

# PUBLIK

## Vierteljahreshefte für Kommunikationsforschung

Inhalt: Community Lost: der Niedergang traditioneller sozialer Beziehungen - Community Liberated: die Bedeutung moderner sozialer Beziehungen - Soziales Kapital - Soziale Netzwerkanalyse - Ein Constrained-Choice-Modell persönlicher Netzwerke - Hypothesen zu Unterschieden zwischen persönlichen Netzwerken - Soziale Unterstützungsanalyse - Ein Austauschmodell informeller sozialer Unterstützung - Hypothesen zu Unterschieden innerhalb persönlicher Netzwerke - Forschungsdesign - Datenbasis - Hypothesenprüfung der drei Netzwerkmodelle - Hypothesenprüfung der vier Unterstützungsmodelle

Herausgegeben von Christina Holtz-Bacha · Arnulf Kutsch  
Wolfgang R. Langenbucher · Ulrich Saxer

Sören Petermann  
**Persönliche Netzwerke in Stadt und Land**  
Siedlungsstruktur und soziale Unterstützungsnetzwerke  
im Raum Halle/Saale  
2002, 232 S., mit 21 Abb. u. 23 Tab. Br. EUR 22,00  
ISBN 3-531-13750-6

Inhalt: Community Lost: der Niedergang traditioneller sozialer

Beziehungen - Community Liberated: die Bedeutung moderner sozialer Beziehungen - Soziales Kapital - Soziale Netzwerkanalyse - Ein Constrained-Choice-Modell persönlicher Netzwerke - Hypothesen zu Unterschieden zwischen persönlichen Netzwerken - Soziale Unterstützungsanalyse - Ein Austauschmodell informeller sozialer Unterstützung - Hypothesen zu Unterschieden innerhalb persönlicher Netzwerke - Forschungsdesign - Datenbasis - Hypothesenprüfung der drei Netzwerkmodelle - Hypothesenprüfung der vier Unterstützungsmodelle

Andreas Bergkapp

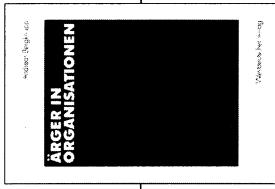
### Ärger in Organisationen

Eine systemische Strukturanalyse

2002, 442 S. mit 51 Abb. u. 13 Tab. Br. EUR 42,00

ISBN 3-531-13744-1

Für ein umfassendes Verständnis von Ärger im Handlungsfeld "Organisation" greift die immer noch dominierende individualpsychologische Sichtweise zu kurz. Durch die Integration von psychologischen, soziologischen und organisationstheoretischen Ansätzen wird in dieser Arbeit ein theoretischer Ausgangspunkt entwickelt, der neue Hinsichten auf Ärger-Episoden in Organisationen ermöglicht. Es wird deutlich, dass Ärger nicht länger als eine auf Individuen beschränkte Randerscheinung oder irrationale Störgröße in einem ansonsten rational funktionierenden Kooperationszusammenhang gesehen werden kann.



**Bestellung per Fax: 0611.7878-400**

**Ja,** ich interessiere mich und bestelle:

Expl. Petermann  
**Persönliche Netzwerke in Stadt und Land**  
EUR 22,00 ISBN 3-531-13750-6

Expl. Bergkapp  
**Ärger in Organisationen**  
EUR 42,00 ISBN 3-531-13744-1

**Meine Anschrift:**  
Vorname / Name \_\_\_\_\_  
Firma \_\_\_\_\_  
Abteilung \_\_\_\_\_  
Straße \_\_\_\_\_  
PLZ, Ort \_\_\_\_\_  
Datum \_\_\_\_\_  
Unterschrift \_\_\_\_\_

## Die flüchtigen Web-Informationen einfangen

Lösungsansätze für die Online-Inhaltsanalyse bei dynamischen Inhalten im Internet

### EINLEITUNG

Es ist nur ein kleiner Klick für den Online-Redakteur, aber ein großes Problem für den Forscher: Gerade noch prangt auf der Startseite des Online-Nachrichtenangebots ein Bild von Arafat mit einer Schlagzeile zum Nahostkonflikt. Wenige Sekunden später zeigt der gleiche Aufmacher nur noch fallende Börsencharts und mahnende Bankanalysten. Arafat ist nach unten auf das Ende der Webseite gerutscht oder ganz vom Bildschirm verschwunden. Große Teile des Internets leben vom ständigen Wechsel der angebotenen Inhalte. Besonders das World Wide Web wird durch einen neuen Publikationsstil geprägt, bei dem die Inhalte dynamisiert, multimedialisiert, engrenzt und partikularisiert werden. Dies stellt die Inhaltsanalyse vor neue Aufgaben. Über Jahrzehnte hinweg hat sie sich als zentrale Methode der Kommunikationswissenschaft erwiesen. Das Fach hat vor allem für die Printmedien und für das Fernsehen »die Inhaltsanalysen kultiviert und auch weiterentwickelt« (Brosius/Koschel 2001: 156). Innerhalb von zehn Jahren ist nun das Internet in der industrialisierten Welt »selbstverständlicher Bestandteil des Medienangebots« geworden (van Eimeren/Gerhard/Frees 2001: 382). Heute nutzen etwa 40 Prozent der bundesrepublikanischen Bevölkerung ab 14 Jahren (vgl. van Eimeren/Gerhard/Frees 2001: 382f.) das »Netz der Netze«, weltweit sind um die 350 Millionen Nutzer auf rund vier Milliarden Webseiten unterwegs (vgl. Focus Magazin Verlag 2001: 1f.). Dieses rasante Wachstum des Publikums und der Inhalte hat das Internet zu einem der zentralen Forschungsgebiete der Kommunikationswissenschaft gemacht. Mit einem Erfolg wenden Kommunikationswissenschaftler das Instrument der Inhaltsanalyse auf die dynamischen Inhalte des World Wide Webs an (vgl. Bucher/Barth 1998; Evans 1998; Klein 1999; McMillan 2000; Rössler 1997; Rössler/Wirth 2001; Seibold 2002; Wirth/Brech 1998). Dazu können sie zwar auf das bewährte Instrumentarium der sozialwissenschaftlichen Inhaltsanalyse zurückgreifen (vgl. Früh 1998; Merten 1995). Dieses sollte allerdings modifiziert und erweitert werden, um den neuen Gegebenheiten des Internets gerecht zu werden. Als Beitrag dazu werden im Folgenden fünf neue theoretische und vier praktische Besonderheiten der Online-Inhaltsanalyse dargestellt. Unter Online-Inhaltsanalyse wird die Untersuchung von Inhalten des Internets verstanden, besonders des World Wide Webs. Nicht gemeint ist die computerunterstützte Inhaltsanalyse

