Diktionärs für die
automatische Codierung freier Antworten

ANNEKATRIN BOCK / HOLGER ISERMANN /
THOMAS KNIEPER 224
Herausforderungen bei der
quantitativen (visuellen) Inhaltsanalyse von Online-Inhalten

4. Reliabilität und Validität bei Online-Inhaltsanalysen

MARTIN R. HERBERS / ANNE FRIEDEMANN 240
Spezielle Fragen der Reliabilität und
Validität bei Online-Inhaltsanalysen

5. Methodenkombination

WOLFGANG SCHWEIGER / PATRICK WEBER 267
Strategische Kommunikation auf Unternehmens-Websites.
Zur Evaluation der Kommunikationsleistung durch eine
Methodenkombination von Online-Inhaltsanalyse und
Logfile-Analyse

SASKIA BÖCKING / FRANÇOIS RÜF / MARKUS JUFER /
CORINNE HÜGLI-BALTZER 291

Analyse des politischen Informationsverhaltens mittels
der von Wählern genutzten Internetinhalte –
Potenziale für die Prognose von Wahlergebnissen

6. Statistisch-automatisierte Inhaltsanalyse

FRANÇOIS RÜF / SASKIA BÖCKING / STEFAN KUMMER 313

Automatisierte Inhaltsanalysen im Internet:
Möglichkeiten und Grenzen am Beispiel des
SINDBAD-Knowledge-Generators

MICHAEL SCHARKOW 340

Lesen und lesen lassen –
Zum State of the Art automatischer Textanalyse

III. PRAXIS

7. Fallbeispiele und Anwendungsfelder

KATHERINE GÜRTLER / ELKE KRONEWALD 365

The Automated Analysis of Media: PRIME web.Analysis:
A Case Study

DORETTE WESEMANN / MARTIN GRUNWALD 387

Inhaltsanalyse von Online-Diskussionsforen
für Betroffene von Essstörungen

NIK STUCKY / FRANÇOIS RÜF / NILS KELLER /
STEFAN KUMMER 409

Image & Co: Eine Online-Analyse der
Markenwahrnehmung am Beispiel von Automarken

MELANIE ARENS / SASKIA BÖCKING / STEFAN KUMMER / FRANÇOIS RÜF	427
Das Meinungsklima zur Klimakonferenz: Wie sich Themenkarrieren im Internet entwickeln und mit welchen Themen sich Politiker positionieren	
KONRAD RÜDIGER / MARTIN WELKER	448
Redaktionsblogs deutscher Zeitungen. Über die Schwierigkeiten diese inhaltsanalytisch zu untersuchen – ein Werkstattbericht	
IV. ANHANG	
Verzeichnis der Autoren und Herausgeber	469

I. EINLEITUNG

MARTIN WELKER / CARSTEN WÜNSCH / SASKIA BÖCKING
ANNEKATRIN BOCK / ANNE FRIEDEMANN / MARTIN HERBERS /
HOLGER ISERMANN / THOMAS KNIEPER / STEFAN MEIER /
CHRISTIAN PENTZOLD / EVA JOHANNA SCHWEITZER

Die Online-Inhaltsanalyse: methodische Herausforderung, aber ohne Alternative

Einleitung

Die Inhaltsanalyse zählt zu den wichtigen Methoden in den Sozialwissenschaften und hat sich insbesondere in der Medien- und Kommunikationswissenschaft als eine zentrale Methode herauskristallisiert (KAMHAWI/WEAVER 2003; LAUF 2006; RIFFE/FREITAG 1997).¹ Bei Erhebungen zur Häufigkeit des Einsatzes empirischer Methoden in der wissenschaftlichen Forschung nimmt die Inhaltsanalyse zwar keinen Spitzenrang, aber einen prominenten Platz ein (vgl. GOULD 2004; WELKER 2007: 47). Die Flexibilität des inhaltsanalytischen Verfahrens erlaubt es, die Methode auf unterschiedliche Medienangebote anzuwenden, sofern diese semiotische Strukturen aufweisen, die inhaltsanalytisch erfasst werden können (vgl. MERTEN 1995: 60). Als empirisches Verfahren zur »systematischen, intersubjektiv nachvollziehbaren Beschreibung inhaltlicher und formaler Merkmale von Mitteilungen« (FRÜH 2007: 27) erlaubt die Methode die kategoriale Zerlegung verschiedener Botschaften in quantifizierbare Einheiten. Text, Bild und Ton können vergleichsweise kostengünstig, komplexitätsreduzierend sowie nonreaktiv erfasst und mit wechselnden Fragestellungen beliebig oft zu unterschiedlichen Zeiten untersucht werden (vgl. FRÜH 2007: 39; KRIPPENDORFF 2004: 40ff.; RIFFE/LACY/FICO 1998: 30ff.; RÖSSLER 2005: 16).