---

Balthas Seibold ist Diplomjournalist. Er untersucht zurzeit für die Gesellschaft für technische Zusammenarbeit (GTZ) und die UNESCO »Online-Content for Development – eine Evaluation der Online-Informationsbedürfnisse in Entwicklungsländern als Beitrag zum Aufbau der geplanten UNESCO-Knowledge Portale« im Rahmen des Stiftungskollegs für internationale Aufgaben der Studienstiftung des deutschen Volkes und der Robert Bosch Stiftung.

(CUI), die mit elektronischer Hilfe »Offline-Inhalte« analysiert. Natürlich ist aber auch die Online-Inhaltsanalyse in hohem Grade computerunterstützt.

Auf den theoretischen und praktischen Besonderheiten der Online-Inhaltsanalyse aufbauend werden schließlich Lösungsansätze aufgezeigt, mit denen man auch flüchtige Informationen des World Wide Webs inhaltsanalytisch einfangen kann. Insbesondere wird argumentiert, dass die Inhaltsanalyse von dynamischen Netz-Inhalten auch die Nutzungsanalyse integriert sollte. Dann bietet die Online-Inhaltsanalyse für die Kommunikationswissenschaft große Chancen: Inhaltsanalyse und Rezeptionsanalyse können so miteinander verknüpft werden, dass die Zusammenhänge von Inhalten, Rezipient und Auswahl sichtbar, messbar und überprüfbar werden.

#### THEORETISCHE HERAUSFORDERUNGEN DER ONLINE-INHALTSANALYSE

Online-Inhalte sind Hypertexte. Die Informationen verteilen sich dabei auf verschiedene Webseiten, die untereinander und oft über ein Navigationssystem miteinander verlinkt sind. Dieser Hypertextcharakter führt zum oft beschriebenen Effekt des »Netzes«, in dem verschiedene Informationen durch Links miteinander verknüpft sind. Der Rezipient klickt sich zum Beispiel von einer journalistischen Webseite mit einer Buchbesprechung zu der Homepage des Autors und von dort zur Bestellsseite eines Online-Buchhändlers. Oft bleibt unklar, wo die Grenzen eines spezifischen Webangebots liegen (vgl. McMillan 2000: 82; Klein 1999; Rössler 1997: 253-255). Schon der erste Schritt der Inhaltsanalyse, das Abgrenzen des Untersuchungsgegenstandes, kann somit zum Hürdenlauf durch Linkebenen werden. »Zu jedem beliebigen Zeitpunkt gibt es eine nahezu unendliche Zahl theoretisch abrufbarer Kommunikationsinhalte« (Rössler 1997: 254f.; vgl. auch Stempel/Stewart 2000: 545). Dazu das Beispiel eines journalistischen Online-Angebotes: Oft enthält die Nachrichtenseite das Ressort Wetter, das von einem fremden Anbieter zugeliefert wird, aber im gleichen Layout gehalten ist. Will der Nutzer aber zum Beispiel das Wetter einer speziellen Stadt wissen, so gelangt er beim nächsten Klick unmerklich auf die Seite des Wetter-Anbieters selbst. Gehört diese nun inhaltsanalytisch zum Angebot der Nachrichtenseite?

Eine sinnvolle Lösung für die Inhaltsanalyse ist die Beschränkung des Untersuchungsgegenstands auf einzelne Teile des Webangebots, zum Beispiel die am häufigsten frequentierte Startseite (vgl. McMillan 2000: 82), die die weitere Rezeption eines Online-Angebots als »advance Organizer« (Bucher 2000: 162) vorstrukturiert. Von dieser ausgehend können zusätzlich Seiten erfasst werden, die durch eine definierte Anzahl von Verweisstritten erreicht werden können (vgl. Rössler 1997: 257). Weitere Anhaltspunkte für die Abgrenzung des Angebots können deutliche Layoutwechsel oder redaktionelle Hinweise sein (zum Beispiel »hier verlassen sie das Angebot von xyz«). Auch die Speicheradresse der jeweiligen Webseiten, die URL,<sup>1</sup> kann ein solcher Wendepunkt sein, der anzeigt, ob man noch »im Untersuchungsobjekt« navigiert. Oft, aber nicht immer, sind die Inhalte eines Anbieters in so genannten Unterordnern

seiner Hauptdomain abgelegt, zum Beispiel der Art <<http://www.hauptdomain.de/unterordner/untersite.htm>>.

Doch wer im Internet schließlich sein Untersuchungsobjekt abgegrenzt hat, kommt vom Regen in die Traufe. Vor ihm steht eine große theoretisch-konzeptionelle Herausforderung des neuen Mediums: Er muss die Analyseeinheiten definieren. Ist die kleinste Einheit der inhaltsanalytischen Untersuchung eine Webseite, eine Kategorie innerhalb der Seite, ein Link, die ersten n-Ebenen einer Link-Hierarchie? Wie werden Navigationselemente integriert? Sind die Kategorien vergleichbar und trennscharf (vgl. Rössler/Wirth 2001: 291-293)? Wie in der traditionellen Inhaltsanalyse sollte die adäquate Analyseeinheit im Wesentlichen vom Forschungsinteresse determiniert werden (vgl. McMillan 2000: 92f.; Friih 1998). Wer also die Nachrichten von journalistischen Online-Angeboten analysieren will, muss einzelne Meldungen als Analyseinheit wählen. Wer Shopping-Angebote von Nachrichtenseiten unterscheiden will, findet auf der Ebene ganzer Webangebote sinnvolle Einheiten (vgl. Rössler 1997: 255). Doch die theoretische Besonderheit des Webs ist grundsätzlicher: »Online-Angebote sind nicht-lineare Medien, das heißt, im Unterschied zu Hörfunk und Fernsehen ist die Rezeptionsabfolge nicht vorgegeben, sondern frei wählbar« (Bucher/Barth 1998: 517). Hierin ähnelt das Web den Printmedien Zeitung und Zeitschrift. Hier wie dort springt der Leser von einem Inhalts-Baustein zum nächsten. Doch während im Printbereich der Artikel als Basisinheit jeder Rezeption (vgl. Evans 1998: 162) feststeht, kann die hierarchische Hypertextzerstückelung des Webs selbst diese Grund-Lineartät aufbrechen. »It makes little sense to conceptualize news stories as the basic content unit«, argumentiert Evans für die Inhaltsanalyse von Nachrichtenseiten im Web (Evans 1998: 162). »When each consumer can create his or her own path through on-line news resources, there may be as many news 'stories' as there are consumers« (Evans 1998: 162). Dazu ein Beispiel: Der Leser eines journalistischen Online-Angebots wählt auf der Startseite eine Meldung über eine Landtagswahl aus. Nachdem er diese gelesen hat, sieht er unter der Meldung einen Link zu einem Hintergrundbericht über dieses Bundesland. Er wählt diesen Link, und kommt schließlich von dem Hintergrundbericht über eine weitere Verknüpfung zu einer erläuternden Landkarte des betreffenden Bundeslandes. Zu beachten ist dabei erstens, dass der Leser diese Trias vermutlich im Nachhinein als eine zusammengehörende Einheit begreifen wird, obwohl seine Auswahl nur eines von vielen alternativen Text-Menüs war. Zweitens hat der Online-Leser auf jeder Station seiner »Online-Wanderung« (Werner 1998: 237) andere Auswahlmöglichkeiten. Auf der Startseite konnte er beispielsweise noch aus 30 Meldungen auswählen, in der nächsttieferen Link-Ebene aus drei weiteren Links, und sein dritter Schritt zur Landkarte war vielleicht der einzige mögliche. Es zeigt sich also, dass die Online-Inhaltsanalyse den Klickpfad der Leser über die Links der Webseiten hinweg beachten sollte, da erst dieser Pfad aus den verschiedenen »Angebotsoptionen« (Rössler/Wirth 2001: 283) den Gesamt-Inhalt der einzelnen Rezeptionen<sup>2</sup> konstituiert.

<sup>1</sup> URL steht für Uniform Resource Locator und bezeichnet die Speicheradresse einer Webseite. Oft wird die URL in einer Leiste des Browsers angezeigt.

<sup>2</sup> Beim Klicken des Rezipienten über verschiedene Seiten hinweg ist natürlich nicht garantiert, dass ein Nutzer jeden Textbaustein (vollständig) liest, den er aufruft. Daher meint »Rezeption« in diesem Zusammenhang zunächst nur den Kontakt des Rezipienten mit den einzelnen Seiten. Wie und wie gründlich ein Nutzer eine Seite liest, hängt von vielen Faktoren wie Involvement und Rezeptionsziel

Natürlich ist diese Art von rezeptionsorientierter Online-Inhaltsanalyse ein komplexes Verfahren, denn es reicht nicht mehr, alle »vorliegenden« Inhalte und deren Verknüpfungen zu untersuchen. Genauso wichtig wird es, den »user-defined path« (Evans 1998: 164) durch das Inhaltsnetz nachzuvollziehen. Die Inhaltsanalyse wird zum »Web Mining« als Kombination aus »Web Content Mining« und »Web Usage Mining« (Bensberg/Weiß 1998: 3f.). Dieser Multimethodenansatz jedoch erweitert die Möglichkeiten der Inhaltsanalyse: Die Verknüpfung mit einer »Selektivitätsanalyse« (Rössler/Wirth 2001: 286) des Nutzerverhaltens führt dazu, dass man treffendere Aussagen über den Zusammenhang von Inhalten und Rezeption machen kann. Insbesondere werden Rezeptionsmuster von Inhaltsbausteinen innerhalb der einzelnen Webseiten deutlich. Dazu ein weiteres Beispiel: Eine Untersuchung der Startseite des Online-Angebots der Süddeutschen Zeitung ([<http://www.sueddeutsche.de>](http://www.sueddeutsche.de)) zeigte, dass ein bestimmter Artikel als oberste Meldung doppelt so viele Leser anzieht wie genau der gleiche Artikel weiter unten auf der Startseite (vgl. Seibold 2002: 115f.). Solche Primacy-Effekte der Linkauswahl sind überhaupt nur im Zusammenspiel von Inhalts- und Rezeptionsanalyse fassbar. Das gleiche gilt für weitere Effekte der Link- und Navigationsstruktur auf die Rezeption wie der Recency- und der Scroll-Effekt. Hier wird eine Bevorzugung der zuletzt rezipierten Textbausteine (Recency-Effekt) oder der zunächst auf dem Bildschirm sichtbaren Inhalte (Scroll-Effekt) vermutet (vgl. Seibold 2002: 56–58; Schweiger 2001: 213–219).

Wo erst die Rezeption das Gesamt-Menü der Inhalte schafft, da ermöglicht die Kontrolle der Inhalte auch die Kontrolle der Rezeption. Ein Extrembeispiel dieser »prinzipielle[n] Reaktivität des World Wide Web« (Rössler/Wirth 2001: 282) stellen personalisierte Online-Dienste dar. Diese präsentieren jedem Nutzer seine persönlichen Nachrichten als »Daily Me« (Riefler 1999: 196). Dabei werden Auswahl, Form und Präsentation der Nachrichten nach den Interessenschwerpunkten strukturiert, die der Rezipient zuvor ausgewählt hat oder durch sein Nutzungsverhalten nahe legt. Bei dieser Form der personalisierten Online-Nachrichten ist eine Inhaltsanalyse ohne Nutzungsanalyse kaum möglich: Ohne zu wissen, ob, in welcher Form und an welcher Position jemand einen bestimmten Wirtschaftsbericht eines personalisierten Nachrichtendienstes bekommt, erscheint es wenig sinnvoll, diesen inhaltsanalytisch zu untersuchen. Und wenn dieser Wirtschaftsbericht immer nach einer bestimmten Kursfrage gelesen wird, so müssen beide eine inhaltsanalytische Einheit bilden. Da die Inhalte des World Wide Web größtenteils als kleine Teil-Einheiten in Datenbanken lagern, ist es für die Anbieter technisch gar kein Problem, diese immer neu zu personalisieren Seiten zusammenzusetzen (zu den technischen Modalitäten vgl. Rössler/Wirth 2001: 283). Die Antwort der Online-Inhaltsanalyse auf diese Daten-Segmentierung sollte die Verknüpfung mit der Nutzungsanalyse sein.