1 Auch viele andere Wissenschaftsdisziplinen nutzen diese Methode (vgl. FRÜH 2007: 14).

Die Online-Inhaltsanalyse grenzt sich von der ›klassischen‹ Inhaltsanalyse durch ihren Bezug auf ihre Forschungsgegenstände ab. Diese sind online, überwiegend im Internet, zu finden. Durch diese Abgrenzung entsteht zwar keine eigenständige Methode der Online-Inhaltsanalyse. Aber aufgrund der spezifischen Merkmale des empirischen Relativs – der ›Online-Inhalte‹ – ergeben sich neue, vielfältige Möglichkeiten aber auch spezifische Probleme. Hinter dem Begriff ›Online-Inhalte‹ – als Gegenstand und Produkt von Online-Kommunikation – verbergen sich zahlreiche und z.T. stark heterogene Angebote, aber auch Strukturen netzwerkbasierter Kommunikation. Die Gemeinsamkeiten dieser Inhalte und Kommunikationsformen finden sich in der zugrunde liegenden technischen Infrastruktur: der computervermittelten Vernetzung verschiedenster Kommunikationsdienste (vgl. z.B. BECK 2006; RÖSSLER 2003). Versuche, Online-Kommunikation oder das Internet zu definieren, können auf die Differenzierung zwischen Medien erster und zweiter Ordnung zurückgreifen (vgl. JOERGES/BRAUN 1994; KUBICEK 1997). Mit Medium erster Ordnung ist ein technisches Vermittlungssystem gemeint. Diese Vermittlungssysteme lassen – insbesondere bei der Online-Kommunikation – verschiedene Realisierungen von Kommunikation zu. Diese Kommunikationsmodi stellen dann die Medien zweiter Ordnung dar (RÖSSLER 1998). Gegenstand der Online-Inhaltsanalyse sind die Kommunikate (Strukturen und Inhalte) dieser Online-Medien zweiter Ordnung. Dazu gehören u.a. Inhalte und Verlinkungen von Weblogs, Kommunikation in Foren, E-Mails, Inhalte von www-Angeboten oder auch der Austausch in sozialen Netzwerken. Bei dieser Liste handelt es sich keineswegs um den Versuch einer vollständigen Auflistung. Aufgrund der Offenheit unserer Definition und der stetigen Weiter- und Neuentwicklungen netzbasierter Kommunikationsmodi kommen kontinuierlich neue Untersuchungsgegenstände für die Online-Inhaltsanalyse hinzu und andere verlieren an Bedeutung.

Da sich aber all diese Modi und Kommunikate derselben technischen Infrastruktur – des Internets – bedienen, unterliegen sie denselben Beschränkungen und weisen die gleichen Potenziale auf (vgl. etwa JANKOWSKI/VAN SELM 2000; NEUBERGER 2000; DEUZE/DIMOUDI 2002; BARDOEL 2002; PÜRER 2004). Die meisten dieser Spezifika sind für die Analyse der Online-Inhalte relevant (vgl. RÖSSLER 1997, 1998, 2001; RÖSSLER/WIRTH 2001; SEIBOLD 2002; WOLLING/KUHLMANN 2003). Wir wollen die wichtigsten dieser Merkmale daher an dieser Stelle knapp vorstellen und in einem

daran anschließenden Überblick die daraus folgenden Konsequenzen und insbesondere Schwierigkeiten für die Online-Inhaltsanalyse diskutieren.