Das World Wide Web liefert dem Forstner dazu die technischen Mittel. Erstmals in der Geschichte der Medien ist es möglich, die Online-Wanderungen jedes einzelnen Mediennutzers nicht-reaktiv und ohne aufwändige Apparaturen zu erleben. Dazu ab (vgl. Schweiger 2001: 84–91). Sehr wahrscheinlich ist beim Navigieren im Web allerdings, dass ein Leser den Text oder das Bild eines Links erfasst, bevor er diesen auswählt (wenn man zufälliges Hervorklicken außer Acht lässt). »Rezeption« und »Lesepfad« sind auch im Folgenden in diesem eingeschränkten Sinne zu verstehen.

muss man das »Gold im Logfile« (Janercko 1999: 87) schürfen. Was ist ein Logfile? Man kann es sich als riesiges digitales Protokoll vorstellen, in das jeder Abruf einer Webseite im Hintergrund notiert wird. Fast jeder Anbieter von Online-Inhalten benutzt solche Logfiles, um die Klickpfade der Nutzer innerhalb seines Angebots zu ermitteln. Die Analyse dieser Logfiles stellt ein wichtiges neues kommunikationswissenschaftliches Untersuchungsinstrument dar (vgl. Seibold 2002: 84–92). Zum einen ist es nicht-reaktiv, da die Nutzungdaten unbemerkt mitprotokolliert werden. Zum anderen sind die Daten sehr präzise und objektiv, da sie vom Computer nach strengen Regeln und Algorithmen erfasst werden (vgl. Nicholas/Huntington 2000; Seibold 2002: 84–92; Wirth/Brech 1998: 152f.). Die Logfile-Analyse kann schließlich die Lesepfade der einzelnen Rezipienten durch die Webseiten eines Angebots liefern.<sup>3</sup> Die so erhobenen Nutzungsdaten können für die Online-Inhaltsanalyse die Einheiten erzeugen, die auf das inhaltsanalytische Forschungsziel hin untersucht werden. Beispielsweise kann es sinnvoll sein, die »Springe« der Nutzer von der Startseite eines Nachrichtenangebots zu den Seiten mit den Einzelmeldungen als Analyseeinheit aufzufassen. Die Logfile-Analyse hält diesen Sprung als aufeinander folgende Abrufe von Startseite und Berichts-Webseite fest. Wenn man diese Logfile-Daten mit den inhaltsanalytischen Informationen kombiniert, dann kann man zum Beispiel sagen, welche Inhalte der Startseite besonders viel Aufmerksamkeit auf sich zogen und wie dies die Rezeption der nachfolgenden Berichts-Seiten beeinflusst (vgl. Seibold 2002: 107–152).

Damit ergibt sich aus dem Hyperertextcharakter, der non-linearen Rezeption einzelner Inhalts-Bausteine im Internet und den Logfiles eine neuartige Chance: Die Verknüpfung von Inhaltsanalyse mit Nutzungsanalyse.<sup>4</sup> Doch nicht nur die Analyseeinheiten sollten durch ihre Dynamik definiert werden. Dies hilft auch über eine weitere Klippe der Online-Inhaltsanalyse: die schon angeregte zeitliche Dynamik der Inhalte. Sehr viele Online-Angebote sind ein »moving target« (McMillan 2000: 80), sie wechseln sehr schnell die Inhalte oder die Struktur und das Layout, verschwinden gänzlich oder entstehen neu. Diese zeitliche Dynamik der Inhalte kann abgebildet werden, indem man auch das Untersuchungsobjekt und die Analyseeinheiten dynamisch definiert. Die Änderung einer Webseite wird also nicht mehr als »Störung« der Inhaltsanalyse begriffen, sondern als wichtige Komponente. Wie oft beispielsweise eine Nachrichtenseite ihren Aufmacher wechselt, wird zum Teil der »Inhalts«-Analyse. »Die Veränderung würde dann als Variable begriffen«

<sup>3</sup> Eine weitere Möglichkeit, um das Navigationsverhalten von Nutzern zu erfassen, ist die so genannte »client-seitige« Protokollierung über Logfiles oder Videoschnittsnotizen beim Nutzer selbst (vgl. Rössler/Wirth 2001: 286f.). Hier werden im Gegensatz zu den »server-seitigen« Logfiles nicht die Aktionen aller Nutzer in *einem* Angebot untersucht, sondern die Aktionen *eines* Nutzers in *allen* beschrittenen Angeboten. Vorteil des »client-seitigen« Verfahrens ist, dass man mehr Aktionen erfassen kann (zum Beispiel Navigation über Browserelemente), allerdings ist es nicht möglich, für ein spezifisches Online-Angebot alle Nutzungsorgänge zu erfassen. Gerade dies ist jedoch bei der Online-Inhaltsanalyse oft die entscheidende Variable.

<sup>4</sup> Bei anspruchsvollen Forschungsdesigns kann die vorgeschlagene Kombination von Inhalts- und Nutzungs-Analyse zusätzlich erreicht werden durch eine Nutzer-Analyse. Eine solche »Nutzerzentrierung« (vgl. Rössler/Wirth 2001: 286f.) ermöglicht genaue personenbezogene Inhaltsnutzungs-Profilen mit den vollständigen »Selektions-Pfaden« der einzelnen Rezipienten durch das Angebot, allerdings sind Datenschutzrestriktionen zu beachten.

(Rössler 1997: 266). Die Aktualisierungen von Webseiten »on the fly« gehorchen fundamental neuen, anderen Regeln als die reglementierte Periodizität von Zeitung, Zeitschrift, Radio und Fernsehen. Solche oft unmerklichen Änderungen an Online-Inhalten fordern von dem Inhaltsanalytiker rigorose Vergleiche über möglichst kleine Zeitintervalle hinweg, wenn er schriftweise die Genese des Inhalts in Abhängigkeit von der Zeit ermitteln und verstehen will. Gleichzeitig gilt es, jeweils neu die Grenze zu bestimmen, ab der ein Inhalt »neu« im Sinne von »verändert« ist. Gelten zum Beispiel einzelne Kürzungen, Rechtschreibkorrekturen oder neu gefasste Formulierungen in einem Fließtext als Änderung? Gelten für die Überschrift und den Vorspann andere »Neuheits-Regeln«? Je nach Forschungsinteresse muss begründet festgelegt werden, welcher Grad von Inhaltsmodifikationen als Änderung gilt. Auch bei einem solchen »Änderungsmaß« hilft wieder ein Rückgriff auf die Rezeption: So kann per Befragung im Experiment ermittelt werden, was die Leser selbst als »anders« empfinden, und welche Änderungen sie nicht tangieren, weil zum Beispiel wenige Nutzer bis zu dem Absatz lesen, der verändert wurde. Wo der Inhalt fließend wird, erscheint es sinnvoll, die Inhaltsanalyse zur intelligenten Zeitreihenanalyse mit Änderungsmaß zu erweitern.

Noch eine selbsterklärende Kategorie der Inhaltsanalyse löst sich im Internet auf: die Autorenschaft. Zwar scheinen sich die Prophezeiungen aus der Anfangszeit des Webs nicht zu bestätigen: Die Rollen von Journalist und Rezipient fallen im Internet nicht als »gleichberechtigte Kommunikatoren« (Rössler 1998: 26; vgl. auch Burkart 1999: 64) zusammen. Dennoch löst sich im Netz der klare Bezug zwischen Kommunikator und Inhalt oft auf (vgl. Rössler 1997: 248). Erstens wird im Web die Mimikry von fremd zugelieferten Inhalten immer weiter perfektioniert. Dadurch ist sowohl für den Leser als auch für den Forsther die tatsächliche Quelle eines Inhalts kaum mehr bestimmbar. Zweitens verzichtet kaum ein Web-Angebot auf interaktive Elemente wie ein Diskussionsforum oder andere Inhalte wie Gästebuch, Online-Umfragen oder Chatforen. All diesen ist gemeinsam, dass der Rezipient entweder Inhalte beisteuert oder verändert, zum Beispiel durch das Abgeben seiner Stimme in einer Online-Umfrage. Wenn nun ein Rezipient ein solches Element nutzt, dann stammt der Inhalt zum Teil von Kommunikatoren, zum Teil von anderen Rezipienten und zum Teil von ihm selbst. Im Extremfall erzeugt der Online-Rezipient die Struktur der Inhalte sogar komplett selbst (vgl. Rössler/Wirth 2001: 283). Dies ist zum Beispiel bei den häufig angebotenen Suchseiten der Fall. Die Webseite mit den Antworten konstituiert sich erst aus den eingegebenen Suchbegriffen des Rezipienten. Wenn man zum Beispiel bei einer Suchmaschine wie <<http://www.google.com>> die Suchbegriffe »Inhaltsanalyse« und »Publizistik« eingibt, dann wird die Antwortseite individuell auf die Suchanfrage hin erstellt. Ein anderer Nutzer, der die Suchbegriffe »Inhaltsanalyse« und »Medientenor« eingeibt, bekommt eine ganz andere Seite mit ganz anderen Suchergebnissen. Sowohl die Reihenfolge der Suchergebnisse als auch die Inhalte hängen genuin von den Aktionen der Nutzer ab. Diese neue Dynamik der Autorenschaft sollte die Online-Inhaltsanalyse einbeziehen, um auch in Zukunft noch alle Inhalte erfassen und einem oder mehreren Urhebern zuschreiben zu können.

Eine letzte theoretische Besonderheit der Online-Inhaltsanalyse bildet schließlich die oft zitierte Multimedialität des Internets, die sich im Wesentlichen durch die

Verflechtung der Medialitäten Text, Ton und Bild zu immer neuen Kombinationen manifestiert (vgl. Rössler 1998: 35). So kann man im Internet zum Beispiel einen Text über Martin Luther King lesen, ein Photo ansehen und gleichzeitig im Hintergrund seine »I Have a dream«-Rede hören sowie in einem weiteren Fenster ein Video über seinen Marsch auf Washington betrachten. Diese Parallelität verschiedener Medialitäten potenziert die Probleme der Inhaltsanalyse. Schon bei der Erfassung von Fernsehen und Hörfunk mussten komplexe Verfahren entwickelt werden, um die Schwierigkeiten von bewegten Bildern und Ton anzugehen (vgl. Rössler 1997: 247, 249). Im Web sind darüber hinaus die dynamischen Medialitäten Film und Ton mit den statischen Medialitäten Text und Bild in beliebigen Kombinationen verknüpft. Wenn diese wechselseitigen Einflüsse beachtet werden sollen, dann wird die Web-Inhaltsanalyse zum komplexen Forschungsprojekt, das enormen Kodieraufwand und Multimedia-Sachverständstand fordert. Besonders die »Wertigkeiten bei der Verrechnung der einzelnen Elemente stellen ein ungelöstes Problem dar« (Rössler 1997: 249). Aus forschungökonomischen Gründen werden sich Online-Inhaltsanalysen daher in der jetzigen »Pionierphase« sinnvollerweise entweder auf wenige mediale Einheiten oder auf ein kleines Sample beschränken müssen und so schrittweise an einer Methodik der multimedialen Inhaltsanalyse weiter arbeiten.