Flüchtigkeit, Dynamik und Transitorik der Inhalte: Online-Inhalte werden kontinuierlich neu erstellt, verändert oder gelöscht. Zu einem bestimmten Zeitpunkt aufzufindende Inhalte stellen daher häufig eine flüchtige Momentaufnahme dar. Dabei wird der jeweils aktuelle Stand zu einem Zeitpunkt i.d.R. nicht archiviert, sodass er nach Modifikationen oder Löschung auch nicht mehr zur Verfügung steht und sich diese Dynamik nachträglich auch nicht rekonstruieren lässt.

Medialität, Multimedialität bzw. Multimodalität: Die technische Infrastruktur des Internets und deren zugrunde liegende Digitalisierung der Kommunikate lässt eine große Vielfalt in der Verwendung verschiedener medialer Zeichenträger (digitale Bild-, Video-, Audio-Formate für unterschiedliche Ausgabegeräte: Bildschirm, Home-Entertainment-Station, mobile Endgeräte etc.) und Zeichenmodalitäten bzw. -systeme (Film, Bild, Sprache, Design/Layout) zu. Dies führt zu einer potenziell hohen Heterogenität der Inhalte.

Nonlinearität/Hypertextualität: Viele, wahrscheinlich die meisten Online-Inhalte liegen in einer nicht linearen Struktur vor. Dies äußert sich durch Verlinkungen, die von einer Webseite auf andere Webseiten verweisen und so kommunikative Einheiten ergeben, welche zu geplanten oder emergenten Netzwerken mit unterschiedlicher Linktiefe führen können. Dadurch verwischen auch die Grenzen zwischen verschiedenen Inhalten.

Reaktivität und Personalisierung: Aufgrund der aktuellen Generierung von Inhalten, die (auch) interaktiv auf Nutzereingaben basieren und so quasi personalisierte Angebote entstehen lassen, sind Online-Inhalte als reaktiv zu charakterisieren. Dies erfolgt bspw. durch datenbankgenerierte Contentangebote gemäß individuell realisierter Suchabfragen, individueller Account-Angebote in Communitys und Portalen, individueller Navigation durch die Online-Angebote etc. sowie durch mögliche Interaktivität, die eine kollektive bzw. usergestützte Inhaltsproduktion in Foren, Blogs etc. bewirkt. Durch diesen Aspekt verstärkt sich auch das von uns bereits genannte Merkmal der Flüchtigkeit von Online-Inhalten: Unter Umständen werden personalisierte Angebote nur einmalig angeboten.

Digitalisierung/Maschinenlesbarkeit: Online-Inhalte liegen in einer digitalen oder digitalisierten Form vor und stehen dadurch zwangsläufig – in unterschiedlicher Quantität und Qualität – einer elektronischen Verarbei-

tung zur Verfügung, ohne dass die Inhalte dafür aufwendig aufbereitet werden müssten.

Quantität: Das Internet speichert bereits heute eine gigantische Menge an Inhalten und stellt diese zum Abruf bereit. Aufgrund der weiteren Entwicklung der Speichermedien, Datenübertragungskapazitäten und neuer Inhalte nimmt diese Menge weiterhin rasant zu. Dies geht mit dem Problem einher, in dieser Masse an Inhalten spezifische Angebote zu finden.

Die Analyse von Online-Inhalten folgt methodisch prinzipiell und mit mehr oder minder großen Schwierigkeiten dem Ablauf der klassischen Inhaltsanalyse (siehe FRÜH 2007). Aufgrund der Maschinenlesbarkeit der Online-Inhalte kommen hier aber besonders häufig auch neue, computerunterstützte Verfahren (wie Text-Mining) zum Einsatz. Diese basieren auf textsemantischen Prozeduren (vgl. FRÜH 2007: 52), können aber kontextabhängig in die (Online-)Inhaltsanalyse integriert werden. So können mithilfe des Computers vor allem Codiervorgänge von Texten automatisiert werden. Auch diesen spezifischen Instrumenten werden wir im vorliegenden Buch Raum geben (vgl. u.a. den Beitrag von KACZMIREK/BAIER/ZÜLL). Zu betonen ist aber, dass diese Instrumente nicht an den Online-Bereich gebunden sind, sondern lediglich mit geringerem Aufwand zum Einsatz kommen können.