Zusammenfassend ergeben sich fünf theoretisch-konzeptionelle Herausforderungen der Online-Inhaltsanalyse: Erstens muss das Untersuchungsobjekt im Netz der verlinkten Inhalts-Bausteinen abgegrenzt werden, zweitens legt die Rezeptions-Dynamik eine Verknüpfung der Inhaltsanalyse mit der Nutzungsanalyse nahe, um den dynamischen Inhalt zu erfassen, bei dem sich die Analyseeinheit erst durch den Lesepfad der Nutzer konstituiert. Drittens sollten die zeitlich-fließenden Änderungen des Online-Inhalts als neue Kategorie der Inhaltsanalyse abgebildet werden. Viertens kann die Autorenschaft im Netz oft nur schwer oder in Rückbezug auf den Rezipienten selbst bestimmt werden. Schließlich fordert fünftens die Multimedialität der Inhalte eine integrirende Analyse, welche die Verhältnisse der einzelnen Medialitäten zueinander berücksichtigt.

Im Anschluss an die Diskussion der theoretisch-konzeptionellen Herausforderungen werden nun einige praktische Besonderheiten der Online-Inhaltsanalyse und ihre Lösungsmöglichkeiten dargestellt. Denn selbst wenn alle konzeptionellen Hürden genommen sind, so ergeben sich aus der technisierten Welt des Online-Mediums einige Schwierigkeiten, die gesuchten Informationen tatsächlich einzufangen, zu extrahieren und zu aggregieren. Die folgenden Lösungsansätze sind mithin als weitere Hinweise dafür gedacht, wie die Methode der Inhaltsanalyse in konkreten Untersuchungen auf die neue Online-Welt angewendet werden kann.

#### PRAKТИSCHE HERAUSFORDERUNGEN DER ONLINE-INHALTSANALYSE

Die ständigen Änderungen der Online-Inhalte führen nicht nur zu den schon ange-sprochenen theoretischen Herausforderungen. Da die Webseiten oft unregelmäßig neu erstellt, aktualisiert oder gelöscht werden, ist es auch forschungstechnisch schwie-rig, alle »Versionen« jeder Seite eines Angebots zu erfassen. Dazu gibt es zwei Lö-sungssätze. Entweder man speichert während des gesamten Untersuchungszeitraums

alle Seiten des Untersuchungsobjektes in einem möglichst kleinen Intervall, zum Beispiel alle fünf Minuten, lokal ab und vergleicht dann die Inhalte aller Versionen.<sup>5</sup> Eleganter als dieses aufwändige und nicht exakt zeitpräzise Speicherungsverfahren ist eine Methode, bei der die Inhalts-Daten zum Zeitpunkt des Online-Gehens aufseiten des Anbieters (»serverseitig«) gespeichert werden. Diese Speicherung findet bei vielen Content-Management-Systemen routinemäßig statt. Wenn man als Forscher mit den Anbietern kooperieren kann, ergibt sich somit eine komfortable und exakte Zeitreihe aller aktualisierten Inhalte. Aus dem Speicherdatum der Dateien lässt sich zudem die Online-Zeit automatisch auslesen.

Ein weiteres praktisches Problem der Online-Inhaltsanalyse ist das Ziehen der Stichprobe (vgl. McMillan 2000: 81; Rössler 1997: 251-253, Rössler/Wirth 2001: 288-291). Als forschungspraktische Lösungen haben sich folgende Praktiken bewährt: Erstens sollte man ein festes Zeitfenster bestimmen, um nur solche Inhalte miteinander zu vergleichen, die gleichzeitig online waren (vgl. McMillan 2000: 84f.). Zweitens wird die vollständige Grundgesamtheit, welche die Voraussetzung für eine unverzerrte Stichprobenziehung ist, im Netz zur »unique challenge[s]« (McMillan 2000: 91). Ständig entstehen neue Angebote, die unter Umständen noch nirgends erfasst sind, und alte Seiten verschwinden, die fälschlicherweise zur Grundgesamtheit gezählt werden. Um möglichst alle relevanten Angebote zu erfassen, sollte man Suchmaschinen, Web-Kataloge, Verzeichnisse und externe Quellen auswertern (vgl. McMillan 2000: 91f., Rössler 1997: 251-253).

Wenn man schließlich eine Stichprobe gezogen oder eine Vollerhebung über einen bestimmten Zeitraum realisiert hat, und die Webseiten lokal abgespeichert wurden, bietet sich eine weiterhin computergestützte Analyse des ohnehin schon digitalisierten Materials an. Dazu müssen in einem nächstem Schritt die inhaltsanalytisch relevanten Daten aus dem HTML-Code extrahiert werden. Dies wird oft als der arbeitsintensive Teil der Web-Analyse bezeichnet (vgl. Klein 1999: 2), da die einzelnen Inhalte in einer Vielzahl von Steuerbefehlen eingebunden sind wie <B class=sz-fließtxt1>Fischer: <A class=sz-fließtxt1 href="http://www.sueddeutsche.de/nachrichten/woch46/fischer.htm">Neue Pläne gegen BSE-Gefahr</A>. Solange die Inhalte in solchen Steuerzeichen des Programmiercodes stecken, können sie oft weder sinnvoll weiterverarbeitet noch kodiert werden. Bei kleinen Datensmengen kann man die Inhalte durch manuelles »Copy and Paste« aus den Webseiten herausziehen oder einfach ablesen. Typischerweise liegen jedoch eine große Menge von Webseiten vor, aus denen Texte, Bilder und andere Elemente extrahiert werden sollen. Dabei helfen »Parser«-Programme (vgl. Seibold 2001: Anhang). Diese werden an die Struktur des HTML-Codes der jeweiligen Seiten angepasst und extrahieren dann automatisch die inhaltlichen Elemente, die in einer Datenbank gesammelt werden. Beispielsweise können dadurch alle Artikel eines journalistischen Online-Dienstes so abgespeichert werden, dass automatisch je eine Spalte für die Überschrift, den Vorspann und den Artikeltext gebildet werden. Die so aufbereiteten Daten sind problemlos mit statistischen Auswertungsprogrammen wie SPSS (Statistical Package of the Social Sciences) weiter analysierbar.

#### ZUSAMMENFASSUNG UND AUSBlick

Als Zusammenfassung soll das Schaubild dienen, welches die zuvor dargestellten Herausforderungen der Inhaltsanalyse im Internet tabellarisch abbildet und die jeweiligen Lösungsmöglichkeiten aufzeigt.

Das Internet wird in der Zukunft immer mehr Bereiche des Lebens und der menschlichen Informationsaufnahme beeinflussen. Unso mehr wird auch das Interesse an wissenschaftlichen Befunden zur Online-Welt wachsen. Die kommunikationswissenschaftliche Inhaltsanalyse kann eine zentrale Rolle in dieser Reflexion des Internets spielen. Im Vorhergehenden wurde gezeigt, dass dazu in vielen Fällen ihr Instrumenten-

als letztes praktisches Problem soll diskutiert werden, dass durch die Änderungen der Webseiten über die Zeit hinweg oft viele »Versionen« einer Webseite vorliegen, die sich nur marginal unterscheiden. Zum Beispiel produziert der Nachrichtendienst <sueddeutsche.de> pro Tag etwa 140 verschiedene Versionen seiner Startseite mit insgesamt über 5.600 Kurzmeldungen. Wenn nun bei der Startseite mit zehn Meldungen nur eine ausgetauscht wurde, dann bleibt der Inhalt aller anderen Meldungen gleich, wird jedoch doppelt oder öfter als »neue Version« abgespeichert. Zu erfassen ist also jeweils nur, welche Einheiten pro »Version« einer Seite neu hinzukommen und welche die Position oder den Inhalt wechseln oder welche gelöscht werden. Das Rohmaterial der einzelnen Zeitreihen muss daher so analysiert werden, dass für jedes Element einer Webseite die Lebensdauer als Online-Zeit und der Rangplatz als Position innerhalb der Seite festgelegt werden. Wenn zum Beispiel die Analyseeinheit eine Meldung ist, dann muss diese wie folgt beschreibbar sein: Die Meldung »Wirtschaft lahmt« bildete von 10:48 Uhr bis 12:23 Uhr den Aufmacher der Seite xyz, wurde dann zur zweiten Notiz im Wirtschaftsressort und schließlich um 15:34 Uhr durch eine Meldung über Aktienverluste ersetzt. Um solche relevanten Daten aus der redundanten Datenflut über die Zeit hinweg zu extrahieren, benötigt man eine aufwändige datentechnische Analyse des Rohmaterials in einer relationalen Datenbank (vgl. Seibold 2001: Anhang). Dabei werden durch Datenbanktransformationen jeweils alle Analyseeinheiten aggregiert, deren Eigenschaften im Sinne der jeweiligen Untersuchung gleich sind. Somit werden in ihrer zeitlichen Abfolge nur die Elemente neu abgespeichert, die sich tatsächlich geändert haben (Neuerschein, Positionswechsel, Löschung).

Zusammenfassend ergeben sich vier praktische Herausforderungen der Online-Inhaltsanalyse: Um erstens alle Versionen eines Online-Angebots zu erfassen, müssen die Daten möglichst lückenlos lokal oder serverseitig abgespeichert werden. Zweitens erfordert das Ziehen einer Stichprobe im Internet, dass die Grundgesamtheit auf ein Zeitfenster eingeengt wird und durch Suchmaschinen, Kataloge und externe Quellen möglichst vollständig erhoben wird. Drittens empfiehlt es sich, die relevanten inhaltsanalytischen Daten manuell oder durch ein Parser-Programm aus dem Programmier-Code der Webseiten in eine Datenbank zu extrahieren. Schließlich müssen viertens die relevanten Inhalte aus der Datenflut der einzelnen Versionen einer Webseite über die Zeit hinweg aggregiert werden. Dies kann eine Datenbank-Analyse des Rohmaterials leisten, die nur geänderte Elemente einer Seite als neu abspeichert.

<sup>5</sup> Zum Vergleich der einzelnen Seiten-Versionen eignen sich computerunterstützte Analysen, da Textverarbeitungsprogramme mittels Ähnlichkeitsermittlung relativ leicht eine große Anzahl von Dokumenten auf (Un-)Gleichheit untersuchen können.

**Herausforderungen und Lösungsmöglichkeiten der Online-Inhaltsanalyse****Schaubild****theoretische Herausforderungen****Lösungsmöglichkeiten**

(1) Untersuchungsobjekte im Hypertext abgrenzen.

(2) Analyseeinheit bestimmen, obwohl Inhalts-Menü erst durch die Klickpfade und Voreinstellungen der Leser konstituiert wird.

(3) Zeitliche Dynamik der Online-Inhalte abbilden (neue, veränderte und gelöschte Inhalte erfassen).

(4) Wechselnde Autorenschaft bestimmen.

(5) Multimediale Inhalte und ihre wechselseitigen Einflüsse erfassen.

**praktische Herausforderungen****Lösungsmöglichkeiten**

(1) Alle Versionen eines Angebots erfassen.

(a) Alle Daten immer wieder lokal abspeichern und vergleichen oder  
(b) Daten zum Zeitpunkt des Online-Gehens serverseitig speichern.

(2) Stichprobe aus der vollständigen Grundgesamtheit ziehen.

(3) Inhaltsanalytisch relevante Daten aus HTML-Code extrahieren.

(4) Relevante Daten aus Datenflut der verschiedenen Versionen einer Webseite aggregieren.