Die einfache Adaption des methodischen Inventars der klassischen Inhaltsanalyse setzt allerdings voraus, dass der inhaltsanalytisch zu betrachtende Forschungsgegenstand eine konstante physische Repräsentation aufweist, sich klar und eindeutig von anderen Kommunikationsangeboten abgrenzen lässt und ferner einem einheitlichen sprachlichen oder visuellen Präsentationsmuster folgt, das der späteren Auswertung als Leitfunktion zugrunde gelegt werden kann. Bei traditionellen Formaten der Massenkommunikation wie Zeitungen, Zeitschriften, Fernsehen oder Radio lässt sich dies bisher als gegeben annehmen. Das Gutenberg-Zeitalter bot gute Voraussetzungen für die Anwendung der Inhaltsanalyse, weil das Untersuchungsmaterial einen recht eindeutigen ›Aggregatzustand‹ aufwies. Die Schwierigkeiten lagen allenfalls in der Beschaffung des physischen Materials, der Forscher musste zur Not in staubige Archive hinabsteigen oder längere Zeit auf Material warten, das er über Fernleihe bestellt hatte.

Das Internet als Medium der Online-Kommunikation hat diese methodische Eindeutigkeit leider überwiegend vertrieben. Zwar waren noch nie so viele Inhalte verschiedenster Kommunikationsformen so

schnell verfügbar und zugänglich wie heute. Insbesondere das World Wide Web bietet mit seiner immensen Größe – 2009 waren mehr als 625 Millionen Hosts online (Internet Systems Consortium, Inc. (isc) 2009) – scheinbar die besten Möglichkeiten für sozialwissenschaftliche Untersuchungen. Leider wachsen mit dieser Verfügbarkeit aber auch die Unübersichtlichkeit und die Uneindeutigkeiten. In der Anfangszeit des Internets waren Inhalte von Websites noch überwiegend textbasiert und konnten mit den etablierten Methoden der Textanalyse untersucht werden. Erst die Ausbreitung von Breitbandinternetanschlüssen ermöglicht seit einiger Zeit die flächendeckende Integration von Audiodateien, Stillbildern, Bewegtbildern und AV-Dateien (vgl. TRAPPEL 2007: 38). Diese Multimedialität des World Wide Webs stellt die Inhaltsanalyse vor besondere Herausforderungen: zum einen, weil der komplexe Prozess der Rezeption und Wirkung von visuellen Inhalten die Analyse erschwert, und zum anderen, wenn verschiedenste Effekte des Zusammenwirkens der unterschiedlichen Medien und die weiteren Charakteristika des www in die Inhaltsanalyse einfließen müssen, zum Beispiel Text-Bild- und Bild-Text-Ausstrahlungs- oder Framing-Effekte (vgl. etwa JANKOWSKI/VAN SELM 2000; NEUBERGER 2000; PÜRER 2004). Diese Formen der Online-Kommunikation bedeuten eine methodische Herausforderung für die Wissenschaftler (vgl. BECK 2006: 11; NEVERLA 2001; WERNER 1998), da die vorgestellten charakteristischen Merkmale des Internets einer linearen Vermessung in herkömmlicher Weise entgegenstehen. »[D]er ureigene Charakter des Mediums [verletzt somit die] Grundaxiome der Inhaltsanalyse« (RÖSSLER/EICHHORN 1999: 267) und macht folglich eine gesonderte Methodenreflexion für den Online-Bereich notwendig.

Diesen Problemen ist das vorliegende Buch vorrangig gewidmet. Denn leider kann weder das Gutenberg-Zeitalter zurückgeholt werden, noch können und sollten Wissenschaftler die Erforschung von Online-Inhalten ausblenden. Unter anderen Vorzeichen hatte diese Forderung Anfang des 20. Jahrhunderts bereits Max Weber in seiner berühmten Rede zum ersten Deutschen Soziologentag in Frankfurt am Main erhoben. Die Analyse des Zeitungswesens, so Weber, sei eine vorrangige Aufgabe der Wissenschaft. Dies sei darin begründet, dass die Massenpresse nahezu jeden Bereich der Gesellschaft beeinflusse und eine Wissenschaft von der Gesellschaft sich also primär mit diesem Gegenstand befassen müsse. Von Online-Medien ließe sich heute, rund 100 Jahre später, das Gleiche sagen. Warum sollte bei der Erforschung und Erfassung bestimmter Wirkungs-