**LITERATUR**

- Bensberg, Frank/Weiß, Thorsten (1998): Web Log Mining als Analyseinstrument des Electronic Commerce. In: Uhr, Wolfgang (Hrsg.): Integration externer Informationen in management support systems. Wirtschaftsinformatik-Fachtagung „Integration externer Informationen“, Dresden, S. 197-214. Auch als elektronische Publikation abrufbar.  
URL: <<http://www.wi.uni-muenster.de/aw/mitarbei/awfirth/wlm.pdfs>>. Stand: 26.9.2001.
- Brosius, Hans-Bernd/Koschel, Friederike (2001): Methoden der empirischen Kommunikationsforschung. Eine Einführung. Wiesbaden.
- Bucher, Hans-Jürgen (2000): Publizistische Qualität im Internet. Rezeptionsforschung für die Praxis. In: Altmeppen, Klaus-Dieter/Bucher, Hans-Jürgen/Löffelholz, Martin (Hrsg.): Online-Journalismus. Perspektiven für Wissenschaft und Praxis. Wiesbaden, S. 153-172.
- Bucher, Hans-Jürgen/Barth, Christof (1998): Rezeptionsmuster der Onlinekommunikation. Empirische Studie zur Nutzung der Internetangebote von Rundfunkanstalten und Zeitungen. In: Media Perspektiven, Heft 10, S. 517-523.
- Burkart, Roland (1999): Was ist eigentlich ein Medium? Überlegungen zu einem kommunikationswissenschaftlichen Medienbegriff angesichts der Konvergenzdebatte. In: Latzer, Michael/Maier-Rabler, Ursula/Sieger, Gabriele/Steinmauer, Thomas (Hrsg.): Die Zukunft der Kommunikation: Phänomene und Trends in der Informationsgesellschaft. Innsbruck, Wien, S. 61-71.
- Eimeier, Birgit van/Gerhard, Heinz/Frees, Beate (2001): ARD/ZDF-Online-Studie 2000: Internetnutzung stark zweckgebunden. In: Media Perspektiven, Heft 8, S. 382-397.
- Evans, William (1998): Content analysis in an era of interactive news: assessing 21st century symbolic environments. In: Borden, Diane/Harvey, Kerrie (Hrsg.): The electronic grapevine. Rumor, reputation, and reporting in the new on-line environment. Mahwah, London, S. 161-171.
- Focus Magazin Verlag (Hrsg.) (2001): Der Markt der Online-Kommunikation. Daten, Fakten, Trends. München. Auch als elektronische Publikation abrufbar. URL: <<http://www.medialine.de/marktanalysen>>. Stand: 26.9.2001.
- Friih, Werner (1998): Inhaltsanalyse. Theorie und Praxis. Konstanz.
- Janerko, Dietmar (1999): Surfer im Visier. Beobachten, Befragen, Belauschen – »Netizens« werden zum Lieblingsopfer von Markt- und Sozialforschern. In: ct. Magazin für Computertechnik, Heft 20, S. 86-92.
- Klein, Harald (1999): Inhaltsanalyse von Webseiten: Probleme und Lösungsansätze. Elektronische Publication. URL: <[http://dgof.de/derband99/pdf/si\\_p/klein.pdfs](http://dgof.de/derband99/pdf/si_p/klein.pdfs)>. Aus: Reips, Ulf/Dietrich/Batting, Bernd/Bandilla, Wolfgang u.a. (Hrsg.): Current Internet science – trends, techniques, results. Aktuelle OnlineForschung – Trends, Techniken, Ergebnisse. Zürich. Elektronische Publication. URL des Sammelbands: <<http://dgof.de/band99/>>. Stand: 19.2.2001.
- McMillan, Sally (2000): The microscope and the moving target: The challenge of applying content analysis to the world wide web. In: Journalism and Mass Communication Quarterly, 77. Jg., S. 80-98.
- Merten, Klaus (1995): Inhaltsanalyse. Einführung in Theorie, Methode und Praxis. Opladen.
- Nicholas, David/Huntington, Paul (2000): Evaluating the use of newspaper web sites logs. In: JMM – The International Journal on Media Management, 2. Jg., Heft 2, S. 78-88. Auch als elektronische Publikation abrufbar.  
URL: <[http://www.mediajournal.org/heracademy/publications.nsf/all\\_pk1785](http://www.mediajournal.org/heracademy/publications.nsf/all_pk1785)>. Stand: 26.9.2001.
- Riefler, Katja (1999): The Daily Me – Personalisierte Informationsdienste im Internet. In: Bundesverband deutscher Zeitungsvorleger e.V. (Hrsg.): Zeitungen '99. Bonn, S. 196-204.
- Rössler, Patrick (1997): Standardisierte Inhaltsanalysen im WorldWideWeb. Überlegungen zur Anwendung der Methode am Beispiel einer Studie zu Online-Shopping-Angeboten. In: Beck, Klaus/Vowe, Gerhard (Hrsg.): Computernetze – ein Medium öffentlicher Kommunikation? Berlin, S. 245-267.
- Rössler, Patrick (1998): Wirkungsmodelle: die digitale Herausforderung. Überlegungen zu einer Inventur bestehender Erklärungsansätze der Medienwirkungsforschung. In: Rössler, Patrick (Hrsg.): Online-Kommunikation. Beiträge zu Nutzung und Wirkung. Opladen, Wiesbaden, S. 17-46.

- Rössler, Patrick/Wirth, Werner (2001): Inhaltsanalysen im World Wide Web. Probleme und Perspektiven. In: Wirth, Werner/Lauf, Edmund (Hrsg.): Inhaltsanalyse: Perspektiven, Probleme, Potentiale. Köln, S. 280-302.
- Schweiger, Wolfgang (2001): Hypermedien im Internet. Nutzung und ausgewählte Effekte der Linkgestaltung (= Reihe InternetResearch, Bd. 3). München.
- Seibold, Balthas (2001): Klick-Magnete. Eine Analyse von Aufmerksamkeitsfaktoren bei journalistischen Online-Angeboten. Diplomarbeit am Institut für Kommunikationswissenschaft, Univ. München (enthält Anhangband mit Dokumentation der Logfile-Analyse, der in der unten aufgeführten Publikation der Arbeit nicht enthalten ist).
- Seibold, Balthas (2002): Klick-Magnete. Eine Analyse von Aufmerksamkeitsfaktoren bei journalistischen Online-Angeboten (= Reihe InternetResearch, Bd. 4). München.
- Stempel, Guido/Stewart, Robert (2000): The internet provides both opportunities and challenges for mass communication researchers. In: Journalism and Mass Communication Quarterly, 77. Jg., S. 541-548.
- Werner, Andreas (1998): Online-Medien: Theoriebereicherung durch neue Forschungsmethoden. In: Rössler, Patrick (Hrsg.): Online-Kommunikation. Beiträge zu Nutzung und Wirkung. Opladen, Wiesbaden, S. 227-241.
- Wirth, Werner/Brechth, Michael (1998): Medial und personal induzierte Selektionsentscheidungen bei der Nutzung des World Wide Web. In: Rössler, Patrick (Hrsg.): Online-Kommunikation. Beiträge zu Nutzung und Wirkung. Opladen, Wiesbaden, S. 147-168.

*Korrespondenzanschrift:* Balthas Seibold, Gundelindenstraße 6, D-80805 München  
E-Mail: seibold@baithas.de